Robert A. Blecker
Mark Setterfield

Elgar

# HETERODOX MACROECONOMICS

## Models of Demand, Distribution and Growth

# Heterodox Macroeconomics

# Heterodox Macroeconomics

## MODELS OF DEMAND, DISTRIBUTION AND GROWTH

Robert A. Blecker

*Professor of Economics, American University, Washington, DC, USA*

Mark Setterfield

*Professor of Economics, The New School for Social Research, New York, NY, USA*

Edward Elgar PUBLISHING

Cheltenham, UK • Northampton, MA, USA

Published by
Edward Elgar Publishing Limited
The Lypiatts
15 Lansdown Road
Cheltenham
Glos GL50 2JA
UK

Edward Elgar Publishing, Inc.
William Pratt House
9 Dewey Court
Northampton
Massachusetts 01060
USA

A catalogue record for this book
is available from the British Library

Library of Congress Control Number: 2019950117

ISBN 978 1 78471 889 3 (cased)
ISBN 978 1 78471 891 6 (paperback)
ISBN 978 1 78471 890 9 (eBook)

Typeset by Servis Filmsetting Ltd, Stockport, Cheshire

# Contents in brief

# Full contents

# Preface and acknowledgements

It is no secret that macroeconomics is in disarray – or at least it should be, given its astonishing and very public failures in recent decades. As the mainstream of the profession has sat by, the majority of working households in advanced capitalist economies have experienced over three decades of real income stagnation that has fuelled burgeoning inequality and gone hand in hand with increased economic insecurity. The *coup de grâce* was delivered a decade ago by the financial crisis and subsequent Great Recession, at which point even Queen Elizabeth II was wont to ask the Royal Economic Society if they knew what they were doing. This has been followed by a long but weak recovery that has done nothing to alter the trends towards ever-increasing income inequality and economic insecurity that preceded the Great Recession and, as a result, has raised the spectre of long-term stagnation. These developments are now commonly linked to social phenomena as various and disturbing as the groundswell of support for far-right populist politicians and falling male life expectancy associated with 'deaths of despair'. The fact that macroeconomics *isn't* in disarray – and that we haven't witnessed a revolution in theory on a par with that in the 1930s in the aftermath of the Great Depression – probably says as much about power and control in the economics profession (and its capacity to protect certain highly ingrained habits of thinking) as it does about the usefulness or applicability of mainstream macroeconomic theory itself.

But even as mainstream thinking has survived these events (and even prospered) by 'patching up' its models with ad hoc extensions on the basis of 20/20 hindsight, and even as this process introduces new tensions (and even contradictions) into mainstream macro analysis (Rogers, 2018a, 2018b), one inescapable fact remains: mainstream thinking is set up to describe capitalism as an innately regular, ordered and tranquil environment, which is rendered disorderly only by the sudden and unexpected imposition of external events ('exogenous shocks'). In other words, mainstream macroeconomics does not envisage a system that, while fairly orderly much of the time, and certainly *resilient* (Holling, 1973), is nevertheless very much capable of *internally* generating tensions, conflict and disruptions. In short, for want of an appropriate pre-analytic vision (Heilbroner and Milberg, 1996), a meaningful sense of what capitalism as an object of analysis actually is (rather than what an idealistic fantasy would like it to be), mainstream macroeconomics lacks (and continues to display little interest in developing) a theory of capitalism as a stratified and contested terrain that is vulnerable to periodic crises. Theorizing that takes these themes seriously, as part and parcel of the ordinary experience of capitalist economies and their functioning, is the wheelhouse of heterodox macrodynamics. It is for this reason that the contents of this book, which is devoted overwhelmingly to the promulgation of heterodox growth theory, is both timely and important.

To be clear, the book is not a treatise on capitalism as a whole, but rather a text that is much more focused on heterodox macrodynamic theory – the core tools that macroeconomists outside the mainstream have developed and used to analyse the motion of capitalism as it grows and transforms over time. Even then, its focus is somewhat narrower than 'heterodox macrodynamics' as a whole. For instance, we focus mainly on 'real' rather than 'nominal' dynamics, associated with the longer-term growth of output and employment. As such, topics like inflation, or money and finance, enter only in a supporting role. Nor do we attempt to discuss every facet of the short-run macroeconomics that is the focus of attention when explaining the booms and busts associated with the business cycle. Instead, we discuss only those theories of short-period fluctuations in activity that are an outgrowth of (and accompaniment to) the long-run theories of growth that are our principal concern. Finally, although macroeconomic policy frequently enters the discussion in the chapters that follow, it is not our main focus in this book and we do not draw out the policy implications of every facet of what follows.

That said, what the book *does* provide is extensive coverage of heterodox growth theories associated, in particular, with the classical-Marxian and post-Keynesian traditions. We begin by explicitly comparing and contrasting these traditions with each other and with the neoclassical mainstream approach. The book is then divided into three parts, which cover (successively): core models of growth and distribution associated with the classical-Marxian and post-Keynesian traditions; extensions to these core models of growth and distribution that introduce (among other things) more nuanced discussion of cyclical growth dynamics, interactions between the real and financial sectors, and broader and more inclusive conceptions of social stratification (including, for example, gender as well as social class); and models of export-led growth in which the focus shifts away from domestic income distribution towards the interaction of trade and growth with structural and technical change (although this last set of models is, by its nature, concerned with convergence or divergence of income levels among different nations).

The analysis in each part of the book follows what might be termed a 'foundational' approach. Modelling traditions and their 'pre-analytic visions' of capitalism are carefully introduced and models are built from scratch in a manner that does not assume extensive prior knowledge of either traditions or specific models. The 'inner workings' of all models are fully explained and their outcomes and implications are discussed so that the reader will come to understand both how the models work and what they tell us about the macrodynamics of a capitalist economy. This foundational approach, together with the broad coverage given to competing models and traditions within heterodox macrodynamics, means that even as some of the topics addressed become quite advanced, no attempt is made to cover every 'frontier' issue or contribution to the literature. The myriad world of extensions and further applications is left for further study on the part of the interested reader. That said, however, the chief virtue of our foundational approach is that the book should prove useful to students of macrodynamics at many levels, from advanced undergraduates through students beginning their graduate studies to members of the profession. In terms of mathematical tools, the book assumes familiarity with basic algebra and multivariate calculus, together with some elements of linear algebra and the analysis of simultaneous differential equations.

In the preparation of certain chapters we have drawn on our previously published work in the following sources, without reproducing any of them in full or in detail:

Chapter 1: Setterfield, M. (2014), 'Neoclassical growth theory and heterodox growth theory: opportunities for (and obstacles to) greater engagement', *Eastern Economic Journal*, 40 (3), 365–86.

Chapter 5: Blecker, R.A. (2016), 'Wage-led versus profit-led demand regimes: the long and the short of it', *Review of Keynesian Economics*, 4 (4), 373–90.

Chapter 7: Blecker, R. and S. Seguino (2002), 'Macroeconomic effects of reducing gender wage inequality in an export-oriented, semi-industrialized economy', *Review of Development Economics*, 6 (1), 103–19.

Chapter 8: Setterfield, M. (2013), 'Endogenous growth: a Kaldorian approach', in G.C. Harcourt and P. Kriesler (eds), *The Oxford Handbook of Post-Keynesian Economics*, vol. I, *Theory and Origins*. Oxford, UK: Oxford University Press, pp. 231–56.

Chapter 9: Blecker, R.A. (2013), 'Long-run growth in open economies: export-led cumulative causation or a balance-of-payments constraint?' in G.C. Harcourt and Peter Kriesler (eds), *The Oxford Handbook of Post-Keynesian Economics*, vol. I, *Theory and Origins*. Oxford, UK: Oxford University Press, pp. 390–414.

Blecker, R.A. and C.A. Ibarra (2013), 'Trade liberalization and the balance of payments constraint with intermediate imports: the case of Mexico revisited', *Structural Change and Economic Dynamics*, 25 (June), 33–47.

Ibarra, C.A. and R.A. Blecker (2016), 'Structural change, the real exchange rate and the balance of payments in Mexico, 1960–2012', *Cambridge Journal of Economics*, 40 (2), 507–39.

Chapter 10: Blecker, R.A. (2016), 'The debate over "Thirlwall's law": balance-of-payments-constrained growth reconsidered', *European Journal of Economics and Economic Policy: Intervention*, 13 (3), 275–90.

We would like to thank the various publishers of these articles and book chapters, together with our co-authors, for their permission to make use of these materials.

Needless to say, the book would not have been completed without the help of many friends, colleagues and students, to all of whom we owe a debt of thanks. Inevitably, the book has benefited extensively from discussions with many of our colleagues in the profession over the course of many years. Special thanks are due to those who made specific comments or suggestions on drafts of our manuscript or presentations of this work, including Mike Cauvel, Steve Fazzari, Carlos Ibarra, Alan Isaac, Gilberto Tadeu Lima, John McCombie, Juan Carlos Moreno-Brid, Torsten Niechoj, Michalis Nikiforos, Özlem Onaran, José Luis Oreiro, Esteban Pérez Caldentey, Gabriel Porcile, Arslan Razmi, Rafael Ribeiro, Jaime Ros, Danilo Spinola, Engelbert Stockhammer, Daniele Tavani, Ramaa Vasudevan and Randy Wray. However, none of them bears any responsibility for any remaining errors or omissions in the book; such responsibility lies solely with the authors.

We must also mention our many dedicated students, who have been exposed to the real-time development of the material set out in this book in courses we have taught at American University, Washington, DC, Trinity College, Hartford, CT, and The New School for Social Research, New York, NY, over the past few decades. Many of these students, too numerous to mention, have provided invaluable feedback by way of the questions they have raised or, in some cases, direct comments and suggestions on earlier drafts of individual chapters in the past four years. Especially, we would like to single out AU PhD student Russell Rollow for a special 'thank you' for his sterling work in compiling and editing the references. Without his assistance, our final preparation of the original manuscript would have been much more difficult and delayed.

Thanks are also due to Alan Sturmer and other members of staff at Edward Elgar Publishing for their support and guidance (not to mention patience, given our repeated passing of deadlines and reconfiguring of due dates) and to our wives, Elizabeth Greenberg and Jennifer MacDonald, for putting up with our distractions and obsessions during the process!

Robert A. Blecker
Washington, DC

Mark Setterfield
New York, NY

February 2019

# Notation guide

## Roman alphabet*

| | | | |
|---|---|---|---|
| $A$ | Technological shift parameter (Hicks-neutral) *or* index of labour efficiency/labour-augmenting technology (1)<br>Constant term in 'Cobb–Douglas' investment function (4)<br>Autonomous component of aggregate demand (6–9) | $a$ | Constant output–capital ratio (productivity of capital) in '$aK$' model (1) |
| $^{A}$ | Superscript, country A (8–10) | | |
| $_{A}$ | Subscript for domestic autonomous demand (7–10) | $a_0$ | Labour–output ratio = $L/Y$ (all) |
| $AD$ | Aggregate demand function (5) | $a_1$ | Capital to full-capacity output ratio = $K/Y_K$ (all) |
| AD | Aggregate demand relationship (5, 7) | $a_1^i = K^i/Y_n^i$ | Capital–output ratio of firm $i$ at normal utilization (4) |
| $ATC$, $AVC$ | Average total cost, average variable cost (4) | $a_2$ | Managers' labour–output ratio (7) |
| | | $a_H$, $a_X$ | Labour–output ratios in home and export goods sectors (7) |
| | | $a_{Ji}$ | Stability conditions in a three-dimensional system of differential equations (7) |
| | | $a_1^f$ | Foreign capital–output ratio, Ros model (10) |
| | | $a^Y$, $a^K$ | Ratios, $a^Y = A/Y$, $a^K = A/K$ (6–7) |
| $B$ | Stock of bonds (debt of firms) (7) | $b$ | Trade balance ratio (ratio of net exports to capital stock) (4) |
| $^{B}$ | Superscript, country B (8–10) | $b_i$ | Partial derivatives of trade balance function (4)<br>Parameters in equation for full-capacity output growth, $\hat{Y}_K$ (5)<br>Parameters in Oreiro's $\eta_M$ function (10) |

| | | | |
|---|---|---|---|
| BP | Balance of payments | $b_{ij}$ | Coefficients in linear versions of the equations for Kaldor's growth laws ($i$ = law, $j$ = coefficient) (8) |
| | | $b_J$ | Stability condition in a three-dimensional system of differential equations (7) |
| $C$ | Consumption (all) | $c$ | Consumption per employed worker = $C/L$ (2–4)<br>Marginal propensity to consume (8) |
| | | $c_r$ | Consumption out of profits, per worker (2) |
| $C_R$ | Total consumption of rentiers (7) | $c_R$ | Marginal propensity to consume of rentiers (7) |
| $C^T$ | Workers' target level of consumption (7) | | |
| $C_w$ | Total consumption of workers (7) | $c_w$ | Consumption out of wages, per worker (2)<br>Marginal propensity to consume of workers (7) |
| | | $\bar{c}$ | Exogenously given constant (1, 6) |
| | | $_c$ | (Subscript) Capitalists in Pasinetti model (3)<br>Final goods (consumption + capital goods) (9)<br>Consumption goods only (10) |
| $D$ | Total debt owed by workers (7) | $d$ | Debt–capital ratio (6) |
| $\dot{D}$ | Borrowing by workers (increase in their debt) (7) | $d_B$ | Firms' debt (corporate bonds) relative to capital = $B/PK$ (7) |
| $D_R$ | Debt owed by workers to rentier households (7) | $d_i$ | Parameters in equation for output growth, $\hat{Y}$ (5) |
| | | $d_R$ | Debt owed to rentier households relative to capital stock (7) |
| DC | Distributive curve (5) | | |
| DR | Demand regime (8–10) | | |
| $E$ | Nominal exchange rate (home currency/foreign currency) (4, 7–10) | $e$ | Employment rate = $L/N = Y/Y_N$ (all)<br>Also used as the base of the natural logarithm, where noted (all) |

| | | | |
|---|---|---|---|
| $\hat{E}$ | Rate of nominal exchange rate depreciation (8–10) | $^{e}$ | (Superscript) Expected variable (all) |
| $E^D$ | Endogenous components of aggregate demand = $C + I$ (6) | | |
| $F$ | Neoclassical production function, $F(K,L)$ (all) | $f$ | Intensive form of neoclassical production function, $f(k)$ (2)<br>Robinson's investment function (3, 6)<br>Financial fragility variable in Stockhammer–Michell model (7) |
| $FC$ | Fixed cost (4) | $f_i$ | Parameters in linearized Robinsonian investment function (3, 6–7) |
| | | $f_R$ | 'The fraction' in Freitas–Serrano model (7) |
| | | $_f$ | (Subscript) Firms (5, 6)<br>Foreign (4, 7–10)<br>Except female for labour and wages (7) |
| $G$ | Government spending (all) | $g$ | Growth rate of capital stock or rate of capital accumulation = $I/K = \hat{K}$ or $\Delta K/K$ (all) |
| $Gap$, $RelGap$ | Income gaps, absolute and relative (8) | $g_A$ | Growth rate of autonomous demand (6–9) |
| | | $g_c$ | Growth rate of capitalists' capital, Pasinetti model (3) |
| $GPM^i$ | Gross profit margin of firm $i$ (4) | $g^d$ | Desired accumulation (growth) rate in Marglin's neo-Marxian/ neo-Keynesian synthesis model (3) |
| | | $g_i$ | Parameters in the Kalecki–Steindl investment function (4, 7) |
| | | $g_w$ | Growth rate of workers' capital, Pasinetti model (3) |
| $H$ | Home goods in Blecker–Seguino gender model (7) | $h$ | Harrodian growth function (implicit) (3)<br>Bhaduri–Marglin investment function (implicit) (4)<br>Investment-to-output ratio in Freitas–Serrano model (7) |
| $_H$ | (Subscript) Heterodox in $Z_H$ (1)<br>Home goods (7) | $h_i$ | Partial derivatives of the Bhaduri–Marglin investment function (4)<br>Coefficients in linearized versions of the same function (4, 7) |

| | | | |
|---|---|---|---|
| $I$ | Investment (net, same as gross assuming no depreciation) = $\Delta K$ or $\dot{K}$ (all) | $i$ | Interest rate (all) |
| $I_c, I_w$ | Investment funded by capitalists' and workers' savings, Pasinetti model (3) | $i^*$ | Risk-free interest rate (6) |
| | | $_i$ | (Subscript) Intermediate goods (9–10) |
| **J** | Jacobian matrix | $j$ | Ratio of raw materials costs to labour costs (4) |
| $J$ | Any exogenous variable (5) | | |
| $_J$ | Subscript for variable $J$ in $Z_J$ (5) | $_j$ | (Subscript) Good or industry (9–10) |
| $K$ | Capital stock | $k$ | Capital–labour ratio = $K/L$ (1–5)<br>Profit share function (6) |
| $K^i$ | Capital of firm $i$ (4) | | |
| $K_c, K_w$ | Capital owned by capitalists and workers, Pasinetti model (3) | $k_A$ | Domestic autonomous expenditures multiplier (9–10) |
| $K_u$ | Capital utilized to produce current output = $uK$ (3, 6) | $k_X$ | Export multiplier (8–10) |
| $L$ | Employment or labour demand (all) | $l$ | Growth rate of employment or labour demand = $\hat{L}$ or $\Delta L/L$ (all) |
| | | $l_m, l_n$ | Growth rates of employment in manufacturing and non-manufacturing sectors (8) |
| $L_0, L_1$ | Employment of production workers and overhead labour, respectively (Appendix 4.1) | | |
| $M$ | Money supply (5)<br>Managerial labour (7)<br>Imports (quantity) (8–10) | $m$ | Growth rate of imports = $\hat{M}$ (8–10) |
| $_M$ | (Subscript) Managers (7)<br>Imports (8–10) | $_m$ | (Subscript) Male labour and wages (7)<br>Manufacturing (8–10) |
| $M_0$ | Constant in import demand function (9–10) | | |
| $MC$ | Marginal cost (4) | | |
| $N$ | Labour force or labour supply (all)<br>Normal distribution (1) | $n$ | Growth rate of labour force or supply = $\hat{N}$ or $\Delta N/N$ (all) |

| | | | |
|---|---|---|---|
| $_N$ | (Subscript) Natural rate of growth (of output or capital) (all)<br>Except also used for neoclassical in $Z_N$ (1) | $^n$ | (Superscript) The North in a North–South trade model (10) |
| $\overline{N}$ | Total number of goods in multisectoral models (9–10) | $_n$ | (Subscript) Normal, for the utilization rate, $u_n$ (all)<br>Non-manufacturing output or employment (8) |
| $N^s$ | Number of goods produced by the South (10) | | |
| $NUC^i$ | Normal unit cost of firm *i* (4) | $ncm^i$ | Net costing margin of firm *i* (4) |
| | | $n_i$ | Coefficients in labour supply growth function (2)<br>Imported raw materials coefficient in sector *i* (7) |
| *NCF* | Net capital inflows (9–10) | *ncf* | Growth rate of net capital inflows (9–10) |
| *NX* | Net exports or trade balance (5, 9–10) | | |
| | | $_o$ | (Subscript) Other goods (primary commodities) (9) |
| *P* | Price level (domestic) (all) | | |
| *Pop* | Population (1) | | |
| $P^i$ | Price of firm *i* (4) | | |
| $\hat{p}$ | Inflation rate (all) | | |
| $P_i$ | Price of good *i*:<br>$i = X, H$ (export and home goods) and $i = n$ (imported raw materials) (7)<br>$i = d$ (domestic), *f* (foreign) *M* (imports), *o* (other exports), *m* manufactures, *X* (exports) (8–10) | | |
| $\hat{P}_i$ | Rate of increase in price $P_i$ (9–10) | | |
| PR | Productivity regime (8–10) | | |
| *Q* | Labour productivity = $Y/L$ (all)<br>Note: this is the same as $1/a_0$ everywhere *except* Appendix 4.1, where there is also overhead labour | *q* | Growth rate of labour productivity = $\hat{Q} = -\hat{a}_0$ (all) |

| | | | |
|---|---|---|---|
| $Q_i$ | Labour productivity in sector *i* (8) | $q_i$ | Coefficients in productivity growth equation (5)<br>Growth rate of labour productivity in sector *i* (8) |
| | | $q_0$ | Constant term in Verdoorn's law equation (8–10) |
| $R$ | | $r$ | Profit rate (all) |
| | | $r_c$, $r_w$ | Rates of profit earned by capitalists and workers, Pasinetti model (3) |
| | | $r_n$ | Normal rate of profit, in target-return pricing model (4) |
| $S$ | Total savings (all) | $s$ | Marginal propensity to save (all) |
| | | $^s$ | (Superscript) The South in a North–South trade model (10) |
| | | $_s$ | (Subscript) Modern service sector (8) |
| | | $s_c$, $s_h$ | Marginal propensities to save of corporations and households (3) |
| | | $s_f$ | Foreign saving propensity, Ros model (10) |
| | | $s_L$, $s_K$ | Marginal propensities to save of production worker and capitalist-manager households, Palley model (7) |
| | | $s_M$ | Marginal propensity to save of managers, Tavani-Vasudevan model (7) |
| | | $s_R$ | Marginal propensity to save of rentiers (7) |
| | | $s_r$ | Marginal propensity to save out of profit income (1–5, 7) |
| $S_w$ | Saving of workers (7) | $s_w$ | Marginal propensity to save out of wage income (3–5, 7) |
| $T$ | Periodicity of the original Goodwin cycle (2) | $t$ | Time (all) |
| | | $_t$ | (Subscript) Time (all)<br>Except – traditional sector (8) |

| | | | |
|---|---|---|---|
| $U$ | Unemployment rate = $1 - e = (N - L)/N$ (1) | $u$ | Capacity utilization rate = $Y/Y_K$ (all) |
| | | $_u$ | (Subscript) Utilized, for the capital stock $K_u$ (3, 6) |
| | | $u_n$ | Normal rate of capacity utilization (all) |
| $ULC$ | Unit labour costs (all) | $uc$ | Unit cost in the cost function dual to the neoclassical production function (2) |
| $V$ | Velocity of money (5) | $v$ | Marginal propensity to import (8) |
| $W$ | Nominal wage rate (all) | $w$ | Real wage rate = $W/P$ (all) |
| $W_i$ | Nominal wage rate of worker type $i$ – could be production workers, managers, male or female, etc. (7) | $w_s$ | Malthusian subsistence wage (2) |
| $\hat{W}$ | Wage inflation rate (all) | $\overline{w}$ | Exogenously given real wage, classical-Marxian and Marglin synthesis models (2–3)<br>Constant/exogenous rate of nominal wage increase, short run of Ribeiro et al. model (10) |
| | | $_w$ | (Subscript) Warranted (in warranted rate of growth, $y_w$), in Harrod model (3, 6)<br>Otherwise, pertaining to workers or wages (all) |
| $X$ | Total output of exportable goods (7)<br>Exports (quantity exported) (8–10) | $x$ | Growth rate of exports = $\hat{X}$ (8–10) |
| $_X$ | (Subscript) Exportable goods (7)<br>Exports (8–10) | $x_i$, $x_j$ | Arbitrary variables (1) |
| $X_0$ | Constant in export demand function (8–10) | | |
| XX | Market-clearing curve for $X$-sector (7) | | |
| $Y$ | Aggregate output or national income (all) | $y$ | Growth rate of output = $\hat{Y}$ or $\Delta Y/Y$ (all) |
| $Y_0$ | Constant in production function (all) | $y^a$ | Realized rate of growth of output (3, 6) |

| | | | |
|---|---|---|---|
| | | $y_B$ | BP-equilibrium growth rate (9–10) |
| | | $y_D$ | Growth rate of aggregate demand (10) |
| $Y^e$ | Expected output (3, 6) | $y^e$ | Expected rate of growth of output (3, 6) |
| $Y_f$ | Foreign output (8–10) | $y_f$ | Growth rate of foreign output = $\hat{Y}_f$ (8–10) |
| $Y^i$ | Output of firm *i* (4) | | |
| $Y_F^i$ | Full-capacity output for firm *i* (4) | | |
| $Y_n^i$ | Normal output for firm *i* (4) | | |
| $Y_K$ | Full-capacity output (all) | $y_K$ | Growth rate of full-capacity output = $\hat{Y}_K$ or $\Delta Y_K/Y_K$ (all) |
| | | $y_m$ | Growth rate of output in manufacturing (8) |
| | | $y_{nc}$ | Growth rate required to satisfy necessary condition for sustainable steady-state growth (8) |
| $Y_N$ | Full-employment output (all) | $y_N$ | Natural rate of growth = $\hat{Y}_N$ or $\Delta Y_N/Y_N$ (all) |
| | | $y^n$, $y^s$ | Growth rates of northern and southern income (10) |
| | | $y_S$ | BP-equilibrium growth rate for a small country, Razmi model (10) |
| | | $y_t$ | Arbitrary time-series variable (5) |
| | | $y_w$ | Harrod's warranted rate of growth (3, 6) |
| $Z$ | Exogenous variable (all) | $z$ | Ratio of price of foreign goods to domestic unit labour costs = $EP_f/Wa_0$; a measure of international competitiveness (4)<br>Rate of growth of world trade, Beckerman model (8) |
| $Z_D$ | Vector of exogenous determinants of aggregate demand (5) | | |

| | | | |
|---|---|---|---|
| $Z_H$, $Z_N$ | Vectors of exogenous growth drivers in heterodox ($H$) and neoclassical ($N$) growth models (1) | | |
| $Z_J$ | Vector of exogenous determinants of variable $J$ (5) | | |
| $Z_q$ | Vector of exogenous determinants of $q$ (1) | | |

## Greek alphabet*

| | | | |
|---|---|---|---|
| | | $\alpha$ | Exponent in production function (1)<br>Exponent in 'Cobb–Douglas' investment function (4)<br>Degree of indexation of wages to inflation (5)<br>Wage response to productivity growth, $\alpha = 1 - (1/\kappa)(\varphi^e - \varphi)$ (10) |
| | | $\alpha_j$ | Parameters in Kiefer–Rada (2015) econometric model (5)<br>Speeds of adjustment in processes that create or 'tame' Harrodian instability (6)<br>Industry $j$ shares in employment (8)<br>Industry $j$ shares in exports (9–10) |
| | | $\alpha_0$ | Constant term in export function, Ros model (10) |
| | | $\beta$ | Constant in production function (1)<br>Exponent in 'Cobb–Douglas' investment function (4)<br>Effect of productivity growth on wage increases (productivity bargaining) (5)<br>Speed of adjustment of workers' borrowing to differences between their target and wage-funded levels of consumption (7)<br>Speed of unconditional convergence in manufacturing in Rodrik's model (8) |
| | | $\beta_i$ | Share of intermediate imports (9–10) |
| | | $\beta_j$ | Parameters in Kiefer–Rada (2015) econometric model (5)<br>Industry $j$ shares in imports (9–10) |

| | | | |
|---|---|---|---|
| $\Gamma$ | Measure of relative competitiveness, Beckerman model (8) | $\gamma$ | Exponent in production function (1)<br>Parameter in wage equation in original Goodwin model (2)<br>Effect of productivity growth in lowering price inflation (5)<br>Speed of adjustment parameter in the Freitas–Serrano model (7)<br>Response of change in trade share to relative competitiveness, Beckerman model (8) |
| | | $\gamma_i$ | Parameters in Stockhammer–Michell model (7) |
| | | $\gamma_X$ | Elasticity of export supply with respect to domestic price (10) |
| $\Delta$ | Change in variable in discrete time (all) | $\delta$ | Error term, Kiefer–Rada econometric model (5)<br>Intermediate import coefficient in Ribeiro et al. model (10) |
| | | $\delta_L, \delta_K$ | Shares of production workers and capitalist-managers in profit income, Palley model (7) |
| | | $\delta_X$ | Elasticity of export supply with respect to capital (10) |
| | | $\varepsilon$ | Transitory shock term (1)<br>Error term in regression equation (elsewhere) |
| | | $\varepsilon_j$ | Price elasticity of demand for good $j$, where $j = X$ (exports), $M$ (imports), $m$ (manufactured exports), $c$ (final imports), $i$ (intermediate imports) (9–10) |
| | | $\zeta$ | Dividend payout rate (7) |
| | | $\zeta_1, \zeta_2$ | Functions for adjustment of growth rate to the natural rate and 'exogenous' growth factors $Z_N$ (respectively) (1) |
| | | $\eta$ | Elasticity of price–cost margin $(1 + \tau)$ with respect to the real exchange rate (4, 7, 10)<br>Parameter in dynamics of Skott model $= \partial y/\partial g$ (6) |
| | | $\eta_i$ | Coefficients in firms' target profit share equation (5, 6) |

| | | | |
|---|---|---|---|
| | | $\eta_j$ | Income elasticity of demand for good $j$, where $j = X$ (exports), $M$ (imports), $n$ (manufactured exports), $c$ (final imports), $i$ (intermediate imports) (9–10) |
| $\Theta$ | Fundamentals in Rodrik's model (8) | $\theta$ | Speed of adjustment of prices in conflicting claims model (5)<br>Ratio of production workers to managers (7)<br>Trade balance ratio = share of exports in BP receipts = $PX/EP_fM = PX/(PX + NCF)$ (9–10) |
| | | $\kappa$ | Research productivity (1)<br>Wage adjustment parameter (10) |
| $\Lambda$ | Denominators in Marglin's synthesis model (3)<br>Variable used to help define a unit root (8) | $\lambda$ | Risk premium (2)<br>Emulation parameter in Setterfield–Kim model (7)<br>Exponential time trend of exports (9) |
| | | $\lambda_i$ | Coefficients in workers' target wage share function (5, 6) |
| | | $\lambda_X, \lambda_M$ | Pass-through coefficients for export and import prices (9–10) |
| | | $\mu$ | Target for price–cost margin $(1 + \tau)$ in open economy neo-Kaleckian model; reflects 'degree of monopoly' (4, 7, 10) |
| | | $\mu_i$ | Elasticity of imports of intermediate goods with respect to manufactured exports (9) |
| | | $\nu$ | Error term in Hamilton's econometric model (5) |
| | | $\xi$ | Parameter in real wage adjustment equation in original Goodwin cycle model (2)<br>Speed of conditional convergence in Rodrik model (8) |
| $\Pi$ | Total profits (3, 4, 7) | $\pi$ | Profit share (all) [but occasionally used for the number pi, as noted] |
| $\Pi_c$, $\Pi_w$ | Capitalists' and workers' profits, Pasinetti model (3) | $\pi_i$ | Relative productivity of sector $i$ (8) |
| | | $\rho$ | Verdoorn coefficient (8–10) |
| $\Sigma$, $\tilde{\Sigma}$ | Denominators of solutions in wage inequality models (7)<br>$\Sigma$ is also used as summation sign (9) | $\sigma$ | Saving–capital ratio or 'saving rate' = $S/K$ (3–4, 6) |

| | | | |
|---|---|---|---|
| | | $\sigma_K, \sigma_L$ | Saving–capital ratios or 'saving rates' for capitalist-managers and workers, Palley model (7) |
| | | $\sigma_\varepsilon^2$ | Variance of transitory disturbance term (1) |
| | | $\tau$ | Markup rate (all) |
| | | $\tau'$ | Proportional rate of change in price–cost margin $(1 + \tau)$ (9–10) |
| | | $\tau_H, \tau_X$ | Markup rates in home and export goods sectors, gender model (7) |
| | | $\upsilon_i$ | Parameters in equation for utilization adjustment, $\hat{u}$ (5) |
| $\Phi$ | Parameter in Marglin's price inflation equation (3)<br>Function for ratio of income elasticities in Cimoli–Porcile model (10) | $\varphi$ | Response of real wage increases to relative growth of labour demand versus labour supply $(l - n)$ (2)<br>Speed of adjustment of nominal wage in conflicting claims model (5)<br>Share of nominal unit labour costs (NULC) in total unit costs ($AVC$), Ribeiro et al. model (10) |
| | | $\varphi^e$ | Workers' expected share of labour in total unit costs, Ribeiro et al. model (10) |
| | | $\varphi_L, \varphi_K$ | Shares of production workers and capitalist-managers in wage income, Palley model (7) |
| | | $\chi$ | Equity–capital ratio, or proportion of investment financed by equity (3, 7) |
| $\Psi$, $\tilde{\Psi}$ | Denominators of solutions in open economy models (4) | $\psi$ | Wage share (all) |
| | | $\psi_L$ | Wage share of production workers (7) |
| | | $\psi_w, \psi_f$ | Target wage shares of workers, firms (5) |
| | | $\psi_L, \psi_i$ | Labour share and intermediate input share, Ribeiro et al. model (10) |
| $\Omega$ | Parameter in Marglin's wage adjustment equation (3)<br>Consumption not directly affected by profit share in Setterfield-Kim model (7)<br>Constant term in demand regime (DR) (8–9) | $\omega$ | Wage inequality (ratio of managers' wage to production workers' wage) (7) |
| | | $\omega_i$ | Parameters in $\psi$ adjustment equation (5) |

| | | | |
|---|---|---|---|
| | | $\omega_A$ | Share of domestic autonomous expenditures (7–10) |
| | | $\omega_X$ | Share of exports (7–10) |

# Notes:

* Chapter numbers are in parentheses. Upper-case Greek letters that are identical to Roman upper case are not used.
For the most part, levels of variables are expressed in upper-case Roman letters. Growth rates of real variables (quantities) are expressed either using a circumflex (ˆ) or as lower-case Roman letters, but growth rates of nominal variables (prices, wages) are generally expressed only with a circumflex (ˆ) over the corresponding upper-case letter. Greek letters are usually used for parameters and ratios. As much as possible, notation is used consistently throughout the book; where this was not possible, differences between chapters are noted. Please note that the same letter may have a different meaning when used with a superscript or subscript.
In most of the book, growth rates are expressed in continuous time, for example, $x = \hat{X} = \dot{X}/X = (dx/dt)/X = d \ln X/dt$.
However, in some of the mathematics in the text, rates of change may be based on discrete time instead, for example, $x = \Delta X/X$.

# 1

# Introduction: competing theories of production, growth and distribution

## 1.1 Introduction: growth and growth theory in perspective

Sustained economic growth is a relatively recent phenomenon, confined to the last few centuries. During that time, economic growth in advanced capitalist economies (those that have the highest per capita incomes and the longest histories of economic growth) has been sufficient to ensure the doubling of per capita incomes once every 40 years or so – that is, within the working lifetime of each generation. On a global scale, however, the pace of growth has been very uneven. Outside the 'club' of advanced capitalist economies – a club still dominated by western European countries and their former North American and Australasian colonies – economic expansion has been slower, with the result that global growth over the past two centuries has involved *divergence*: the early start and relatively rapid growth of the advanced economies has seen them 'forge ahead' of the rest of the world, which has 'fallen behind' in a process of increasing global income inequality. There has been evidence of the possibility of 'catching up', however. Since the middle of the twentieth century, Asian economies have grown almost twice as quickly as the advanced capitalist economies, closing the per capita income gap. This catching up process is particularly evident in a small number of very successful East Asian economies such as Japan and South Korea, which have grown so rapidly since the 1950s that they have joined the elite club of advanced capitalist economies, while China is now progressing in the same direction.[1]

These are some of the most basic 'stylized facts' of the capitalist growth process that all growth theorists strive to interpret and explain. Perhaps not surprisingly given the central importance of its subject matter, growth theory has a long and illustrious history in economics, and has occupied some of the discipline's great minds. Growth was central to economic analysis in the work of classical economists such as Adam Smith, David Ricardo and John Stuart Mill, as well as Karl Marx, all of whom understood capitalist expansion as a wide-ranging and transformative process affecting the structure and fabric of whole societies. Growth quickly re-emerged as a topic of interest following the Keynesian revolution of the 1930s, as early post-Keynesians such as Joan Robinson and Nicholas Kaldor sought to extend the analysis of John Maynard Keynes – which had been located in the Marshallian short run – to a long run characterized by capital accumulation that would have been recognizable, as an analytical terrain, to the classical economists. This project was foreshadowed by the macrodynamics of Roy Harrod, who himself is often seen as the progenitor of 'modern' growth theory following the turn towards focusing on value theory that characterized the marginalist revolution in economics of the late nineteenth and early twentieth centuries. Harrod's dynamics were quickly interpreted as part of the Keynesian project and as providing a nascent theory of cyclical growth. More recently, Robert M. Solow was awarded the 1987 Nobel Prize in Economics for his contributions to the theory of economic growth, which pioneered the neoclassical approach to growth theory that has prospered ever since, and Paul Romer won the 2018 prize for later contributions to that paradigm (described later in this chapter). The emergence and development of neoclassical growth theory has been accompanied by continued interest in, and further development of, the classical and post-Keynesian traditions in macrodynamics, spawning (among others) neo-Marxian, neo-Harrodian, neo-Kaleckian and neo-Kaldorian analyses of growth.

As these developments suggest, growth theory as a field is now characterized by a great many different specific models of growth. These models can be divided into two broad types: neoclassical growth theory (NGT) and heterodox growth theory (HGT).[2] The main purpose of this chapter is to compare and contrast the structure of NGT and HGT and in so doing provide a basis for the more exclusive focus on, and further development of, HGT in the chapters that follow. Before we set about that task, however, we begin in section 1.2 by outlining some of the basic concepts and definitions that will be employed throughout the book. Section 1.3 then identifies and explores the competing visions of growth that inform the key distinctions between and within NGT and HGT. Particular attention is paid to differing conceptions of the production process and the different understandings of technical change and potential output to which they give rise. The focus on steady and balanced growth analysis (as opposed to cyclical and/or unbalanced growth) that characterizes most (although not all) of the models developed in this book is also discussed. Next, section 1.4 provides an overview of the core models that comprise NGT and HGT. All models are developed so as to emphasize the way they describe the basic 'mechanics' of growth (Jones, 2002). The level of generality so achieved enables us to describe the core insights of the neoclassical and heterodox traditions in terms of just five structural models, representing the three successive generations of NGT (exogenous, endogenous and semi-endogenous) and the two main branches of HGT (classical-Marxian and post-Keynesian). Section 1.5 examines reconciliation of supply and demand in the theory of long-run growth, before section 1.6 offers some reflections on why a specific focus on the study of HGT is worthwhile. Finally, section 1.7 concludes.

## 1.2 Some basic definitions and concepts

Before we begin to examine competing traditions in growth theory, it is useful to define some basic terms and concepts that the reader will encounter throughout the book. It is also important to appreciate how these terms and concepts relate to both equilibrium analysis (the predominant form of theoretical analysis throughout the economics discipline, extensive use of which is made throughout this book) and other methodologies that attempt to reflect the actual historical experiences of capitalist growth (for example, models of cyclical instability or cumulative causation).

Two basic but important concepts in growth theory are *steady* and *balanced* growth. Steady growth occurs when a variable grows at the same constant rate over time. Consider, then, any variable denoted by $x_i$. Using continuous time for mathematical convenience, steady growth in $x_i$ means that

$$\hat{x}_i = \frac{\dot{x}_i}{x_i} = \bar{c} \tag{1.1}$$

where $\hat{x}_i = \dot{x}_i/x_i = d\ln x_i/dt$ is the proportional rate of growth of $x_i$ (where $\dot{x}_i = dx_i/dt$ denotes the increase in the same variable at any instant of time $t$), and $\bar{c}$ is an exogenously given constant.[3] Related to steady growth is the concept of balanced growth, which occurs when *two or more* variables grow at the same constant rate. For example, for any two variables $x_i$ and $x_j$, balanced growth means that

$$\hat{x}_i = \hat{x}_j = \bar{c} \tag{1.2}$$

Very often, balanced growth is a necessary implication of steady growth in economic models, at least over a sufficiently long interval of time. This is because many important economic variables are ratios that are bounded both above and below. One important example is the employment rate ($e$), defined as the ratio of all employed workers ($L$) to the total labour force ($N$):

$$e = \frac{L}{N} \tag{1.3}$$

The employment rate is bounded above and below: it cannot exceed one or be negative. This means that the employment rate itself cannot grow steadily (at a constant rate) in the long run – to suggest otherwise would be to claim that the value of $e$ can eventually exceed unity or fall below zero, either of which is impossible by definition, since $L \leq N$.

Ultimately, then, the only plausible *steady rate of growth* of the employment rate is zero, which value would ensure that the employment rate itself remains constant over time. Note that this, in turn, means that the rate of growth of total employment must equal the rate of growth of the labour force, since it follows from equation (1.3) that

$$\hat{e} = \hat{L} - \hat{N} = 0 \tag{1.4}$$
$$\Rightarrow \hat{L} = \hat{N}$$

In other words, if the employment rate is to remain within its logical bounds at all points in time, steady growth of the employment rate, which satisfies the condition for steady growth in equation (1.1), implies that we must observe *balanced growth* of total employment and the total labour force. In other words, equation (1.4) satisfies the condition for balanced growth in equation (1.2). As will be discussed further below, equation (1.4) may be considered either a 'full employment' condition or, more broadly, a condition that ensures a constant equilibrium rate of employment, $0 < e^* \leq 1$ (a superscript * will be used to indicate the equilibrium level of a variable throughout most of the book).

These seemingly narrow and technical points are important because of their relationship to equilibrium analysis, extensive use of which is made throughout the chapters of this book. More specifically, many (although by no means all) of the models in the chapters that follow are *stable, steady-state equilibrium* models. A steady-state equilibrium occurs when the equilibrium outcome of a model produces a constant rate of growth in the variable (or variables) of interest,[4] while such outcomes are said to be stable if conditions of disequilibrium result in adjustments that move the system towards its steady-state equilibrium path.[5] Our extensive use of stable, steady-state equilibrium analysis is motivated by pedagogical considerations: it is easier, in the first instance, to construct such models if the objective is to clearly characterize the fundamental causal relations of a growth theory. In particular, the properties of stable, steady-state equilibrium models can be analysed using the method of *comparative dynamics*. Like comparative statics, comparative dynamics involves contemplating the impact on the equilibrium configuration of a system of some parametric (exogenous) change.[6] The stability of equilibrium is important (although not essential) for comparative static and comparative dynamic methods alike, because it ensures that once processes of disequilibrium adjustment have 'played out', the reconfiguration of equilibrium associated with a parametric change doubles as a description of how the same parametric change will affect *actual* system outcomes.

It is important to realize, however, that while convenient for pedagogical purposes, our use of stable, steady-state equilibrium models will frequently impose upon the analysis a vision of the growth process as steady and (frequently also) balanced. Hence note that, as defined above, a steady-state equilibrium will produce a constant rate of growth that remains unchanged unless the economy is hit by an unexplained exogenous shock. In other words, a steady-state equilibrium rate of growth is a *steady* rate of growth. Meanwhile, and as noted earlier, the prevalence of bounded ratios among economic variables means that steady growth will often imply balanced growth.

There are certainly exceptions to all this. As will become clear later in this chapter, our exploration of HGT demands that we sometimes abandon stable, steady-state equilibrium analysis, while even those heterodox models that can be analysed in such terms produce results that require careful contextualization, or even extension and/or further analysis, before a steady-state equilibrium can be considered 'final' or 'fully adjusted'.[7] Also, an apparently stable steady-state equilibrium can be upset if forces that compel change (such as balance-of-payments deficits or unsustainable financial positions) that are not explicitly accounted for in a given model build up during the period of steady growth. The realization that such destabilizing forces can emerge has led to new types of theory that attempt to incorporate them, as we shall see in some later chapters. But these caveats notwithstanding, it is important to bear in mind that we will make frequent use of steady-state equilibrium analysis throughout this book, and that as we do so, and whatever its pedagogical virtues, such analysis does involve a specific (and contestable) characterization of the real-world capitalist growth process that we are attempting to model.

This is not to say that describing growth as steady (and even balanced) is demonstrably at odds with the historical growth record. On the contrary, there is some empirical evidence to suggest that capitalist economies are – or at least, for long periods, can be and have been – characterized by steady and balanced growth. For example, Kaldor's celebrated stylized facts point to numerous instances of steady and/or balanced growth in mid-twentieth-century advanced capitalist economies (Kaldor, 1957). Among these stylized facts are constancy of the capital–output ratio – implying balanced growth of real output and the capital stock – and steady growth of the level of output per worker or labour productivity.

Nevertheless, the notion that observed capitalist growth is best characterized as steady and balanced is controversial. For example, many observers contend that long-run growth is inherently *cyclical* rather than constant (steady). This observation is based not only on the commonly accepted phenomenon of the business cycle, but also on claims that growth is characterized by fluctuations of a considerably longer period, such as Kuznets swings (lasting 25 or more years) or Kondratieff waves (lasting approximately 50 years).[8] Indeed, many growth theorists – going at least as far back as Marx – have preferred to conceptualize growth as an inherently cyclical process. None of this is adequately captured by steady-state equilibrium growth models. We will consider models of business cycles in various places in this book, including the neo-Marxian limit cycles of Goodwin (1967) in Chapter 2, the neo-Goodwin cycles of Barbosa-Filho and Taylor (2006) in Chapter 5 and various neo-Harrodian models of cyclical growth (due to Skott, 1989 and Fazzari et al., 2013, among others) in Chapter 6.

Other observers, meanwhile, contend that growth is inherently unbalanced – that different sectors of the economy grow at different rates, so that the growth process is necessarily also a process of *structural change* (changes in the composition of economic activity).[9] In this view, associated especially with the 'growth laws' of Kaldor (summarized by Thirlwall, 1983), growth is always led by key sectors, such as manufacturing (or particular industries), and is always uneven across sectors. This means that studying (for example) deindustrialization and the rise of the service sector is all part and parcel of studying growth. Again, none of this is adequately captured by steady-state equilibrium growth models. Although structural change will not be a major focus in the earlier parts of this book, we will address it in the context of Kaldorian growth models in parts of Chapters 8–10.

Ultimately, the point is not that phenomena such as cyclical growth or structural change are unimportant, but rather that the simplifications that will often assist our exposition of many of the growth models in the chapters that follow involve some amount of sacrifice. Our frequent reliance on steady-state equilibrium analysis means that some of the models we construct cannot satisfactorily capture all of the features of growing capitalist economies that are thought to be relevant by students of the historical growth record. That said, such models often provide the simplest, clearest, and therefore best introduction to the essential causal mechanisms that drive the process of growth in the various theories we consider. Furthermore, as already noted, we will not limit ourselves exclusively to models of steady-state growth in the chapters that follow. Models of cyclical growth, cumulative causation and structural change will also be explored, and we will put considerable emphasis on processes of convergence (or lack of convergence) to equilibrium growth paths.

## 1.3 Competing visions of growth

### 1.3.1 The neoclassical, classical-Marxian and post-Keynesian visions of growth

Turning now to growth theory itself, it is useful to begin by considering, at a very general level, how competing theoretical traditions envision the capitalist growth process. Table 1.1 summarizes the overarching visions of growth that characterize NGT on the one hand, and the two main branches of HGT (classical-Marxian and post-Keynesian analyses) on the other.

**Table 1.1** Fundamental distinctions between competing approaches in long run macroeconomics

| | Conception of the 'long run' | Demand side matters for long-run growth? | Theory of distribution | Characterization of supply side |
|---|---|---|---|---|
| Classical-Marxian | Equalization of rates of return; convergence to 'normal' rates of profit, utilization | No | Surplus approach | Technical and social |
| Neoclassical | Steady state | No | Marginal productivity theory | Technical |
| Post-Keynesian | Evolutionary sequence of short-/medium-run 'episodes' of macro performance | Yes | Surplus approach | Technical and social |

The first thing that is evident from Table 1.1 is the fundamental differences between neoclassical, classical-Marxian and post-Keynesian growth theories with regard to their conceptions of the long run. In neoclassical growth theory, the long run is associated with a steady-state (and balanced) equilibrium position defined and reached independently of the path taken towards it. Shocks that dislodge the economy from its steady-state time path are resolved by negative feedback mechanisms that cause the economy to return relatively quickly to its steady-state equilibrium.[10] In classical-Marxian analysis, meanwhile, the long run is associated with a 'fully adjusted position' characterized by the equalization of rates of return across sectors of the economy and equality between actual and target or 'normal' values of variables (such as the actual and normal rates of capacity utilization). The rate of adjustment towards such a position is not supposed to be rapid, however; indeed, the economy may never actually operate in a fully adjusted position. Instead, it will exhibit tendential gravitation towards such a position over very long periods of calendar time.

Finally, post-Keynesians view the long run as a historical outgrowth of sequences of short- or medium-run episodes of actual performance. They are critical of both neoclassical and classical-Marxian analyses for regarding the long run as a fixed point towards which the economy is inevitably and inexorably moving, regardless of the characteristics of its traverse path (that is, the precise route taken). Although (as we will see) not averse to identifying long-run equilibrium positions and even imbuing them with attractor-like properties, post-Keynesians are inclined to view any long-run equilibrium as the path-dependent product of the sequence of events leading up to it. They are, moreover, inclined also to consider self-reinforcing positive feedbacks as dominating self-correcting negative feedbacks in the event that the economy departs from equilibrium. This thinking is particularly important in the neo-Harrodian and neo-Kaldorian traditions, where it gives rise to celebrated phenomena such as Harrodian instability and cumulative causation.

It is important to remember that despite the systematic differences in vision identified here, the models in the chapters that follow – including those found in several chapters devoted to classical-Marxian and post-Keynesian growth theory – will often be presented as stable, steady-state equilibrium models. Fundamental differences between NGT and HGT as regards their conceptions of the long run and the role of 'history versus equilibrium' will sometimes be made abundantly clear, as (for example) when we discuss models of cyclical behaviour (such as the Goodwin model in Chapter 2), models with unstable disequilibrium dynamics (such as the Harrod model in Chapter 3) and models that eschew equilibrium as an 'organizing concept' in favour of non-equilibrium constructs such as cumulative causation (such as the neo-Kaldorian model in Chapter 8). At other times, however, these differences will be suppressed by our reliance on stable, steady-state equilibrium analysis. The reader is therefore encouraged to bear in mind that when we do make use of such analysis, there will be, on occasion, some loss of fidelity to the underlying vision of the long run that accompanies the model we are developing. As noted earlier, losses of this sort are sometimes a 'necessary evil'. Since equilibrium models often provide clearer exposition of the essential causal relations that characterize particular theories of growth, the loss of fidelity just described is a worthwhile sacrifice to make when providing foundational analyses of the sort that populate this book.

Elsewhere in Table 1.1 we notice sources of overlap between competing theories that occasionally produce seemingly counter-intuitive 'alliances'. For example, both neoclassical and classical-Marxian theories are essentially *supply-led* visions of the long-run growth process (although the importance attached to realization crises by Rosa Luxemburg and others in the 'underconsumptionist' tradition means that the demand side is not altogether passive in some variants of the classical-Marxian tradition). Post-Keynesians, meanwhile, view growth as a fundamentally *demand-led* process, expansion in the availability and productivity of resources serving at most to place an upper bound on the rate of growth of economic activity without acting as the fundamental determinant or driver of growth. Indeed, both the availability and productivity of resources are often seen as being endogenous to the demand-determined actual rate of growth in post-Keynesian theories. In this way, even the seemingly supply-determined limits to economic activity at any point in time are, in fact, likely to be influenced by the demand side of the economy.

Both post-Keynesian and classical-Marxian theories, meanwhile, are inclined to characterize the supply side of the economy in terms of both *social* and *technical* relations of production, whereas in neoclassical theory, the supply side is regarded as a purely *technical* (engineering) construct. Indeed, the treatment of even purely *technical* features of the production process differs between classical-Marxian and post-Keynesian theories on the one hand, and neoclassical growth theory on the other, with important consequences for the ways these competing traditions conceptualize technical change and even the meaning of potential output.[11] Finally, both post-Keynesian and classical-Marxian theories explicitly adopt (or else implicitly allude to) some variant of the classical *surplus approach* to the theory of value and distribution in the course of analysing growth. In contrast, neoclassical growth theory generally adheres to *marginal productivity theory* for determining 'factor prices' (prices of basic inputs such as capital and labour) and hence the distribution of income.

### 1.3.2 Alternative approaches to production, technical change and potential output

In spite of their other differences, the classical-Marxian and post-Keynesian approaches characteristic of HGT share a conception of the production process that differs markedly from that found in NGT. In order to better understand these differences and their significance, it is useful to begin with some terminology and concepts that are common to both HGT and NGT, before turning to what distinguishes them and how the two variants of HGT in turn differ from each other.

#### *Technology and production*

First, a *technique* is a single method of producing a good (such as the set of ingredients required for baking a cake, combined with the labour, energy and capital – the oven and building – required to produce it). More formally, a technique is a vector of inputs that produces a given level of output (or potentially, a vector of outputs, if a production process produces more than one good, in which case it is called 'joint production' – such as meat, skin and fat from an animal). The *technology* at any point in historical time is the set of all available or known techniques that firms can choose from.

*Technological change* or *innovation* (also called *technical progress*) is the introduction of a new technique and/or product, in other words, the improvement or expansion of the technology set. *Process innovations* are new ways of making the same goods, while *product innovations* are new types of goods (which in turn require new production methods). Innovation is a complex process, which involves various steps such as invention, refinement and diffusion of a new product or process. Usually, products go through much further development after their initial invention and there is a non-trivial amount of knowledge and investment required even to adopt existing state-of-the-art techniques in new locations. Some firms may stick with older techniques if it is too costly to adopt new ones (especially considering that some or all of the fixed costs associated with the former are 'sunk costs'),[12] while some old techniques eventually get abandoned and forgotten and are no longer in the effective technology set.

Where HGT and NGT differ is on how best to characterize the technology available at any point in time and how technologies change over time. The standard NGT view is expressed in a continuous aggregate production function of the form

$$Y = F(K, N) \tag{1.5}$$

where $Y$ denotes real output and $K$ represents the capital stock. The level of output $Y$ on the left-hand side of equation (1.5) can be considered a measure of *potential* output, that is, the maximum output that can be produced in principle at any point in time, given the availability and productivity of the factors of production. In equation (1.5), this notion of potential output involves *both* full employment of labour *and* full utilization of the capital stock (which will be defined more precisely below). The term full employment is used here (and throughout this book) to denote the maximum level of employment, at any point in time, that the economy can achieve.[13] Full employment so defined may fall short of 100 per cent of the total labour force because of turnover in the labour market, contributing to so-called frictional unemployment. In this case, there will be a maximum rate of employment associated with full employment, denoted by $e_{max}$, such that $e_{max} < 1$.[14] For the sake of simplicity, however, it is generally useful to normalize $e_{max}$ to a value of one, so that full employment implies $e_{max} = 1$ or $L = N$. We will generally make this simplifying assumption throughout the rest of this book. Note that the assumption is already implicit in equation (1.5) as stated above.

Equation (1.5) is usually assumed to have certain 'well-behaved' properties. First, in order for factors of production to be paid their marginal products while maintaining the assumption of zero excess profits (perfect competition), this function must exhibit constant returns to scale: $xY = xF(K, N) = F(xK, xN)$, where $x > 0$ is any positive number. Second, the function must exhibit positive but diminishing marginal products and obey all 'second-order' conditions for cost minimization ($F_K > 0$, $F_N > 0$, $F_{KK} < 0$, $F_{NN} < 0$, $F_{KK}F_{NN} - F_{KN}{}^2 > 0$). On these assumptions, we can draw a unit isoquant (along which $Y = 1$), and all other isoquants (for higher output levels $Y > 1$) are radial or proportional 'blow-ups' of this unit isoquant. As shown in panel (a) of Figure 1.1, the unit isoquant $Y = 1$ has the standard properties of being downward sloping and convex to the origin.

Most importantly, this specification assumes that *there is an infinite choice of techniques with any given technology*: firms can choose a cost-minimizing point anywhere along the unit isoquant (or any multiple of it) based only on relative factor prices, so that a higher ratio of real wages to the real rate of interest, for example, will induce the (costless and effortless) substitution of capital for labour and the choice of a technique with a higher capital–labour

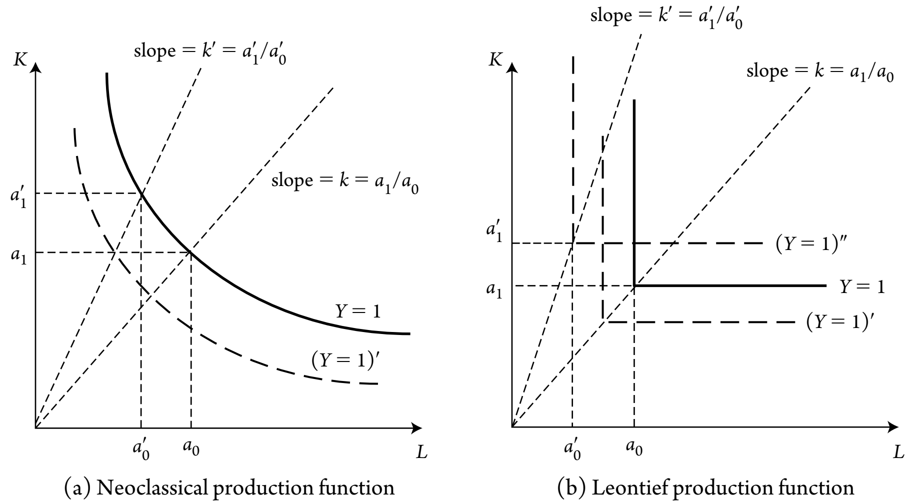



**Figure 1.1** Unit isoquants and technological progress with alternative production functions

ratio. To see this, first note that a ray from the origin to any point on the unit isoquant has the slope

$$k = \frac{K}{N} = \frac{K/Y}{N/Y} = \frac{a_1}{a_0} \tag{1.6}$$

where $a_0$ and $a_1$ denote, respectively, the quantity of labour used to produce one unit of output and the quantity of capital used to produce one unit of output (again assuming full employment of labour and full utilization of capital), and $k$ is the capital-to-labour ratio or 'capital intensity' of production.

Note that if a higher ratio of real wages to the real rate of interest were to give firms an incentive to substitute capital for labour, this change in technique would be reflected in a movement up and to the left along the $Y = 1$ isoquant in panel (a) of Figure 1.1, since the capital intensity of production as defined in equation (1.6) varies continuously along the slope of any given isoquant. As is clear from equation (1.6), such movement requires adjustment in the values of $a_0$ and $a_1$. This is captured in Figure 1.1 by the simultaneous increase in $a_1$ to $a_1'$ and reduction in $a_0$ to $a_0'$, which together increase the value of $k$ to $k'$. The developments just described reveal several other features of the continuous production function used in NGT. First, capital is 'putty-like', the same quantity of capital having the capacity to be utilized by different quantities of workers. Second, factors of production are substitutes for one another: the same level of output $(Y = 1)$ can be produced by more capital and less labour, or less capital and more labour. Finally, a corollary of this substitutability is that individual factors of production have well-defined marginal products. The ability to combine capital and labour in different proportions means that it is possible to produce more output by increasing one factor input while holding the other constant.

HGT, in contrast, typically uses a fixed-coefficients or Leontief production function, named after the famous Russian-American economist Wassily Leontief.[15] The 'fixed coefficients' referred to here are $a_0$ and $a_1$, which are now understood to be the quantity of labour *required* to produce one unit of output and the quantity of capital *required* to produce one unit of output, respectively. For purposes of defining the productive capability (maximum feasible output) of an economy, the Leontief production function takes the form

$$Y = \min\left(\frac{N}{a_0}, \frac{K}{a_1}\right) \tag{1.7}$$

Note that both of the terms in parentheses on the right-hand side of equation (1.7) are quantities of output: $N/a_0$ denotes the maximum output that can be produced with available labour resources; similarly, $K/a_1$ is the maximum output that can be produced using the available capital stock. The output that it is possible to produce – the economy's potential output – is then derived as the *smaller* of these two values, in accordance with the 'min' operator on the right-hand side of (1.7). For example, suppose that baking a cake requires three eggs and one pound of flour, and that a kitchen is currently stocked with nine eggs and four pounds of flour. Four pounds of flour is sufficient to produce four cakes, but nine eggs are sufficient to produce only three cakes, as a result of which total production is limited to a potential output of three cakes.

As this simple example of baking cakes demonstrates, the Leontief production function does not provide an infinite choice of techniques at any given moment in time. Instead, it describes just one technique represented by the kink-point of an L-shaped unit isoquant like $Y = 1$ in panel (b) of Figure 1.1. There could be more than one such technique in existence at any point in time, but normally there are only a limited number of techniques that could be represented by a small number of L-shaped isoquants. In order to significantly change the capital intensity of production, therefore, a firm must normally engage in technological change by either inventing or adopting a new technique of production that was not previously available to it (for example, new machinery, equipment or software). Neither invention nor adoption is costless, of course. In the heterodox view, high labour costs (wages) can induce a search for labour-saving, capital-intensive techniques, but this generally requires innovative effort rather than a mere movement along an existing isoquant.

These observations draw to attention several other salient features of Leontief production functions. First, capital is 'clay-like': just as clay that has already been baked into a certain shape cannot be remoulded into another form, similarly capital that has been invested in a particular type of machinery or equipment cannot be transformed into other types that could be utilized by different quantities of workers. Second, factors of production are strict complements: the same level of output $(Y = 1)$ can only be produced by one particular vector of capital and labour inputs, the ratio of which strictly coincides with the fixed ratio $a_1/a_0$ determined by the state of technology embodied in the production process. Finally, a corollary of this lack of substitutability is that individual factors of production do not have marginal products. It is impossible to produce more output by increasing one factor input while holding the other constant. For example, adding more eggs to a fixed quantity of flour in an effort to produce more cakes would instead succeed only in creating a rather unpalatable omelette. This is especially significant because it means that production according to the principles of a Leontief production function will usually mean that some quantity of productive inputs (using the numbers given previously, one pound of flour) will lie idle. We will return to discuss this observation in greater detail below.

Of course, any theoretical model is an abstraction, and the two extreme cases depicted in Figure 1.1 of only a single technique or an infinite array of techniques are merely theoretical devices. Most heterodox economists, however, believe that a focus on a small number of discrete techniques (with innovation required to develop new ones) is a better first approximation to reality than the assumption of an infinite range of techniques with a given technology. As a result, implicit or explicit appeal to Leontief production functions will be seen to proliferate in the chapters that follow.

The two isoquant diagrams in Figure 1.1 can also be used to show how innovation is differently conceived in the different approaches to production theory favoured by NGT and HGT. The neoclassical view requires that the entire production function, or the whole set of isoquants, must shift simultaneously inward, so that fewer inputs are required per unit of output. Mathematically, this can be represented by inserting the technological shift parameter $A > 0$ into the production function (1.5), which becomes $Y = AF(K, N)$. A rise in $A$ is represented in panel (a) of Figure 1.1 by an inward shift of the unit isoquant to $(Y = 1)'$, thus creating a new infinitely large set of available techniques that is entirely superior to the old set.[16]

In contrast, the heterodox view – which coincides with the perspective of many technology experts, such as Rosenberg (1976) – is that firms will not seek to innovate across the entire spectrum of possible techniques, but instead will focus their (costly) innovative efforts on an economically relevant range of techniques – such as ones that would replace some current workers with machines or robots in countries where wages are high. This is depicted in panel (b) by two possible such innovations: a 'Hicks-neutral' one, shown by $(Y = 1)'$, in which $a_0$ and $a_1$ both decrease proportionately so that the kink-point of the Leontief isoquant moves down along a ray towards the origin; and a 'Marx-biased' one, shown by $(Y = 1)''$, which reduces $a_0$ but increases $a_1$ so that the new kink-point is above and to the left of the initial one. Note that a series of innovations of the latter type might appear to trace out a neoclassical-looking isoquant, but this would be a false appearance because the more capital-intensive techniques can only be reached through a process of innovation, not by selecting from a pre-existing choice of techniques.[17] These and other types of innovations will be discussed in more depth in section 2.6 of Chapter 2.

### *Potential output, capacity utilization and employment*

The Leontief production function in equation (1.7) actually implies two *different* conceptions of potential output, the identification of which will prove important throughout this book. Depending on whether labour or capital is the binding constraint on production, equation (1.7) can imply *either*

$$Y_N = \frac{N}{a_0} \tag{1.8}$$

or

$$Y_K = \frac{K}{a_1} \tag{1.9}$$

where $Y_N$ denotes 'full-employment output' (the measure of potential output that is constrained by the availability of labour) and $Y_K$ denotes 'full-capacity output' (the measure of potential output that is constrained by the size of the capital stock). In the rest of this book, we will consistently define $a_1$ as the ratio of capital to full-capacity output, $a_1 = K/Y_K$, as implied by equation (1.9). However, $a_0$ will be *defined* as the ratio of *employed* labour to *actual* output, $a_0 = L/Y$, which we will generally assume to be constant regardless of whether or not labour is fully employed – so that $a_0 = L/Y = N/Y_N$.[18] The reasons for this asymmetry in how the two input–output coefficients are defined will become clear when we consider models of less than full utilization of capacity (capital), below and in later chapters.

The two measures of potential output described in equations (1.8) and (1.9) will only coincide, at any point in time, as a special case, when (by pure coincidence) $N/a_0 = K/a_1$. Only in this very special case would *all* capital and *all* of the labour force (economically active population) be engaged in productive activity simultaneously. Since that is not generally observed in practice (and is theoretically unlikely with a fixed-coefficients technology), the *normal* situation is that *either* some labour is unemployed *or* some capital is underutilized (or both). Recall that we previously defined the maximum possible rate of employment of the labour force associated with full employment as $e_{max}$. NGT models generally assume that $e = e_{max}$ at all times, but HGT models do not in general, and indeed vary from one to another with respect to what determines $e < e_{max}$ as we shall see in the chapters that follow.

Analogous to the employment rate, we can define the rate of capacity utilization (or, more briefly, 'utilization rate') as

$$u = \frac{Y}{Y_K} \tag{1.10}$$

which is the ratio of actual output to potential output defined by equation (1.9), that is, output with full utilization of the capital stock. NGT models typically assume that $u = 1$ (or, equivalently, $Y = Y_K$) at all times, and reconcile this with full employment of labour ($e = e_{max} = 1$ or $L = N$) by assuming that factor substitution along a production function of the form described by (1.5) allows for the simultaneous achievement of full employment and full utilization by varying capital–labour proportions.[19] HGT models, in contrast, vary in their assumptions about capacity utilization, but many of them allow for an equilibrium utilization rate that may be less than 100 per cent ($u < 1$). Note that *any* equilibrium for the utilization rate represents a specific type of steady-state, balanced-growth equilibrium, since $\hat{u} = 0 \Rightarrow \hat{Y} = \hat{Y}_K = \hat{K}$ (where the last of these equalities holds only if $a_1$ is constant).

For any equilibrium utilization rate $u \leq 1$, we can define the variable $K_u = uK$ as the amount of the available capital stock, $K$, that is utilized to produce the current actual level of output, $Y$, at any point in time. Note that actually utilized capital $K_u = uK$ is analogous to the level of employment $L = eN$. It follows from the definition of $K_u$ and equation (1.10) that

$$\frac{K_u}{Y} = \frac{uK}{Y} = \frac{Y}{Y_K} \cdot \frac{K}{Y} = \frac{K}{Y_K} = a_1 \tag{1.11}$$

In other words, and given the current state of technology, just as $a_0$ remains constant whether or not labour is fully employed, so $a_1$ remains constant whether or not the capital stock is fully utilized. Nevertheless, it should be borne in mind that, as defined here, $a_1$ is a constant ratio of capital to a measure of *potential* output (full capacity), while $a_0$ is a constant ratio of labour to *actual* output.[20]

Many HGT theories (especially those of classical-Marxian or Harrodian inspiration) assume that the economy converges to a 'normal' rate of capacity utilization, denoted by $u_n$, in a long-run (or long-period) equilibrium. In principle, this normal rate need not be 100 per cent utilization of capacity. A normal rate of capacity utilization below 100 per cent ($u_n < 1$) can arise as a matter of choice on the part of firms, who may deliberately seek to maintain excess capacity to insulate themselves against unforeseen variations in product demand or for strategic reasons in oligopolistic rivalry, and may therefore be unwilling to allow the actual rate of capacity utilization to drift above or below this normal rate.[21] However, for some theories (especially the classical-Marxian and neo-Keynesian growth models covered in Chapters 2 and 3) it is useful to 'normalize' the normal rate $u_n$ by assuming $u_n = 1$ for mathematical convenience. In that case, the Leontief production function in equation (1.7) applies, but capital is generally considered to be the binding constraint so that actual output is $Y = Y_K = K/a_1 \leq L/a_0$ and in general there may be unemployed labour.

Other HGT models, especially those inspired by Kalecki (1971b), tend to assume simply that industrialized economies typically operate below a maximum technically feasible rate of utilization ($u < 1$), and do not have any tendency to converge to a predetermined normal rate of utilization in the long run. Thus, the neo-Kaleckian models we will cover in Chapters 4 and 5 assume that there is a maximum utilization rate of unity, but the actual utilization rate can settle at *any* level below this maximum even in a long-run equilibrium. In neo-Kaleckian or other post-Keynesian models that allow for underutilization of capacity and in which output and employment are determined by aggregate demand, the Leontief production function for *actual* output effectively becomes

$$Y = \min\left(\frac{L}{a_0}, \frac{K}{a_1}\right) \tag{1.7'}$$

where $L \leq N$ is the actual level of employment and $Y = L/a_0 \leq Y_K = K/a_1$ so that there is generally excess capacity or underutilized capital.

The distinction between the two heterodox approaches to output determination (equations 1.7 and 1.7′) and the different definitions of potential output in equations (1.8) and (1.9) can also be visualized graphically. Figure 1.2 depicts typical classical-Marxian and post-Keynesian conceptions of how the

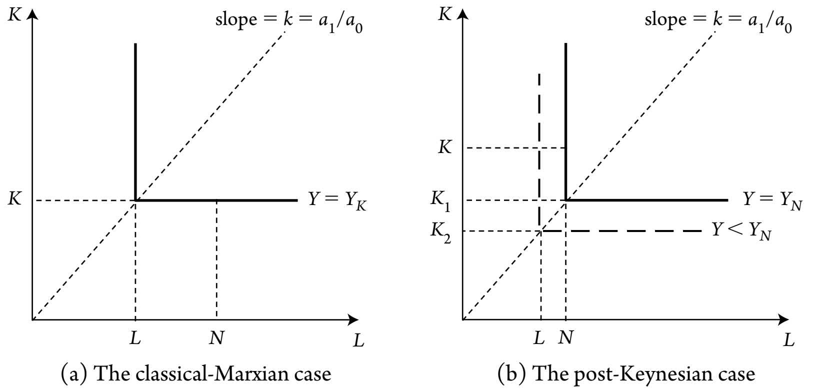



**Figure 1.2** Different concepts of potential output and resource underutilization in HGT.

same economy, endowed with $K$ units of capital and $N$ units of labour at a particular point in time, is likely to function. In the classical-Marxian case in panel (a) of Figure 1.2, $Y_K$ denotes the economy's potential output which (on the simplifying assumption that $u_n = 1$) is also its *actual* output, the economy in question being *capital constrained*.[22] Meanwhile, given the available labour force $N$, an abundance of labour clearly exists over and above the amount $L$ required to produce $Y_K$ in conjunction with the available capital stock. Given the capital stock $K$, the actual level of employment $L$ is determined by $L = a_0 Y_K = (a_0/a_1)K$, or in other words, employment is proportional to capital. The abundance (or excess supply) of labour, given by $N - L$, is the Marxian 'reserve army' of unemployed workers.

The typical post-Keynesian economy, meanwhile, is *demand constrained*. An outcome of this sort is depicted in panel (b) of Figure 1.2. To begin with, the economy's potential output is associated with $Y_N = N/a_0$, which, in accordance with the required input proportion $k = a_1/a_0$, involves a capital requirement of $K_1 < K$. In other words, the economy's potential output is associated with conditions of full employment rather than full capacity utilization, there being idle capital even when the economy reaches its full-employment output at $Y_N = N/a_0 = K_1/a_1 < K/a_1$. However, as is clear from Figure 1.2, the economy will not typically operate at its full-employment output level, but instead at the demand-determined level of output $Y < Y_N$ at which (again in keeping with the necessary input proportions $k = a_1/a_0$) the capital and labour requirements are $K_2 < K$ and $L < N$, respectively. As in the classical-Marxian case, there is once again underutilization of labour resources. In the post-Keynesian economy, however, the source of the abundance of labour $N - L$ is the deficiency of aggregate demand for final output (which determines $Y$ independently of the resource-utilization outcomes depicted in Figure 1.2). This abundance of labour is involuntary unemployment. And of course, most post-Keynesian models solve for some degree of capacity underutilization ($u < 1$), which may or may not (depending on the specific model) coincide with a preordained 'normal' degree of excess capacity.

Thus, the classical-Marxian and post-Keynesian branches of HGT differ on which concept of potential output is typically most relevant, how a capitalist economy usually functions with respect to proximity to its potential output, and the interpretation of any excess of productive resources this involves.[23] The debate about whether the utilization rate is freely flexible or should be assumed to settle at some predetermined normal rate (or within a normal range) in the long run will be covered in depth in Chapter 6. Without resolving that debate here, for present purposes we can think of $u_n$ as defining 'full-capacity output', regardless of whether that is understood as a normal rate including some excess capacity or a true technological maximum. In this case, the specifications of full-employment output and full-capacity output in equations (1.8) and (1.9) respectively can be thought of as taking these additional resource constraints into account by normalizing the values of both $e_{max}$ and $u_n$ to values of 1, so that 'full' employment and utilization imply $e_{max}N = N = L$ and $u_nK = K = K_u$, respectively. We will relax these assumptions in some later chapters, but for now the normalizations just described serve the useful purpose of minimizing algebraic notation throughout the analysis in this chapter.

### *Implications for growth*

The maximum *levels* of output stated in equations (1.8) and (1.9) have important implications for the *growth* of output, by imposing 'ceilings' on the output path of a growing economy that are related to resource constraints. Hence it follows from equation (1.8) that

$$\hat{Y}_N = y_N = n - \hat{a}_0 \tag{1.12}$$

where $n = \hat{N}$ is the rate of growth of the labour force and $y = \hat{Y}$ is the growth rate of output (with the $N$ subscript indicating potential output defined by full employment).[24] Let us define labour productivity ($Q$) as output per worker-hour, or the reciprocal of the labour-to-output ratio, $a_0$:

$$Q \equiv \frac{Y}{L} = \frac{1}{a_0} \tag{1.13}$$

It follows that the growth rate of labour productivity can be expressed as

$$q = \hat{Q} = -\hat{a}_0 \tag{1.14}$$

Then, substituting equation (1.14) into (1.12), we arrive at

$$y_N = q + n \tag{1.15}$$

where, following Harrod (1939), $y_N$ is called the 'natural rate of growth'.

Thus, the natural rate of growth is simply a maximum rate of growth that can be achieved in the long run, given labour supply constraints and the rate of growth of labour productivity (although, as discussed later in this chapter, these 'constraints' may be flexible and can endogenously adjust – they are not necessarily rigid or exogenously fixed constraints). As such it is not to be confused with constructs such as the Wicksellian natural rate of interest or Friedman's natural rate of unemployment, which have the properties of stable equilibria (fixed points towards which the economy will automatically gravitate). Alternatively, as noted earlier, the natural rate of growth can be thought of simply as the growth rate that maintains a constant employment rate, given the (possibly endogenous) rate of increase in *effective* labour supply, $n + q$.

Similarly, it follows from equation (1.9) that the growth rate of potential output defined by full utilization of the capital stock is given by

$$\hat{Y}_K = y_K = \hat{K} - \hat{a}_1 \tag{1.16}$$

It is often assumed that, as a stylized fact, $\hat{a}_1 = 0$ – that is, the full-capacity capital-to-output ratio is constant in the long run.[25] Under this assumption, equation (1.16) can be simplified to

$$y_K = g \tag{1.17}$$

where $g = \hat{K}$ is the rate of growth of the capital stock or the *rate of accumulation*. As we will see throughout the chapters that follow, much of the enterprise of HGT is devoted to explaining the determination of the equilibrium growth rate ($y$ for output or $g$ for the capital stock, depending on the specific model). At this point, we need focus only on the fact that, however determined, the rate of capital accumulation will define the full-capacity growth rate $y_K$ – the maximum rate of growth consistent with full utilization of the capital stock – as in equation (1.17), provided that $a_1$ remains constant in the long run. But we should note that some HGT models, especially those of Marxian origin, dispute that $a_1$ tends to remain constant in the long run (the consequences of changes in $a_1$ will be analysed in Chapter 2).

Both $y_N$ and $y_K$ define maximum rates of growth that can be achieved in the long run without the economy exceeding its capacity to produce at any point in time. Either of these maxima can, however, be exceeded in the short run, if the economy begins from a *level* of output *below* potential, and is therefore underutilizing its resources. Hence we can have $y > y_N$ or $y > y_K$, without violating the technical requirements of production given by the fixed values of $a_0$ and $a_1$, by increasing the rate of employment or the capacity utilization rate, and so expanding output faster than the rate at which the labour force or the capital stock are, themselves, growing. This cannot be achieved indefinitely, of course: eventually the economy will reach full employment or full capacity utilization and from then on will be constrained to grow no faster than the maximum rates given by (1.15) and (1.17).

At the same time, it is important to realize that even if the actual rate of growth is equivalent to either the full-employment or full-capacity growth rates in (1.15) and (1.17), so that $y = y_N$ or $y = y_K$, this does not necessarily mean that the economy is moving along its potential output path – that is, the ceiling traced out by the expansion over time of full-employment output or full-capacity output. To see this, suppose that $y = y_N$ but that $Y_0 < Y_{N0}$ in some initial period 0. It follows that

$$Y_t = (1 + y)^t Y_0 = (1 + y_N)^t Y_0 \tag{1.18}$$

(since $y = y_N$), and

$$Y_{Nt} = (1 + y_N)^t Y_{N0} \tag{1.19}$$

Combining the information in (1.18) and (1.19) tells us that

$$\frac{Y_t}{Y_{Nt}} = \frac{(1 + y_N)^t Y_0}{(1 + y_N)^t Y_{N0}} = \frac{Y_0}{Y_{N0}} < 1 \tag{1.20}$$

By appeal to similar reasoning, we can state that if $y = y_K$ but $Y_0 < Y_{K0}$ initially, then

$$\frac{Y_t}{Y_{Kt}} = \frac{Y_0}{Y_{K0}} < 1 \tag{1.21}$$

Now, since $Y = L/a_0$ and $Y_N = N/a_0$ by the definition of $a_0$, it follows from (1.20) that

$$e_t = \frac{L_t}{N_t} = \frac{L_0}{N_0} = e_0 < 1 \tag{1.22}$$

In other words, we will observe a constant rate of employment (and hence a constant rate of unemployment, defined as $U = 1 - e$). Meanwhile, since the rate of capacity utilization is defined as $u = Y/Y_K$, it follows from (1.21) that

$$u_t = \frac{Y_t}{Y_{Kt}} = \frac{Y_0}{Y_{K0}} = u_0 < 1 \tag{1.23}$$

Then, recalling that $e_{max} = u_n = 1$ by hypothesis, we can see that equations (1.22) and (1.23) imply (respectively) that $e_t < e_{max}\ \forall t$ and $u_t < u_n\ \forall t$ (where $\forall t$ means 'at all times $t$'). In short, $y = y_N$ or $y = y_K$ only guarantees that the economy realizes a *constant* rate of employment (unemployment) or capacity utilization, which could be below the rates consistent with full employment or full capacity utilization. Only in the special cases where $y = y_N$ *and* $Y_0 = Y_{N0}$ initially, or $y = y_K$ *and* $Y_0 = Y_{K0}$ initially, will the economy move along its full-employment or full-capacity output path, and so experience growth consistent with full employment or full capacity utilization.

In sum, whether of classical-Marxian or post-Keynesian pedigree, HGT models are similar in their treatment of the technical conditions of production on the supply side of the economy, and typically predict that capitalist economies will move along their actual output paths (in accordance with their underlying rates of growth) in a manner that, consistent with the technical conditions of production, involves a permanent abundance of labour. They differ as to the precise mechanisms that bring about these common outcomes, however. This, in the process, can be associated with differing conceptions of what typically constitutes the economy's potential output (full-capacity output or full-employment output), and gives rise to different interpretations of the abundance of labour itself, as either a reserve army of unemployed workers (in the capital-constrained classical-Marxian economy), or involuntary unemployment (in the demand-constrained post-Keynesian economy).

## 1.4 Competing models of growth: a preliminary overview

The chapters that follow develop multiple heterodox growth theories in the classical-Marxian and post-Keynesian traditions at length. The purpose of this section is to briefly outline the analytical structure of models in these two main heterodox traditions, and in so doing to contrast them with the analytical structure of mainstream neoclassical models (that are not the subject of analysis in later chapters). We begin, then, with an overview of the neoclassical tradition in growth theory.

### 1.4.1 The neoclassical tradition

*The Solow model*

Beginning with the work of Solow (1956) and Swan (1956), what became known as the Solow model first emerged in response to the problems of dynamic instability that Harrod (1939) identified as likely to encumber a growing economy.[26] The model was subsequently treated as descriptive of the actual dynamics of a capitalist economy and, in retrospect, can be thought of as the first generation of NGT.

The core of the Solow model can be written in a very stripped-down fashion as:

$$y_N = q + n \tag{1.24}$$

$$q = \bar{q} \tag{1.25}$$

$$n = \bar{n} \tag{1.26}$$

$$y = y_N \tag{1.27}$$

Equation (1.24) will immediately be recognized as restating Harrod's natural rate of growth (the full-employment growth rate) in equation (1.15), which was derived from the full-employment output level associated with a Leontief production function.[27] However, the derivation of Harrod's natural rate need not be associated with a Leontief production technology, but instead follows from the accounting relationship

$$Y_N \equiv \frac{Y_N}{e_{max}N} \cdot \frac{e_{max}N}{N} \cdot \frac{N}{Pop} \cdot Pop \tag{1.28}$$

$$\Rightarrow Y_N \equiv \frac{Y_N}{N} \cdot \frac{N}{Pop} \cdot Pop$$

given our assumption that $e_{max} = 1$, where $Pop$ is the total population and the labour force participation rate $N/Pop$ is assumed to be exogenously given in the long run. In the neoclassical context, $y_N$ represents the proportional rate of growth of the level of output derived from the continuous production function in equation (1.5) which, as previously noted, involves *both* the full employment of labour *and* the full utilization of capital stock. This means that in neoclassical growth theory, $y_N = q + n = y_K$. Equations (1.25) and (1.26), meanwhile, treat the rates of growth of productivity and the labour force as exogenously given constants. Finally, equation (1.27) equates the actual rate of growth of output, $y$, to the natural rate, and so equates actual output growth to the expansion of potential output, consistent with the full employment of labour and full utilization of capital, determined on the supply side of the economy.

From a neoclassical perspective, equation (1.27) is a causal statement:[28] $y_N$ causes $y$, so that growth is strictly supply-determined in the long run. Put differently, there is no role for aggregate demand in the determination of long-run growth outcomes in the Solow model. Demand is instead assumed to automatically adjust to accommodate the economy's supply-determined potential output, which means that describing the expansion of the latter suffices to describe the actual rate of growth of the economy.[29] The simplest variant of this story is the corn economy, in which any corn not consumed in one period is, by definition, corn seed for use in the following period's production. In this case, equation (1.27) can be written as an identity. More generally, in a closed economy with no fiscally active public sector, leakages from and injections into the circular flow of income amount to saving and investment, respectively. In this case, the neoclassical causal interpretation of equation (1.27) can be represented as

$$S = sY \tag{1.29}$$

$$I \equiv S \tag{1.30}$$

implying

$$I = sY \tag{1.31}$$

Equation (1.31) states that investment ($I$) is determined by household saving ($S$) which, in turn, depends on total income and a uniform household propensity to save, $s$ $(0 < s < 1)$.[30]

Solving the model in equations (1.24)–(1.27) yields the result

$$y = \bar{q} + \bar{n} \tag{1.32}$$

According to equation (1.32), the rate of growth of output is *determined* by the rates at which the productivity and availability of labour increase in the long run. Note that, despite its important role in the neoclassical causal interpretation of equation (1.27), the saving behaviour of households has no effect on the steady-state growth rate in (1.27). Both productivity growth and the rate of growth of the labour force are taken as exogenously given and are therefore unexplained within the model itself. As a result, the first-generation neoclassical growth model is commonly described as an *exogenous growth* model: in the long run, the rate of growth is ultimately imposed upon the system from without.

In order for output to grow at the rate given by equation (1.32), the Solow model requires flexible adjustment of the capital–labour and capital–output ratios in the process of production to facilitate adjustment towards the steady-state equilibrium. Of course, this flexibility arises naturally from the properties of a continuous (neoclassical) aggregate production function of the sort described by equation (1.5). To see this more clearly, consider a slight reformulation of the production function in equation (1.5), of the form

$$Y = F(K, AN) \qquad (1.5')$$

Here, $A$ is used as an index of labour efficiency or labour-augmenting technology, such that $A$ grows at the rate $q$, and $AN$ measures labour input in efficiency units (or 'effective labour' inputs).[31] In what follows, we will assume that $\bar{q} = 0$ in equation (1.25), so that there is no labour-saving technical change and the value of $A$ remains constant. This is a simplifying assumption designed to avoid having to rewrite the model in terms of 'efficiency units', and which in turn facilitates comparisons with other models covered later.

Since the production function in (1.5′) exhibits constant returns to scale, we can multiply both $Y$ and the arguments in the production function $F(\cdot)$ on the right-hand side of (1.5′) by $1/N$, which gives us the production function in intensive form:

$$Q = f(k, A) \qquad (1.33)$$

Recalling equation (1.26), it follows from the definition of $k$ that

$$\hat{k} = \hat{K} - \bar{n}$$
$$\Rightarrow \dot{k} = \frac{K}{N}\frac{\dot{K}}{K} - \bar{n}k$$

where it should be noted that we are ignoring depreciation of the capital stock, so that all investment constitutes net investment in new capital. Then, since it follows from (1.30) and (1.31) that $\dot{K} = I \equiv S = sY$, and since $Y/N = Q = f(k, A)$ by equation (1.33), we can write

$$\dot{k} = sf(k, A) - \bar{n}k \tag{1.34}$$

The equilibrium condition for this first-order differential equation is $\dot{k} = 0$, from which it follows that

$$sf(k, A) = \bar{n}k$$
$$\Rightarrow \frac{k}{f(k, A)} = \frac{s}{\bar{n}}$$

or

$$a_1^* = \frac{s}{\bar{n}} \tag{1.35}$$

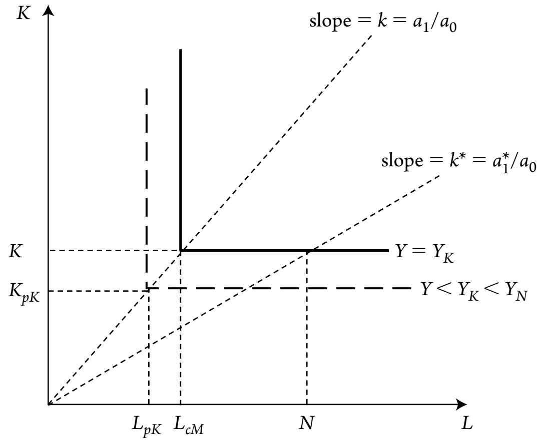
given that

$$\frac{k}{f(k, A)} = \frac{k}{Q} = \frac{K/N}{Y/N} = \frac{K}{Y} = a_1$$

In other words, the capital–output ratio $a_1 = K/Y$ must adjust to a unique equilibrium level $a_1^*$ in order for the equilibrium growth rate (1.32) to be achieved.

This result draws attention to the fact that although (as previously noted) the saving propensity does not affect the long-run, steady-state growth rate in the Solow model, it does (positively) affect other outcomes in the model – including the long-run equilibrium capital–output ratio, as is evident from the right-hand side of equation (1.35), and the steady-state equilibrium level of output per worker.[32] The result in (1.35) also affords further insight into the contrast between the treatment of technical features of the supply side in NGT and HGT. Figure 1.2 shows that in HGT, using Leontief technology to describe the technical structure of the supply side, *factor utilization* (either the rate of unemployment or the rate of capacity utilization, or both) is the adjusting variable that reconciles the technical demands of production with the level of activity at any point in time – regardless of whether that latter is supply-determined (as in the classical-Marxian case) or demand-determined (as in the post-Keynesian case).

Now consider Figure 1.3, which shows how NGT can be contrasted with both variants of HGT using the Leontief technology associated with HGT as a framework of reference. To facilitate this diagrammatic comparison,




Note: The subscript *cM* refers to the classical-Marxian variant of HGT, while *pK* refers to the post-Keynesian variant

**Figure 1.3**
Facilitating adjustments on the supply side in NGT and HGT

although it may be somewhat artificial, we will assume that the labour coefficient $a_0$ is exogenously fixed for all of these theories, so that all the variation in capital–labour proportions in NGT has to come from variations in the capital coefficient $a_1$ (but the latter coefficient is exogenously fixed in both versions of HGT). In the capital-constrained classical-Marxian case ($cM$), the economy operates at the corner of the $Y = Y_K$ isoquant facilitated by the adjustment of labour input to $L_{cM}$, consistent with a reserve army of size $N - L_{cM}$. The demand-constrained post-Keynesian ($pK$) economy, meanwhile, operates at the corner of the $Y < Y_K < Y_N$ isoquant, which outcome is facilitated by the adjustment of capital input to $K_{pK}$ and labour input to $L_{pK}$, consistent with both underutilization of the capital stock ($u < u_n = 1$) and involuntary unemployment ($N - L_{pK}$). Finally, the neoclassical economy operates at a level of output consistent with *both* the full utilization of the capital stock ($u = u_n = 1$) *and* the full employment of labour ($e_{max} = 1$, or $L = N$). Adjustment to this outcome is facilitated by adjustment of the capital–labour ratio to the value $k^* = a_1^*/a_0$, where $a_1^*$ is given by equation (1.35) and, bearing in mind that we are holding $a_0$ constant for the purposes of this comparative exercise,[33] the $k$ ratio adjusts to this equilibrium level solely through changes in $a_1$. This occurs at the point ($N$, $K$) in Figure 1.3, which (given that the capital coefficient has adjusted to $a_1^*$) represents full employment of both inputs ($Y = Y_N = Y_K$).

## *Neoclassical endogenous growth theory*

Associated with Romer (1986, 1990) and Lucas (1988), among others, neoclassical endogenous growth theory (NEGT) builds on the Solow model to create a second generation of neoclassical growth theory, primarily by developing a theory of technical change to replace the assumption of an exogenously given growth rate of labour productivity.[34] The primary purpose of this second-generation neoclassical model is to move beyond the exogenous growth feature of the first-generation (Solow) model.

The essential claims of NEGT can be represented by modifying slightly the system of equations used previously to outline the Solow model. Retaining equations (1.24), (1.26) and (1.27), we replace equation (1.25) with

$$q = q(Z_N) \tag{1.36}$$

and

$$Z_N = \overline{Z}_N \tag{1.37}$$

where equation (1.36) is a technical progress function in which $Z_N$ is a vector of variables that affect the resources devoted to and/or the incentives to produce technological change. The malleability of the precise specification of $Z_N$ and, indeed, of the precise functional form of (1.36) is what gives rise to the great variety of models associated with NEGT.[35]

Solving the model in equations (1.24), (1.26), (1.27), (1.36) and (1.37) now yields

$$y = q(\overline{Z}_N) + \overline{n} \tag{1.38}$$

According to equation (1.38), the rate of growth is endogenous in the sense that, given the rate of growth of the labour force, it is driven by technological change that (1) is explicitly described by the technical progress function in (1.36); and (2) occurs at a rate that is amenable to change in response to variation (by private decision makers and/or policy makers) in the vector $Z_N$ (Roberts and Setterfield, 2007, pp. 14–16). The latter may include such choice variables as tax rates, saving propensities, intertemporal discount rates and/or human capital accumulation, depending on the construction of the specific NEGT model at hand.

To demonstrate this more concretely, we consider here a simple '*aK*' model of neoclassical endogenous growth.[36] We begin by first giving the continuous aggregate production function in (1.5′) an explicit (Cobb–Douglas) form:

$$Y = K^{\alpha}(AN)^{1-\alpha} \tag{1.39}$$

Now assume that the labour-augmenting technology, $A$, varies directly with the capital–labour ratio:

$$A = \beta \frac{K}{N} \tag{1.40}$$

In other words, rather than being determined exogenously, $A$ is now the result of capital accumulation by firms, the assumption being that workers learn from (and become more productive from) the experience of using more (and more advanced) capital. Consider, for example, an office in which typewriters are replaced with computer-based word processors. Use of the former requires only that workers learn to type. Use of latter, however, demands that they *also* become familiar with computer operating and file-management systems. Ostensibly engaged in the same basic task (letter writing, for example), workers themselves become more skilled and therefore more productive independently of the advantages for the production process that result from the new capital itself.

Substituting (1.40) into (1.39), it follows that

$$Y = aK \tag{1.41}$$

where $a = \beta^{1-\alpha}$. It will now be self-evident to the reader why this is called an '$aK$ model'!

Next, we need to transform equation (1.41) from the level of output to the growth rate of output. Recalling once again that, from (1.30) and (1.31), we have $\dot{K} = I \equiv S = sY$,[37] it follows from (1.41) that

$$\dot{Y} = a\dot{K} = asY$$

and hence

$$\frac{\dot{Y}}{Y} = y = as \tag{1.42}$$

According to equation (1.42), variations in the household saving rate are chiefly responsible for variations in long-run growth. Note the contrast with the Solow (exogenous growth) result derived previously. Household economic behaviour – specifically, saving behaviour, which was claimed to be irrelevant for the determination of the long-run growth rate in the Solow model – is now central to the determination of that rate.

The reader may be left to wonder exactly how this result has been achieved, since the differences between NEGT and the Solow model (equations 1.36

and 1.37 rather than 1.25) are few indeed. To understand the mechanisms at work, begin by dividing both sides of (1.41) by $N$, yielding

$$Q = ak \tag{1.43}$$

Equation (1.43) expresses the $aK$ production function in intensive form, and is the counterpart to the intensive-form production function in the Solow model in equation (1.33). If we replace the expression for $Q$ in the Solow model with equation (1.43) in the differential equation for $k$ in (1.34), we arrive at

$$\dot{k} = sak - \overline{n}k \tag{1.44}$$

Note that $s$, $a$ and (of course) $\overline{n}$ are all constants, so that if $sa > \overline{n}$, the result in (1.44) is $\dot{k} > 0$ for all values of $k$: the capital–labour ratio will increase continuously, as a result of which the value of $A$ in equation (1.40) and hence the level of real output in equation (1.39) will keep expanding. In contrast, if, in equation (1.34), we observe $sf(k, A) > \overline{n}k \Rightarrow \dot{k} > 0$ initially, the rise in $k$ will cause the term $\overline{n}k$ to increase at the constant rate $\overline{n}$, whereas the term $sf(k, A)$ will increase at a *decreasing* rate because of one of the basic properties of the production function $f(k, A)$: the law of diminishing marginal returns ($f_{kk} < 0$). As a result, the difference between $sf(k, A)$ and $\overline{n}k$ will fall continuously as $k$ rises, and eventually be eliminated so that $sf(k, A) = \overline{n}k \Rightarrow \dot{k} = 0$. Looking again at equation (1.43) we see the basis for the contrary result in the NEGT model: when $f(k, A) = ak$, $f_{kk} = 0$. The NEGT model has thus succeeded in eliminating the law of diminishing returns (which, it should be noted, most HGT models do not embody from the outset).

The consequences for growth outcomes are profound. With $f_{kk} < 0$, we must eventually have $\dot{Q} = f_k\dot{k} = 0$ (because we must eventually have $\dot{k} = 0$), so that $0 = \dot{Q}/Q = \hat{Q} = y - \overline{n} \Rightarrow y > \overline{n}$, whereas with $f_{kk} = 0$ and hence $\dot{Q} = f_k\dot{k} > 0$ (because $\dot{k} > 0$ for all values of $k$), $f_k\dot{k} = \dot{Q}/Q = \hat{Q} = y - \overline{n} \Rightarrow y = f_k\dot{k} + \overline{n} > \overline{n}$. The economy continually surpasses the exogenous rate of growth that would be achieved in the Solow model. On the face of it, this is because of the capacity of the NEGT model to endogenously generate technical change.[38] But as the foregoing discussion demonstrates, what is ultimately responsible for our result is the absence of diminishing marginal returns to capital in the NEGT model.[39]

Referring back to the basic structure of the $aK$ model in equations (1.39) and (1.40) reveals exactly how this has been achieved. As firms accumulate capital, output is enhanced directly in keeping with the marginal product of capital – which, in the context of equation (1.39), is diminishing in the level of the capital stock, as in conventional neoclassical production theory. But the accumulation of capital then has a *second*, indirect effect on output, by enhancing the value of $A$ in equation (1.40) and hence the quantity of labour in efficiency units in equation (1.39). It is the combination of these direct and indirect effects that makes output proportional to the capital stock as in equation (1.41) and thwarts the onset of diminishing marginal returns. The indirect effect of capital accumulation acting via equation (1.40) can be considered an externality or 'spillover' effect, which explains frequent reference to externalities and/or spillovers in the NEGT literature.

Suppose, however, that technical progress is crucially dependent on the supply of labour devoted to research and development (R&D) activities, so that instead of equation (1.37) we have

$$Z_N = Z_N(n) \tag{1.45}$$

Solving the NEGT model under these new conditions now yields

$$y = q(Z_N(\bar{n})) + \bar{n} \tag{1.38'}$$

The result is 'semi-endogenous' growth (Jones, 1995, 2002), so-called for two reasons. On the one hand, the rate of growth depends on the rate of technical progress as explicitly modelled in (1.36), a feature that the result in equation (1.38′) shares with the basic NEGT model. On the other hand, however, the rate of technical progress and hence the rate of growth is no longer obviously amenable to change by either private or public decision makers within the economy, since both depend ultimately on the rate of growth of the population which is not a (narrowly defined) economic variable.[40] The resulting exogeneity of the growth rate is instantly recognizable as a distinguishing feature of the Solow model discussed earlier.[41] The fact that the result in (1.38′) satisfies the first but not the second sense in which the rate of growth in (1.38) is endogenous – or in other words, that it hybridizes the results in (1.32) and (1.38) associated with the Solow model and NEGT, respectively – is what gives 'semi-endogenous' growth its name.

The derivation of this semi-endogenous growth result is easily demonstrated by introducing a small modification to equation (1.40). Hence suppose we write

$$A = \beta\left(\frac{K}{N}\right)^{\gamma} \tag{1.40'}$$

Note that if $\gamma = 1$ we are back to the specification of (1.40) in the $aK$ model. But this is now a special case. Suppose that instead $\gamma < 1$, a typical assumption for an exponent in a production function in neoclassical production theory (except here applied to knowledge generation). To see the implications, we first rewrite equation (1.40) as

$$Q = \frac{Y}{N} = A^{1-\alpha}\frac{K^{\alpha}N^{1-\alpha}}{N^{\alpha}N^{1-\alpha}} = A^{1-\alpha}k^{\alpha} \tag{1.46}$$

Substituting equation (1.40′) into (1.46) now yields an intensive production function of the form

$$Q = (\beta k^{\gamma})^{1-\alpha}k^{\alpha} = ak^{\alpha+\gamma(1-\alpha)} \tag{1.47}$$

Finally, note that on the basis of (1.47), we have

$$f_{kk} = a(\alpha + \gamma[1-\alpha] - 1)(\alpha + \gamma[1-\alpha])k^{\alpha+\gamma(1-\alpha)-2} < 0$$

since $\gamma < 1 \Rightarrow \gamma(1-\alpha) < (1-\alpha) \Rightarrow \alpha + \gamma(1-\alpha) - 1 < 0$. In other words, Jones has restored the law of diminishing marginal returns.

Recall that, as previously demonstrated, the law of diminishing returns ensures that the economy cannot permanently grow faster than the exogenously given rate of expansion associated with the Solow model. This Solovian (exogenous growth) result is thus restored in the semi-endogenous growth model, despite its explicit modelling of the process of technical change. The intuition is straightforward: equation (1.40′) insists that, as a general case, the law of diminishing returns applies to the production of $A$ (sometimes referred to in NEGT as 'knowledge').[42] The intuition is that while the production of knowledge benefits from additional resources devoted to research, as ever more resources are added it becomes harder and harder for researchers to produce new ideas (rather than just replicating old ideas, for example).

Thus, the knowledge-producing sector represented by equation (1.39′) is no more immune to being subject to the standard neoclassical laws of production than the goods-producing sector represented by equation (1.39). However, we may certainly question whether knowledge generation is inevitably subject to diminishing returns as Jones claims. In fact, there may be increasing returns in the early stages of development of any given technological paradigm, such as the steam engine or information technology, when the creation of new innovations calls forth additional innovations at an ever-more-rapid pace. And even if diminishing returns eventually set in for any given technological paradigm, they may be offset by the arrival of new technologies that displace older ones (such as aeroplanes displacing railways for long-distance travel, or personal computers and laptops replacing earlier mainframe computers). Hence, it is not possible to say whether $\gamma < 1$ or $\gamma = 1$ is the more general case, and it is even possible that $\gamma > 1$ for significant periods of time.

### 1.4.2 The heterodox tradition: classical-Marxian and post-Keynesian theory

In principle, the term HGT can be used to refer to anything that does not fit into the rubric of neoclassical growth theory as outlined in the previous subsection. In this book, however, the term is used in a narrower and more focused sense. Specifically, it refers to models in the classical-Marxian and post-Keynesian traditions that, at their cores, emphasize the importance of class conflict over the distribution of income and the principle of effective demand (respectively) as central to capitalist macrodynamics.[43] As noted earlier in this chapter, there are various examples of both commonality and difference between classical-Marxian and post-Keynesian theories of growth and distribution. These will become more apparent in subsequent chapters where we explore models associated with these traditions in greater detail. To begin with, however, and in order to better effect comparison and contrast with the neoclassical tradition, we start with a more generic representation of the growth process designed to capture its general properties as envisaged by HGT.

#### *The canonical heterodox model*

The basic tenets of HGT can be represented by a canonical model of the form

$$y_N = q + n \tag{1.24}$$

$$q = \bar{q} \tag{1.25}$$

$$n = \bar{n} \tag{1.26}$$

$$y = y(Z_H) \tag{1.48}$$

$$Z_H = \bar{Z}_H \tag{1.49}$$

where $Z_H$ is a vector of variables that determines *either* the rate of growth of saving (classical tradition) *or* the level and/or the rate of growth of autonomous demand (post-Keynesian tradition), and all other variables are as previously defined. It should be noted immediately that equations (1.25) and (1.26) are used here strictly for the sake of simplicity as we develop our initial representation of HGT: there is a long tradition in both classical-Marxian and post-Keynesian strands of HGT of regarding both $q$ and $n$ as endogenous to the growth process. So important is this theme that we will return to address it immediately below.

Solving the model in equations (1.24)–(1.26) and (1.48)–(1.49) now yields *two* distinct growth rates. Hence solving (1.24)–(1.26) yields

$$y_N = \overline{q} + \overline{n} \tag{1.50}$$

while substituting (1.49) into (1.48) we arrive at

$$y = y(\overline{Z}_H) \tag{1.51}$$

Equation (1.50) represents the natural rate of growth, while equation (1.51) is the actual (steady-state equilibrium) rate of growth. Per Harrod (1939), and unlike the NGT approach discussed earlier, in HGT the natural rate of growth is limited to describing a 'ceiling' that sets an upper limit to the actual (equilibrium) rate of growth in the long run. The actual (equilibrium) rate of growth is determined independently of the natural rate, as in equation (1.51) – either by the determinants of demand formation (in post-Keynesian models) or capital formation (in classical-Marxian models). Equations (1.50) and (1.51) immediately draw attention to a generic feature of growth outcomes in HGT: the distinct possibility (indeed, likelihood) that we will observe $y \neq y_N$, since $\overline{q}, \overline{n}$ and $\overline{Z}_H$ are determined independently of one another. The inequality of the equilibrium and natural rates of growth is called the first Harrod problem.[44] This result permits a range of growth outcomes in HGT that are not observed in neoclassical models, where as we have seen the actual long-run rate of growth always corresponds to the natural rate (and, indeed, always entails full employment).

In HGT, either by virtue of a demand constraint (in post-Keynesian models) or a capital constraint (in classical-Marxian models), the economy's expansion path will most likely be characterized by a chronic failure to fully utilize the productive potential of labour.[45] As such, heterodox growth models distinguish between two different types of growth regimes. The first is characteristic of labour-constrained – or, following Robinson (1956), 'golden age' – economies, where $y = y_N$ and growth conforms to the same pattern that would be observed in the neoclassical models outlined earlier (except, unlike the neoclassical models, and as noted earlier, there is no assumption that the economy operates at full employment in HGT). The second regime is characteristic of non-labour-constrained or 'dual' economies (Skott and Ryoo, 2008), where $y = y(\overline{Z}_H) \neq y_N$ and the first Harrod problem is observed. Dual economy models, which originated with Lewis (1954), assume the existence of a periphery from which surplus labour can be drawn into the modern or industrial sector as required (and, in the original version, at a constant real wage) whenever growth in the latter increases. Hence, even if the overall rate of growth of the labour force is given, the rate of labour force participation in the industrial sector varies, so that the unemployment rate in the industrial sector need not exhibit a secular trend even when $y \neq y_N$. In such economies, the *logical* boundaries of the rate of labour force participation in the industrial sector are understood not to impose a *practical* constraint on the growth process.[46]

The classical-Marxian and post-Keynesian variants of HGT produce substantively different solutions for the equilibrium rate of growth in (1.51), this being a product of the different (supply- versus demand-side) drivers of growth that they posit as constituents of the vector $Z_H$. This will become clear in the chapters that follow as we develop successive growth models of classical-Marxian and post-Keynesian pedigree. At this juncture, and in anticipation of the analysis that follows in later chapters, it suffices to draw attention in passing to two features of classical-Marxian and post-Keynesian growth outcomes that distinguish these traditions from each other and (in turn) from NGT. These concern the sensitivity of the equilibrium rate of growth to the saving rate and to the distribution of income.

Saving out of profit is the wellspring of capital accumulation and growth in the classical-Marxian vision, as a result of which either an increase in the propensity to save out of profits or a redistribution of income towards profit (which is the unique source of saving in the classical-Marxian approach) will stimulate aggregate saving and investment by capitalists, and hence capital accumulation and growth. Therefore, in classical-Marxian analysis, the equilibrium rate of growth varies directly with both the profit share of income and the propensity to save out of profits. In the classical-Marxian tradition, then, growth is *profit-led* (a redistribution towards profits boosts growth) and *supply-driven* (an increase in saving funds additional capital accumulation, which boosts growth).

In post-Keynesian models, in contrast, the equilibrium rate of growth varies inversely with the propensity to save out of profits. This result is a long-run analogue of Keynes's famous paradox of thrift, and is a reflection of the fact that, in post-Keynesian models, growth is *demand-driven*. Because an increase in the saving propensity comes at the expense of reduced consumption spending, it must therefore diminish aggregate demand, which ultimately retards the rate of growth. In addition, some (but not all) post-Keynesian models produce the result that the equilibrium rate of growth may vary inversely with the profit share of income: a decrease in the profit share – or in other words, an increase in the wage share of income – will stimulate long-run growth, which therefore can be *wage-led*.[47] This possibility arises from the demand-led nature of growth in post-Keynesian theory, if a redistribution of income towards wages stimulates the consumption spending of workers (whose wage income rises) by an amount that exceeds the decline in consumption by capitalist households and also any possible reduction in investment by firms (as profit income falls). This, in turn, is possible because working households are assumed (realistically) to have a much higher marginal propensity to consume than capitalist households, owing to their lower absolute income levels.[48] However, it is also possible for demand-driven economies to have profit-led growth, if the stimulus from higher profitability to investment exceeds the loss in consumption demand, and the possible negative impact of higher labour costs on net exports make this outcome even more likely in highly open economies. These distinctions will be discussed at length in Chapters 4–5.

Heterodox growth models thus differ as to whether long-run growth is a supply- or demand-driven process, and as to the precise role of distribution in the determination of growth. However, in the neoclassical model used in the previous subsection to exemplify NEGT,

$$\frac{dy}{ds} = a > 0$$

and

$$\frac{dy}{d\pi} = 0$$

where $\pi$ is the profit share. The first result reasserts our earlier claim that growth is supply-driven in neoclassical growth theory (even when the growth rate is endogenous) and is in keeping with the result associated with classical-Marxian analysis. The second implies that distribution has no causal effect on growth in standard neoclassical models.[49] This latter claim constitutes another important difference between NGT (either original or NEGT) and HGT, since for HGT, regardless of whether growth is profit-led or wage-led, the distribution of income is more likely to play an important role as a causal determinant of the long-run growth rate.[50]

### *Endogenizing the natural rate of growth in HGT*

An important extension of the canonical heterodox growth theory model described above involves treating the natural rate of growth in equation (1.50) as endogenous to the *actual* rate of growth, $y$ (see León-Ledesma and Thirlwall, 2000, 2002; León-Ledesma and Lanzafame, 2010; and Perrotini-Hernández and Vázquez-Muñoz, 2017 for empirical evidence). For example, labour productivity growth can be described as a function of actual output growth by appeal to the *Verdoorn law*, so-named because of Verdoorn's (1949) pioneering empirical finding of a positive correlation between the growth rates of labour productivity and output in manufacturing across countries.[51] In this case, the canonical heterodox growth model outlined in the previous section must be rewritten as

$$y_N = q + n \tag{1.24}$$

$$n = \bar{n} \tag{1.26}$$

$$y = y(Z_H) \tag{1.48}$$

$$Z_H = \bar{Z}_H \tag{1.49}$$

$$q = q(y) \tag{1.52}$$

where equation (1.52) – which replaces equation (1.25) – represents an aggregative version of the Verdoorn law.[52] Solving the model once again yields two different growth rates. As before, combining (1.48) and (1.49) produces equation (1.51). But combining (1.24), (1.26) and (1.52) and bearing in mind the result in (1.51), we now find that

$$y_N = q(y(\bar{Z}_H)) + \bar{n} \tag{1.53}$$

Equations (1.51) and (1.53) once again distinguish between the actual and natural rates of growth, respectively. But rather than acting as an exogenously given growth ceiling, the natural rate in (1.53) now sets a maximum value of the growth rate at any point in time that is directly influenced by the equilibrium growth rate. This creates a form of path dependence in the model, in the sense that the natural rate of growth will depend on the actual growth history of the economy, as captured by variations in the equilibrium rate of growth in (1.51) (see Setterfield, 2009, pp. 42–4).[53]

## 1.5 Reconciling aggregate demand and aggregate supply in the theory of long-run growth

### 1.5.1 Say's law versus 'Say's law in reverse'

As previously noted, both neoclassical and classical-Marxian growth analyses conceive of growth as a supply-driven process, whereas post-Keynesian models describe growth as demand-driven. The result is that growth models typically focus on modelling either the supply side or the demand side to the neglect of the other, which is assumed to passively adjust in order to accommodate the dictates of either the demand-side or the supply-side drivers of the growth process. In this way, growth models tend to present stylized views of the long run as being characterized either by Say's law (demand always accommodates supply) or 'Say's law in reverse' (supply always accommodates demand) (Cornwall, 1972).

In fact, neither view is entirely satisfactory. Whatever the driver of the growth process (demand or supply), there ought to be – and sometimes, as we will see, *needs* to be – an explicit account made as to how and even whether the accommodating factor (supply or demand – or possibly both) adjusts. Otherwise, we cannot be sure that the required accommodation will even take place, which event would, of course, constrain and therefore influence final growth outcomes.

### 1.5.2 Interactions between demand and supply

We have so far argued that HGT generally admits two distinct growth regimes, corresponding to golden age and dual economy conditions. Further consideration, however, suggests that in a steady-state equilibrium framework, only golden age conditions are truly sustainable.

To see this, consider the ratio of the actual and full employment level of output at any point in time, $Y/Y_N$. This ratio can be written as

$$\frac{Y}{Y_N} = \frac{\frac{Y}{L}}{\frac{Y_N}{N}\cdot\frac{N}{L}} = \frac{L}{N} = e \tag{1.54}$$

since $Y/L = Y_N/N = a_0$. It therefore follows from (1.54) that

$$\hat{e} = (y - y_N) \tag{1.55}$$

The employment rate, *e*, is, however, a bounded variable, and as discussed earlier in this chapter, the only plausible constant proportional rate of growth of a bounded variable in the long run is zero. Imposing this condition on (1.55) yields

$$y = y_N \tag{1.56}$$

Equation (1.56) is observationally equivalent to equation (1.27) in NGT, but its interpretation is quite different. Recall that from a neoclassical perspective, equation (1.27) is a *causal* statement: $y_N$ causes $y$, so that growth is determined on the supply side by the rate of expansion of full-employment output. As derived above, however, equation (1.56) states only that, in a long-run, steady-state equilibrium, the actual rate of growth is constrained to grow at the same rate as the full-employment growth rate in order to prevent the employment rate from exceeding its logical bounds. In other words, only the labour-constrained or 'golden age' variant of the canonical HGT model is ultimately sustainable as a representation of steady-state equilibrium growth outcomes in the long run. As noted earlier, dual economy variants of HGT work on the assumption that full utilization of labour resources does not impose a practical constraint on economic activity. But mature or advanced capitalist economies do periodically operate at or near full employment, so in these cases it is not safe to ignore the labour constraint on growth and equation (1.56) must be considered binding. Note that, even with an endogenous natural rate of growth, the model developed in the previous subsection does not break free of this issue. Hence, except in the special case where comparison of (1.51) and (1.53) reveals that

$$y(\bar{Z}_H) = q(y(\bar{Z}_H)) + \bar{n}$$

we would still expect the first Harrod problem ($y \neq y_N$) to arise.

In view of all this, it is not surprising that various strands of the HGT literature have sought to identify mechanisms by which we might come to observe $y = y_N$, and hence the onset of golden age conditions, without abandoning the principle that the actual rate of growth is determined independently of $y_N$ by the equilibrium solution for $y$ in equation (1.51) (for example, Cornwall, 1972; Palley, 2002c; Setterfield, 2006b). This literature is post-Keynesian in orientation. It is concerned with the possibility that the equilibrium rate of growth in (1.51) is demand-determined, and that it will remain so even in the long run. Otherwise, it would be straightforward to invoke the notion of a labour-constrained economy as discussed earlier, and in so doing to accept that we must eventually accept $y = y_N$ as a causal statement. Coupled with the traditional interpretation of $y_N$ as an exogenous natural rate of growth determined on the supply side, however, this would involve giving up the notion of demand-led growth.[54]

Therefore, HGT models that accept the importance of equation (1.56) as a long-run equilibrium condition nevertheless often begin with the proposition that the natural rate is endogenous to the actual rate of growth. These models are then structured so that the actual and natural rates of growth will be equalized (bearing in mind the endogeneity of the latter to the former) in equilibrium. The question that remains is, how is this reconciliation achieved? Does the natural rate of growth adjust to accommodate the actual rate, or vice versa? It transpires that different heterodox models have been proposed that have answered this question in different ways, some of which stress the natural rate of growth as the adjusting variable, others emphasizing the accommodating role of the actual rate. For those HGT models that make the natural rate of growth an adjusting variable, there are also differences between the classical-Marxian tradition, in which endogeneity of labour force growth $n$ is a key feature (as discussed in Chapter 2), and the Kaldorian branch of post-Keynesian theory, in which technological progress and hence the rate of labour productivity growth $q$ is the key adjusting factor (as discussed in Chapters 8 and 10). Rather than further extending our canonical heterodox growth model to accommodate adjustment mechanisms of these sorts, however, we will instead defer further discussion of the reconciliation of demand and supply in heterodox growth theory to the chapters that follow, where it will taken up on a case-by-case basis in the context of specific variants of HGT.

One further comment is in order at this stage, however. Despite its identification above with heterodox growth theory, some models in the neoclassical tradition have also recently taken up the theme of the natural rate of growth being endogenous to the actual rate of growth. Recall that, although the natural rate of growth is usually understood as an equilibrium towards which the actual rate of growth eventually adjusts in the conventional NGT approach, it is possible in neoclassical theory for the actual rate of growth to depart from its (equilibrium) natural value due to an exogenous shock. Such events, creating departures from the long-run trend rate of growth, would be associated with the periodic recessions that characterize the business cycle. Hence interest in the endogeneity of the natural to the actual rate of growth in neoclassical growth theory can be associated with the discovery by neoclassical growth theorists (or at least some of those working in the NEGT tradition) that trend and cycle may interact, with cyclical disturbances causing departures from trend now associated with playing a critical role in the determination of the subsequent trend itself (see Gaggl and Steindl, 2008; Cerra and Saxena, 2017). This way of expressing the matter differs from the fundamental inequality of the equilibrium and natural rates of growth posited by the first Harrod problem in heterodox growth theory, but would cause no great consternation among heterodox growth theorists well versed in the notion that trend and cycle interact (with the latter playing a causal role in determining the former). Hence consider, for example, the sentiments expressed by Kalecki in the mid-twentieth century:[55]

> the long-run trend is but a slowly changing component of a chain of short-period situations; it has no independent identity, and the two basic relations mentioned above [the multiplier and the accelerator] should be formulated in such a way as to yield the trend cum business-cycle phenomenon. (Kalecki, 1968, p. 263)

In fact, the similarities between NGT and HGT that result from consideration of the interaction of trend and cycle reach far beyond a shared vision. To see this, we begin by rewriting the stylized NEGT model from section 1.4.1 as follows:

$$y_N = q + n \tag{1.24}$$

$$n = \bar{n} \tag{1.26}$$

$$q = q(Z_N) \tag{1.36}$$

$$y = y_N + \zeta_1(Y - Y_N) + \varepsilon \tag{1.57}$$

$$\dot{Z}_N = \zeta_2(Y - Y_N) \tag{1.58}$$

where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ describes a transitory disturbance term that is normally distributed about a mean of zero with a variance of $\sigma_\varepsilon^2$, $\zeta_1' < 0$ and $\zeta_2'$ can have either sign. In keeping with the supply-determined character of neoclassical macroeconomics, full-employment output $Y_N$ is also understood as the equilibrium level of output at any point in time. As described in (1.57), which replaces equation (1.27) in the 'baseline' NEGT model developed earlier, $\varepsilon \neq 0$ can separate the actual rate of growth from the natural rate of growth at any point in time, a difference that will be propagated subsequently by the economy operating (temporarily) off its full-employment output path ($Y \neq Y_N$).[56] Equation (1.58), meanwhile, replaces equation (1.37) in the baseline NEGT model. It suggests that the supply-side determinants of technical progress are sensitive to the difference between $Y$ and $Y_N$.

Equation (1.58) is also consistent with the NEGT hypothesis that aggregate fluctuations can have *either* a positive *or* a negative impact on the pace of growth, depending on whether the activities responsible for improving aggregate productivity, as captured by equation (1.36), are complements to or substitutes for the process of producing goods and services (Blackburn, 1999, p. 68).

Hence in one strand of NEGT models (for example, Aghion and Howitt, 1992), recessions stimulate growth by 'shaking out' inefficient practices (Schumpeter's creative destruction) and by encouraging firms to devote more resources to innovation, so that $\zeta_2' < 0$. In another strand, however (for example, Martin and Rogers, 1997), increased engagement in production stimulates learning by doing, which has positive spillover effects on the productivity-enhancing activities modelled in (1.36) – so that $\zeta_2' > 0$.[57] It should also be noted that the mechanisms emphasized above are not mutually exclusive. Hence there exists a class of NEGT models in which $\zeta_2'$ can take either sign, depending on whether 'internal' learning (which results from explicitly devoting resources to learning and is countercyclical, as in creative destruction models) dominates 'external' learning (which results from learning by doing, and is procyclical) – see, for example, Blackburn (1999), Blackburn and Galindev (2003) and Galindev (2009).

Under the equilibrium conditions $\varepsilon = 0$ and $Y = Y_N$, we get $y = y_N$ from (1.57) and $\dot{Z}_N = 0$ from (1.58). Solving the model under these conditions yields

$$y = q(Z_N^*) + \bar{n} \tag{1.59}$$

where $Z_N^*$ denotes the initial equilibrium value of $Z_N$. This solution is superficially similar to that of the standard NEGT model, in the sense that the potential rate of growth appears to determine the actual rate of growth. However, note that it is now the case that if $\varepsilon \neq 0$, the result will be $y \neq y_N$ in (1.57). Suppose, for example, that the economy experiences a transitory shock $\varepsilon < 0$. This implies that $y < y_N$ in (1.57), as a result of which we will subsequently observe $Y < Y_N$ as the economy slips into recession.[58] Two things will now happen. First, $Y < Y_N$ implies $y > y_N$ in (1.57): growth increases above potential as the economy immediately begins to recover from recession. Second, as a result of equation (1.58), and depending on the sign of the derivative of this equation, either

$$\dot{Z}_N < 0 \Rightarrow \downarrow q \Rightarrow \downarrow y_N$$

or

$$\dot{Z}_N > 0 \Rightarrow \uparrow q \Rightarrow \uparrow y_N$$

In either event (and given suitable stability conditions), the gap between $Y$ and $Y_N$ will narrow as a result of $y > y_N$, and when the steady state is regained (with $\varepsilon = 0$, $Y = Y_N$, and $\dot{Z}_N = 0$) we will have

$$y = q(Z'_N) + \bar{n} \tag{1.60}$$

where $Z'_N \neq Z^*_N$, so that the steady-state growth rate is either higher or lower (depending on the sign of the derivative of equation 1.58) than its value prior to the transitory shock with which we began. Hence, despite the appearance of a conventional neoclassical result in (1.59), the potential rate of growth (and hence the long-run actual rate of growth) is now sensitive to departures of the actual rate from its prior trend – so that the potential rate of growth is essentially path dependent.

Moreover, if $\varepsilon$ captures the real effects of *demand* shocks (due, for example, to nominal rigidities in the goods market), the result is that the long-run equilibrium rate of growth will be demand-determined, even though it satisfies the neoclassical causal interpretation of equation (1.27).[59] Mainstream studies that analyse how the growth of productivity (and hence potential output) can be reduced by persistent demand shortfalls have become more prominent in the aftermath of the financial crisis and Great Recession of 2007–09. A few examples include Reifschneider et al. (2015), who emphasize endogenous negative responses of labour supply and business investment, and Anzoategui et al. (2016) and Benigno and Fornaro (2018), who emphasize endogenous declines in innovative activity. These are, of course, very post-Keynesian results, but it should be noted that such results are found only in a very small slice of the NEGT literature, whereas they are prominently featured in HGT (especially models in the Kaldorian strain).

## 1.6 Why study heterodox theories of growth and distribution?

As the discussion of Table 1.1 in section 1.3 suggested, there are certain important *similarities* between neoclassical and heterodox growth theories. Neoclassical and classical-Marxian theories are allied in their essentially supply-side visions of the long run, with only post-Keynesian theorists insisting on the primacy of demand in the determination of long-run growth.

Meanwhile, post-Keynesians are critical of both neoclassical and classical-Marxian visions of the long run, which they see as too deterministic compared to their own (path-dependent) vision of a long run in which 'history matters'. In other words, it would be inaccurate to suggest that there is a simple, bimodal contrast between NGT and HGT.

Furthermore, as the discussion in the preceding section suggests, some of the emerging similarities between NGT and HGT have come about as a result of the NEGT branch of NGT moving towards themes traditionally emphasized in HGT, such as the importance of aggregate demand and income distribution in the determination of long-run growth, and the path dependence of long-run outcomes. With regard to demand effects, for example, the grounds for this claim of 'convergence' can be elucidated further if, in the NEGT model in equations (1.24), (1.26), (1.36), (1.57) and (1.58) above, we interpret $\varepsilon \neq 0$ as arising from demand shocks and $\zeta_2' > 0$ as the product of induced, factor-biased technical change (see Foley et al., 2019, chap. 8; Duménil and Lévy, 1995, 2003). According to the latter process, if the economy 'runs hot' $(Y > Y_N)$, the resulting low rate of unemployment will produce a 'profit-squeeze' (a fall in the profit share of income and hence, given the output–capital ratio, the rate of profit). This, in turn, induces firms to engage in technical change designed to displace labour and alleviate the profit-squeeze. In so doing, the capital–labour ratio and hence (given the output–capital ratio) labour productivity increases. The result, then – as originally shown by Dutt (2006a, 2010b) – is a neoclassical model with a classical technical progress function that produces post-Keynesian (demand-determined) long-run growth outcomes! This 'grand synthesis' combines classical and neoclassical insights about the supply side while allowing for fundamentally post-Keynesian (demand-determined) long-run growth outcomes.

The discussion in this section suggests that even when we confine attention to the abstract mechanics of steady-state models, there exist important examples of theoretical overlap between NGT and HGT. So why study HGT at all if (some) mainstream models can effectively 'mimic' (some of) the main results of the heterodox approach?

One answer to the question just posed lies in the fact that the theoretical overlap between NGT and HGT is *incomplete*, while recognition of such partial overlap as does exist is *asymmetric*. Starting with this last point, it is important to understand that there is little or no genuine interaction between NGT and HGT on such themes as the interaction of trend and cycle, and the possibility that variations in aggregate demand influence long-run growth.

Such recognition of theoretical overlap that exists is highly asymmetric, and is found disproportionately in the HGT literature.[60] This lack of genuine interaction is important in its own right: it suggests that it is only through immersion in HGT that students are likely to become aware of the relationship between recent developments in NGT and themes – both established and emergent – in HGT.

The fact that the theoretical overlap between NGT and HGT is incomplete is also as suggested by the references to 'some' mainstream models and 'some of' the main results of HGT in the question posed above. Hence, for example, the theme of trend–cycle interaction and the idea that demand matters for long-run growth, as discussed in section 1.4.2, is only a feature of a particular (and relatively marginal) branch of NEGT. It is not part of the 'hard core' of NGT thinking, and is not emphasized 'front and centre' in standard accounts of NGT or even NEGT models. These themes (the inseparability of trend and cycle and the central role of aggregate demand and income distribution in the determination of long-run growth) *are*, however, part of the 'hard core' of HGT, at least in its post-Keynesian variants. These themes are not emphasized in equal measure in every individual model associated with HGT, which, like NGT, is internally heterogeneous to a considerable degree. But they are taken seriously by all contemporary contributors to HGT, as evidenced by the routine engagement among scholars associated with HGT on these themes. In short, a student is only guaranteed to encounter themes associated with HGT by studying HGT itself.

Furthermore, notwithstanding the elements of theoretical overlap between NGT and HGT, there remain important differences between these traditions. As a cursory reference back to Table 1.1 and the discussion thereof in section 1.3 demonstrates, there remain various themes uniquely associated with HGT that NGT has not absorbed, and in some cases (such as the value-theoretic foundations associated with HGT, and its admission of the possibility that conflict rather than harmony is characteristic of competition in capitalism) *cannot* absorb. There is also the matter of the obsession in contemporary NGT with methodological individualism – specifically, its insistence on deriving all results from 'microfoundations' based on intertemporal optimization by essentially omniscient households. This also has the unfortunate and unwarranted consequence of making households the exclusive decision-making drivers of growth, because firms are reduced to the passive status of being mere agents of their owners.

HGT instead builds models based on structural behavioural relationships that are themselves rooted in stylized facts – that is, abstractions based on observation of capitalism as it actually is, rather than as it is imagined to be. These methodological foundations are also anti-reductionist, recognizing categories such as social class, race and gender that provide 'macrofoundations' for individual behaviour. This suggests that a student is only guaranteed to encounter *all* of the themes associated with HGT – and hence *all* of the themes that are important for understanding long-run capitalist macrodynamics – by studying HGT itself.

As this brief discussion suggests, then, there are still potentially important lessons to be learned about the character of the long-run growth process that will only likely be learned, or else can only be learned, by devoting time to the study of HGT in its own right. This, essentially, is the 'value added' to a student's economics education of a book such as this one.

## 1.7 Conclusions and outline of the book

This chapter has identified the basic characteristics of the long-run growth analysis that will be explored in more depth throughout the remainder of this book. Apart from defining and exploring some key terms and concepts, it has focused on sketching out the basic structures of both neoclassical and heterodox growth theories. A number of simple structural models have been developed that describe the essential mechanics of growth in NGT and HGT. Interpretation of these models suggests that they are not simply a study in contrast: there are examples of theoretical overlap between the ostensibly competing neoclassical and heterodox growth traditions, and there is also considerable diversity within each of them. This overlap is far from complete or symmetrical, however, and in order to focus squarely on the heterodox tradition, the remainder of this book is devoted exclusively to developing, in depth, the various classical-Marxian and post-Keynesian models associated with HGT.

Part I of the book presents the core models of growth and distribution associated with the HGT tradition. Chapter 2 begins by focusing on the classical-Marxian approach. An equilibrium model of growth is constructed, subject to a constant state of technology, and different specific solutions are shown to follow from alternative closures that can be used to complete the model. Two enduring features of the model are revealed by studying its response to changes in saving behaviour and the distribution of income – specifically, that increasing the propensity to save generally boosts the equilibrium rate of growth, while increasing the share of profits in total income has a similar effect. Technological change is then introduced, and in particular Marx-biased technological change, which is understood to be simultaneously labour-saving but capital-using. Technological change of this sort is shown to be associated with Marx's famous prediction of the tendency towards a falling rate of profit in capitalism. The chapter also covers Ricardo's (1821 [1951]) analysis of the stationary state and Goodwin's (1967) neo-Marxian model of cyclical growth.

Chapter 3 is devoted to the neo-Keynesian models associated primarily with the Cambridge growth theorists Joan Robinson, Nicholas Kaldor and Luigi Pasinetti. It begins, however, with a discussion of Roy Harrod's macrodynamics, which are prototypically post-Keynesian in the sense that they arise from the independence of investment (by firms) from saving (by households). Harrod's macrodynamics are shown to give rise to the possibility that the equilibrium rate of growth may be unstable – that is, any departure from equilibrium will be self-reinforcing rather than self-correcting. This provides the basis for the neo-Harrodian models (and the contemporary Harrodian instability debate) that are taken up in Chapter 6. Next, the model of Kaldor and the Pasinetti and neo-Pasinetti theorems are briefly explored. These draw attention to the important role of income distribution in the Cambridge neo-Keynesian models, a theme that is amplified by the subsequent development of the Robinson model in which an otherwise post-Keynesian model of growth is shown to rely on prices (rather than quantities of output) responding to excess demand and supply in the goods market, as a result of which the distribution of income becomes the key adjusting variable enabling the model to achieve equilibrium. The chapter ends with a discussion of the model developed by Marglin (1984a, 1984b), which synthesizes themes associated with the classical-Marxian and neo-Keynesian analyses found in Chapters 2 and 3.

The complaint that neo-Keynesian models are 'insufficiently Keynesian' because of their emphasis on wages and prices (rather than quantities of output and employment) bearing the brunt of disequilibrium adjustment is often understood as the motivation for the neo-Kaleckian models, inspired by the work of Michał Kalecki, that are discussed in Chapter 4. This chapter begins by exploring the microfoundations of goods markets dominated by oligopolies and, in particular, the maintenance of excess productive capacity by firms, and the practice of markup pricing. A basic growth model in this tradition is shown to yield the unique result that growth is *wage-led* – that is, a redistribution of income towards wages will increase both the utilization of capacity and the rate of growth. Extensions of this basic model are, however, shown to complicate the picture, with the possibility of a result more in keeping with classical-Marxian analysis (in which a redistribution towards profit boosts growth) re-emerging. Ultimately, neo-Kaleckian models suggest that long-run growth may be either wage- or profit-led, a claim that has sparked a debate in HGT about the causal influence of distribution on growth that endures to this day.

So important is this debate, in fact, that it frames the first chapter in Part II of the book, which explores various extensions of and new directions in the thinking associated with the core models of growth and distribution that were the focus of Part I. Chapter 5 begins by considering a conflicting claims model of inflation in which the distribution of income – treated as an exogenous given in most of the neo-Kaleckian models in Chapter 4 – becomes endogenous. Incorporating conflicting claims into the neo-Kaleckian framework means that demand (as reflected in the rate of capacity utilization) affects distribution (the share of wages in national income) even as distribution affects demand, and the resulting interaction of demand and distribution is shown to give rise to a model of business cycles comparable to (but somewhat different from) the neo-Marxian limit cycle model of Goodwin (1967). The remainder of the chapter is then devoted to a survey of the empirical literature that has sought to establish whether real-world capitalist economies – either individually or collectively – have wage-led or profit-led demand and whether they exhibit a 'profit-squeeze' in distribution (a positive effect of increased demand on the wage share) or not. A key finding of this survey is that different methods of estimation seem to incline findings one way or the other, leading to often conflicting results even for the same countries.

Chapter 6 returns to the theme of Harrodian dynamics and the possibility that the equilibrium rate of growth is unstable. Neo-Harrodian models embrace this possibility but argue that the potential for divergence is limited and reversible, with the result that the economy experiences cyclical growth. Two archetypes of neo-Harrodian analysis are identified. The first postulates 'ceilings and floors' that contain and reverse movement away from the equilibrium growth rate. The second postulates mechanisms endogenous to the process of divergence itself that fetter and eventually reverse the forces of divergence, so that the economy follows a limit cycle trajectory around its equilibrium growth rate. Finally, Chapter 6 takes up the extensive Harrodian instability debate that has gripped modern HGT. According to some neo-Harrodians, the conditions for Harrodian instability can arise even in non-Harrodian growth models with seemingly stable equilibrium rates of growth. This has provoked responses from numerous classical-Marxian and post-Keynesian authors, who propose various mechanisms that purport to 'tame' Harrodian instability if and when it does arise.

Chapter 7 surveys four of the most recent themes to have emerged and flourished in HGT. The first of these is dimensions of income distribution other than the broad functional distribution between wages and profits (or labour and capital income broadly defined) emphasized elsewhere in this book. These include inequality among different recipients of labour income (for example, based on education, occupation, race or gender), a theme motivated by observed increases in wage inequality since the 1980s. In particular, managerial compensation has soared, and while official statistics include such compensation in measures of total wages and salaries, it is characteristically regarded as being closer to residual earnings (profits) in HGT. These concerns have spawned models in which either capitalists themselves, or else a third class of 'managers', receive a share of the wage bill. The existence of a gender gap is another important dimension of wage inequality. As an example, Chapter 7 covers a model in which the distribution of wage income between male and female workers affects economic performance in an export-led economy.

Two other themes taken up in this chapter are financialization, and models that produce profit-led growth outcomes without being formally profit-led (at least in the sense of the models outlined in Chapters 2–5). Financialization is a broad term used to characterize the relationship between firms as institutions and the suppliers of financial capital (lenders and equity owners, otherwise known as 'rentiers'). In particular, the concept of financialization draws attention to the rise to *dominance* of financial interests over industrial interests in capitalist economies. Various models are outlined that introduce either creditors or absentee owners into core HGT models of the type that populate Part I of the book. A key issue is how the balance of power between financial interests on the one hand, and the interests that represent the development of the firms' productive capabilities on the other, affect the rate of growth. The appearance of profit-led growth outcomes, meanwhile, is shown to arise in models that are not conventionally profit-led. Such outcomes can arise because of the interaction of the real and financial sectors in a process akin to Minsky's (1982, 1986) financial instability hypothesis, or because of the 'financialization of the household': the process whereby working class households have resorted to credit markets to supplement stagnant wage incomes in order to pursue their consumption aspirations so as to 'keep up with the Joneses'. Models of this nature demonstrate that patterns in the data that suggest capitalism is profit-led may, in fact, be the product of mechanisms that involve a different or more complicated relationship between distribution and growth.

Finally, Chapter 7 surveys the recent resurgence of interest in supermultiplier models. A concept originally associated with Hicks (1950), the supermultiplier has recently become a prominent feature of Sraffian growth models. Inspired by the work of Sraffa (1960), the principle objective of Sraffian (or neo-Ricardian) economics is to perfect the development of classical value theory and simultaneously repudiate the logical basis of neoclassical marginal productivity theory. Sraffian growth models, meanwhile, form part of a larger Sraffian project designed to integrate the classical theory of value and distribution with the post-Keynesian principle of effective demand. Its conceptual ties to Sraffian economics notwithstanding, interest in the supermultiplier concept has blossomed recently because of its appeal to sources of autonomous demand growth and the capacity of the latter to contain Harrodian instability. This is evident in some neo-Harrodian 'ceiling and floor' models, in which the growth of autonomous demand creates a 'floor' that prevents an economy subject to Harrodian instability from collapsing. The supermultiplier approach has also found favour with some neo-Kaleckians because of the capability of autonomous demand to tame Harrodian instability and restore the stability of equilibrium in a neo-Kaleckian model.

While the interplay of growth and distribution are a major preoccupation in HGT, the interplay of international trade and the growth of national economies are also prominent themes, thanks in large part to the neo-Kaldorian models that are the focus of Part III of the book. Ironically, the basis for neo-Kaldorian analysis is a supermultiplier, but one in which it is specifically the rate of growth of *exports* that drives the rate of growth of domestic output, in a model of regional growth. This neo-Kaldorian approach is first examined in Chapter 8, which focuses on modelling growth as a process of export-led cumulative causation (ELCC). Following a discussion of Kaldor's vision of growth as a demand-driven, non-equilibrium process, as encapsulated in his famous 'growth laws', a model of cumulative causation is developed. Extensions of the model are introduced to make its growth outcomes formally path dependent, and the possibility of reconciling the actual and potential rates of growth is discussed. The debate over the importance of relative prices in export-led growth and the so-called 'Kaldor paradox' is also considered. Finally, various criticisms of the neo-Kaldorian ELCC model are assessed, especially the inattention to the regional trade imbalances (balance of trade surpluses and deficits) to which its dynamics can give rise. This criticism gives rise to the neo-Kaldorian models of balance-of-payments-constrained growth (BPCG) that are the focus of Chapters 9 and 10.

The BPCG model outlined in Chapter 9 is designed to reconcile certain basic Kaldorian principles (that growth is a demand-driven, export-led process) with the notion that balance of trade deficits (and hence, by extension, surpluses elsewhere in the world) cannot persist indefinitely. This is because the capital-account consequences of trade deficits (growing foreign indebtedness and/or the sale of assets to foreigners) will not be tolerated, either by deficit countries themselves or else by their foreign creditors. It is shown that the original BPCG model of Thirlwall (1979) – which produces the outcome commonly known as 'Thirlwall's law' – can be extended to incorporate net financial inflows (thus relaxing the strict trade balance requirement of the original model), and that structural change will have a decisive effect on the equilibrium growth rate in a multisectoral BPCG framework. This latter observation revives the emphasis on structural change that was a key part of Kaldor's growth laws, but which was not emphasized in either the ELCC model or the original BPCG apparatus. The chapter also presents a BPCG model for two large countries, demonstrating the importance of repercussion (feedback) effects between them. Finally, it is shown that the process of cumulative causation and the role of relative prices can be reintroduced into a modified BPCG framework, despite their absence from the original Thirlwall model.

Chapter 10 provides further discussion of the BPCG model developed in Chapter 9, and in so doing draws the book to a close. Various critiques and defences of Thirlwall's law (or empirical tests thereof) are examined, some of which urge reconsidering potential influences on growth – such as relative prices – that were present in the original neo-Kaldorian model of ELCC, but are 'suppressed' in the standard formulations of Thirlwall's law. The endogeneity of the critical income elasticities (of imports and exports), in terms of which the equilibrium BPCG rate is defined, is also re-examined from both neoclassical and heterodox perspectives. The possibility of reconciling the actual and potential rates of growth within a BPCG framework is discussed, as are alternatives to the canonical BPCG framework that consider how balance-of-payments constraints may affect growth differently in countries of different sizes (small, medium and large). The chapter concludes by demonstrating the possibility of a 'grand synthesis' that combines insights from the ELCC and BPCG branches of neo-Kaldorian growth theory as well as neo-Kaleckian models, where different causal mechanisms are modelled as prevailing over different time horizons (short, medium and long run).

The book as a whole is intended to be 'foundational': chapters assume no prior knowledge of the subject matter, and models are developed from the 'ground up' so as to give the reader the clearest possible sense of the assumptions and 'inner workings' on which their final results and implications rest.

It is hoped that this approach, together with the breadth of models and approaches that are covered throughout the book, will suffice to equip a new generation of students to understand new contributions in this field and to further advance the development of heterodox growth theory.

## ? STUDY QUESTIONS

1) How do overarching visions of the growth process differ between NGT and HGT, and between the classical-Marxian and post-Keynesian variants of HGT?
2) Compare and contrast neoclassical and heterodox theories of production and technical change.
3) Discuss the relationship between the actual and natural rates of growth in HGT. Is there reason for heterodox theorists to seek to resolve any inequality between the actual and natural rates?
4) In what ways can NGT (especially NEGT) be seen to have incorporated themes traditionally associated with HGT? Why is study of HGT nevertheless warranted?

## NOTES

1 See, for example, Maddison (1991, 2008) on these and other basic facts of the historical growth record.

2 See, for example, Aghion and Howitt (2009) and Setterfield (2010) for recent overviews of NGT and HGT, respectively.

3 In general, $\dot{x}_i = dx_i/dt$ denotes the *rate of change* of $x_i$ at any instant of time *t*, but in the context of theories of long-run growth $\dot{x}_i$ will usually be positive – hence our description of it as the *increase* in $x_i$.

4 The term 'steady state' is sometimes defined in more restrictive terms than those outlined above, especially in NGT where it necessarily involves not just steady but balanced growth. The value of the less restrictive definition adopted here will become obvious in due course. In the meantime, note that as defined above, a steady-state equilibrium, which produces steady growth in a variable of interest, can be contrasted with a stationary state, in which a variable of interest takes on a constant equilibrium value.

5 The stability of equilibrium can be defined more broadly to include any non-divergent behaviour (such as a limit cycle). In common with most economic theory, we adopt a more restrictive definition in this book, so that stability is associated exclusively with convergence towards equilibrium.

6 The essential difference between these techniques is that whereas comparative statics applies to systems set up to describe the *levels* of variables, comparative dynamics applies to systems describing how variables change over time (for example, as the result of the proportional rate of growth associated with a steady-state equilibrium growth path).

7 An example of this latter phenomenon is the general manifestation of the so-called first Harrod problem in HGT, as discussed in section 1.3.2 below, which in turn gives rise to the HGT literature seeking to reconcile the steady-state equilibrium rate of growth with the potential rate of growth, as discussed in section 1.4.2.

8 See Chapter 8 as well as Gordon et al. (1982) and Maddison (1991).

9 See, for example, Cornwall (1977), Pasinetti (1981) and – in a more mainstream methodological framework – Rodrik (2014).

10 Just how quickly this adjustment occurs depends on the context. Disturbances associated with the business cycle, for example, might be eliminated quickly. Adjustment to secular trends, meanwhile, such as a change in the rate of growth of the population (something that is integral to neoclassical growth theory, as we will see in section 1.4) may take considerably longer.

11 We will return to this theme in detail in the next subsection.

12 Sunk costs are fixed costs that have already been incurred and cannot be recovered.

13 In the context of NGT, the term full employment can be treated as synonymous with the 'natural rate of unemployment' or non-accelerating inflation rate of unemployment (NAIRU).

14 The (maximum possible) level of employment associated with full employment will then be $L = e_{max}N$.

**15** Leontief was perhaps best known for his contributions to input–output analysis, a technique that (as its name suggests) traces the interdependencies within an economy between the outputs of various sectors and the inputs used to produce them (many of which are, of course, the outputs of other sectors). He won the Nobel Prize for his work in 1973. See Pasinetti (1977) for an exposition of input–output models.

**16** In Chapter 2, this type of innovation, which enhances the productivity of all factor inputs proportionately, will be defined as 'Hicks-neutral'. The other types of innovation defined there can also be modelled using neoclassical production functions by having separate shift factors multiplying the $K$ and $L$ terms inside $F(\cdot)$.

**17** See Foley and Michl (1999, pp. 123–7) for further discussion. Foley and Michl term the resulting 'pseudo-isoquant' described in the text a *fossil production function*, because it traces out a past history of technical change rather than a menu of contemporaneous alternative techniques.

**18** Thus, the implicit assumption in most of the book is that all labour consists of production workers whose labour time is proportional to current output. An exception will be found in Appendix 4.1 in Chapter 4, where we discuss the role of overhead or fixed labour costs; portions of Chapter 7 will also distinguish different types of labour (for example, capitalist-managers versus production workers).

**19** In the background, NGT models also implicitly assume some demand-side mechanism (such as fine-tuning of monetary and/or fiscal policies) that maintains maximum feasible employment and utilization.

**20** When we say 'constant' here, we mean with respect to the employment rate for labour and the utilization rate for capital. With technical progress, as discussed below, either ratio $a_0$ or $a_1$ may change over time. Also, the labour–output ratio for all workers becomes variable in the short run in the presence of fixed 'overhead' labour, as discussed in Appendix 4.1 in Chapter 4.

**21** Note that carrying excess capital (capacity) is costly: it involves incurring fixed costs of production or acquisition for a factor that lies idle (or partly idle) and thus generates no (or reduced) income. As such, firms may be equally unwilling to allow the actual rate of capacity to drift permanently below the normal rate, in which case the latter can also be associated with a *minimum* rate of capacity utilization, again in the long run. For more on the reasons for firms to have a desired or normal level of excess capacity, see Chapter 4.

**22** Since we have normalized $u_n$ to a value of one here, there is in effect some allowance for a 'normal' degree of underutilization of the capital stock in panel (a) of Figure 1.2, even if this is not immediately apparent.

**23** As noted earlier, production according to the principles of a Leontief production function will typically involve *some* quantity of productive resources lying idle. Also note that the early 'neo-Keynesian' models covered in Chapter 3 are more similar to the classical-Marxian approach than to later post-Keynesian models (such as neo-Kaleckian ones) in regard to the way production is modelled.

**24** The variable $n$ is sometimes also referred to as the rate of growth of the population. This is because

$$N = \frac{N}{Pop} Pop$$

where $Pop$ denotes the size of the population and $N/Pop$ is the labour force participation rate. The latter is commonly assumed to be constant in the long run (not least because, as a bounded variable, it cannot grow indefinitely at a constant rate – so that ultimately $N$ and $Pop$ can only grow at the *same* constant rates). This implies that the rate of growth of the labour force is the same as the rate of growth of the population in the long run – hence the interchangeable use of these terms in reference to $n$.

**25** There are important exceptions to this, not the least of which results from the possibility of 'Marx-biased' technological change that is capital-using, and so results in secular increases in the value of $a_1$. This will be taken up in the discussion of classical-Marxian analysis in Chapter 2.

**26** Although the American economist Solow received more attention (and a Nobel Prize), essentially the same model was invented independently and published in the same year by Australian economist Trevor Swan. See Chapter 3 for a discussion of Harrod's original model and Chapter 6 for more recent neo-Harrodian models.

**27** This is true because equation (1.15) incorporates (1.12), which is the growth rate form of (1.8), which in turn was derived from (1.7).

**28** As will become clear later in this chapter, other interpretations of equation (1.27) – as stating no more than an equilibrium condition consistent with long-run, steady-state growth – are possible.

**29** The neoclassical growth framework does allow for actual output to fall below potential output in the short run due, for example, to temporary negative demand shocks. But the consequences of these shocks are understood to be rapidly self-correcting as the economy reverts towards its equilibrium expansion path associated with the growth of potential output. See, however, section 1.4.2 below for a recent change in (some) neoclassical thinking in this regard.

**30** As previously mentioned, and as will become clear throughout this book, the neoclassical treatment of households as undifferentiated stands in marked contrast to their treatment in heterodox growth theory, where households are assumed to be stratified by social class relations. Hence while in neoclassical theory there exists a generic propensity to save from income (regardless of source), heterodox growth theory distinguishes between different types of income (rent, profits, wages) received by households in different social classes (rentiers, capitalists, workers), and ascribes different consumption and saving behaviours to these class-stratified households.

**31** In other words, both the quantity of labour ($N$) and the 'quality' of labour ($A$) now enter into the determination of total labour input, $AN$.

**32** This last result follows directly from the equilibrium condition $\dot{k} = 0$, which implies a constant (equilibrium) value of the capital–labour ratio $k^*$, and consequently a constant (equilibrium) level of output per worker, $Q^* = f(k^*, A)$. Note also that if the level of Q is constant at $Q^*$, then $\hat{Q} = y - \bar{n} = 0$, from which it follows that $y = \bar{n}$, which is consistent with the result previously stated in (1.32) given that we are assuming $\bar{q} = 0$.

**33** Note that the fixity of $a_0$ is not a general property of the Solow model, in which adjustment to $k^*$ would ordinarily be achieved by means of flexibility in the values of both $a_1$ and $a_0$.

**34** Indeed, Solow (1994, 2007) argues that its contributions to the economics of technical change are likely to prove to be the most enduring contribution of NEGT.

**35** See, for example, Aghion and Howitt (2009, pp. 13–18) for a survey of the essential varieties of NEGT. See Jones (2002, pp. 164–6) for an illustration of how the precise functional form of (1.36) can affect the results of NEGT even as the vector $Z_N$ remains unchanged.

**36** This is usually called the '*AK*' model in the literature, using an upper-case *A*, but given that we have defined *A* differently in this chapter, we use a lower-case *a* instead for the productivity of capital. Also, with apologies for any possible confusion, it should be noted that (on the assumption of full capacity utilization at a 'normal' rate $u_n = 1$) the '*a*' in this model equals $1/a_1$ in our heterodox model of production – which highlights the fact that both models assume constant rather than decreasing returns to capital.

**37** The assumption made here about saving behaviour abstracts from the process of intertemporal utility maximization, from which the saving behaviour of households is characteristically derived in contemporary NEGT. The reader is reminded that the focus here is on the basic mechanics of growth in competing neoclassical and heterodox traditions, for which appeal to intertemporal optimization is unnecessary. See also Solow (1994, 2007) on the perils of ornamenting neoclassical growth theory with behavioural assumptions based on dynamic optimization.

**38** Recall that we have assumed $\bar{q} = 0$ in the process of deriving results for the Solow model.

**39** As we shall see, this is a significant (rather than purely technical) observation.

**40** It is not, for example, thought to be influenced by household saving behaviour – so the link between the latter and the equilibrium rate of growth established in NEGT is, once again, lost in semi-endogenous growth theory.

**41** As in the Solow model, parametric change in, for example, the saving rate will produce 'level effects' (changes in the level of output per worker and output per capita in all future periods) in semi-endogenous growth models, even as it has no effect on the long-run growth rate.

**42** Note that equation (1.40′) implies $f_{kk} = (\gamma - 1)\gamma\beta k^{\gamma-2} < 0$, since $\gamma < 1$.

**43** Harris (1974), Asimakopulos (1975), Del Monte (1975), Rowthorn (1981) and Dutt (1984) mark the origin of the modern Kaleckian strand of the post-Keynesian literature (on which see Blecker, 2002a; Lavoie, 2010), while Dixon and Thirlwall (1975) and Thirlwall (1979) began the modern Kaldorian strand (on which see Blecker, 2013b; Setterfield, 2013a). These strands, in turn, trace their origins to first-generation post-Keynesian growth theory in the work of Robinson (1956, 1962) and Kaldor (1966a [1989]), respectively. There is also a modern Harrodian strand of post-Keynesian growth analysis that can be associated with the work of Fazzari (1985), Fazzari et al. (2013) and Skott (1989), although in keeping with the focus of Harrod himself on the instability of the equilibrium ('warranted') rate of growth, this analysis does not fit into a stable, steady-state equilibrium framework of analysis. While we set aside Harrodian concerns with instability in this chapter, then, we will nevertheless have reason to return to this theme in the chapters that follow. The modern classical-Marxian tradition, meanwhile, can be traced back through Foley (1986), Marglin (1984b) and Harris (1978) to authors such as Goodwin (1967) and Sweezy (1942). Foley and Michl (1999, 2010) and Foley et al. (2019) provide contemporary statements of this tradition. A good general reference on HGT that also involves comparison and contrast with neoclassical macrodynamics is Taylor (2004).

**44** We will return to discuss the Harrodian origins of the first Harrod problem in Chapter 3.

**45** Capital may also be chronically underutilized, but whether or not this is deliberate (that is, the utilization rate corresponds to a value that is chosen a priori by firms) and the degree to which underutilization of capacity can persist in the long run is a matter of debate between various schools of post-Keynesian and classical-Marxian growth theory. The nature of this debate will become evident in Chapters 2–6.

**46** Models of this genus are common in both classical-Marxian and post-Keynesian (especially neo-Kaldorian) macrodynamics. See Porcile and Lima (2010) for a recent example.

**47** In some versions, the wage-led growth result is accompanied by what is known as the 'paradox of costs', in which an increase in the real wage (or share of labour in national income) leads to an increase in the equilibrium or 'realized' rate of profit – a seemingly paradoxical result since wages are a cost of production and a rise in wages would therefore appear to be inimical to profitability. Wages are also a source of demand, however, so that a rise in wages funds an increase in the demand for consumption goods which in turn stimulates the realization of profits (provided that there is not an offsetting decrease in investment, as discussed in the text). See Lavoie (2014) and Chapters 3 and 4 for further discussion of these various paradoxes (of thrift, costs and wage-led growth).

**48** Under some circumstances investment spending may also be stimulated by these developments, owing to the operation of so-called accelerator effects, according to which investment is spurred by increases in output produced as firms attempt to adjust productive capacity to keep pace with the expansion of the market. See the discussion of this point in relation to neo-Kaleckian models in Chapter 4.

**49** Although distribution has no *causal* role in standard NGT or NEGT models, it may nonetheless be true that income distribution has to adjust endogenously in order for the economy to reach a long-run, steady-state equilibrium with growth at the natural rate. For example, in the Solow model, both the rate of return to capital $r$ and the real wage $w$ are 'factor prices' that have to adjust to the 'right' levels in order for the capital–output ratio to attain its long-run equilibrium level $a_1^*$ and for growth to occur at the natural rate. Also, there are some new variants of NEGT that do allow for distributional effects on long-run growth (see, for example, Halter et al., 2014).

**50** This observation should also be treated with some caution. First, in neoclassical macroeconomics (as in classical-Marxian analysis), a change in the real wage is associated with a change in the profit rate of opposite sign. This could, in principle, affect the propensity to save – a possibility that is eliminated from the model used to exemplify NEGT in the previous subsection by virtue of its inclusion of a simple proportional saving function in which the propensity to save is constant. This having been said, empirical evidence does not suggest that changes in the real interest rate (synonymous with the profit rate in neoclassical macroeconomics) have a marked effect on saving behaviour (see Campbell and Mankiw, 1989; Carroll and Summers, 1991). Second, not all post-Keynesian models posit a *causal* role for distribution in the determination of long-run growth. In some neo-Keynesian models, for instance, the distribution of income is, instead, an *adjusting* variable: changes in the profit share facilitate the adjustment of the economy between equilibrium states following a change in investment or saving behaviour by firms or households.

**51** Alternatively, building on the dual economy tradition in heterodox growth theory, the rate of growth of the labour force can be made endogenous to the actual rate of growth. See, for example, Cornwall (1977) for a discussion. See Chapter 8 and McCombie et al. (2003) for further discussion of the Verdoorn law.

**52** This ignores, for simplicity, the original focus of the Verdoorn law on manufacturing industries. See Chapter 8 for further discussion.

**53** An alternative to the Verdoorn law approach used to derive the endogenous natural rate in (1.53) can be found in Palley (1996b, 1997, 2002b) who, drawing on Kaldor (1957) and Scott (1989), specifies a technical progress function in which the rate of technical progress depends on the capital–labour ratio and the level of investment spending. Since investment is a variable that determines the level of autonomous demand in the Keynesian tradition, and is therefore an element of $Z_H$ in this tradition, the Palley approach can be captured by a technical progress function of the form:

$$q = q(Z_H, Z_q) \tag{1.52'}$$

where $Z_q$ is a vector of other variables unrelated to the growth and/or level of autonomous demand. Combination of (1.24), (1.26), (1.49) and (1.52′) now yields:

$$y_N = q(\overline{Z}_H, Z_q) + \overline{n} \tag{1.53'}$$

The natural rate is once again endogenous, but this time because the demand-side determinants of the actual rate of growth also enter into the determination of productivity growth and hence the natural rate. Note, however, that as shown by Dixon and Thirlwall (1975, p. 209), the Verdoorn law can be derived from Kaldor's (1957) technical progress function, so Palley's approach to modelling the endogenous natural rate can be encompassed by the Verdoorn law approach adopted in the text.

**54** Note that the specific references in the text are to the neo-Kaldorian tradition in post-Keynesian growth theory. See, for example, Dutt (2006a, 2010b) for a parallel concern with reconciling the rates of growth of aggregate demand and aggregate supply within the neo-Kaleckian tradition, in which the supply constraint on growth emanates (in the first instance) from the availability of capital and the target or 'normal' rate of capacity utilization chosen by firms, rather than the availability of labour. However, reconciliation of aggregate demand and aggregate supply in this tradition may still give rise to the first Harrod problem as in the canonical HGT model, and thus leave open the question of reconciling the actual and natural rates of growth that is explored above.

**55** The astute reader will notice the conformity between Kalecki's description of trend-cycle interaction and the post-Keynesian conception of a path-dependent long run constituted by a sequence of short- and medium-term adjustments, described at the start of this chapter.

**56** According to equation (1.58), if $Y = Y_N$ initially, then with $\varepsilon < 0$ we will observe $y < y_N$, which will mean $Y < Y_N$ subsequently as a result of the economy having been 'knocked off' its full-employment expansion path. To see this, note that if $Y = Y_N$ initially and $y < y_N$, then $\Delta Y = Y + (1 + y)Y < Y_N + (1 + y_N)Y_N = \Delta Y_N$, with the result that $Y < Y_N$ subsequently. Assuming that $\varepsilon = 0$, $Y < Y_N$ will then mean that $y > y_N$ as the economy grows faster than the natural rate in the course of moving back towards its full-employment expansion path. Once the latter has been regained ($Y = Y_N$), and assuming that there are no further shocks ($\varepsilon = 0$), then according to (1.58) we will once again have $y = y_N$.

**57** A simple example of the mechanisms operative in these different strands of the NEGT literature can be developed by appealing to the Romer (1990) model of technical change, in which (1.36) takes the explicit form:

$$q = q(Z_N) = \kappa L_A$$

where $L_A$ is research effort and $\kappa$ is research productivity – specifically, the rate at which it is possible to transform the existing stock of knowledge into new ideas, per unit of research effort (see Jones, 2002, pp. 101–6). Hence in 'creative destruction' models, $y < y_N$ will eliminate inefficient practices and stimulate innovation, thus increasing $\kappa$ and $L_A$ (respectively), and hence $q$. In 'learning by doing' models, meanwhile, $y > y_N$ will stimulate learning by doing which will have spillover effects on the efficiency of research activity, thus increasing $\kappa$ and hence $q$.

**58** As previously noted, starting from the equilibrium position $Y = Y_N$ and with $y < y_N$, $\Delta Y = Y + (1 + y)Y < Y_N + (1 + y_N)Y_N = \Delta Y_N$, the result being that $Y < Y_N$ subsequently.

**59** These sorts of results are analogous to the effects observed in NAIRU or natural rate of unemployment models of the labour market in which hysteresis effects are postulated (for example, Cross, 1995).

**60** See Setterfield (2014, pp. 379–80) for further development of this theme and, in particular, various explanations for this state of affairs.

# Part I

# Core models of growth and distribution

# 2

# Classical-Marxian models

## 2.1 Introduction

Modern heterodox models of long-term growth and income distribution build on a foundation that originated with the classical economists and continued in the Marxian tradition. Both the classical economists (especially Adam Smith and David Ricardo) and Karl Marx strongly emphasized the long-run growth and development of the capitalist economy in their theories.[1] Smith (1776 [1976]), who titled his magnum opus *An Inquiry into the Nature and Causes of the Wealth of Nations*, emphasized the forces that made some nations richer than others. Ricardo (1821 [1951], p. 5) stated that 'To determine the laws that regulate this distribution of income [between rents, profits and wages], is the principal problem in Political Economy'; his analysis of income distribution in turn was the key to his model of long-run growth leading to a 'stationary state'. Later, Marx (1867 [1976], p. 92) sought to uncover what he called 'the economic law of motion of modern society', which revolved around the 'class struggle' between labour and capital and its impact on economic growth and technological change.

This chapter presents a set of simplified models that can be used to represent the core ideas of Smith, Ricardo and Marx on growth and distribution.[2] Our purpose here is to develop a framework for representing the basic logical structure of their theories, not to provide a textual exegesis of the original versions. Given our macroeconomic focus, we will not enter the long-running debates over the labour theory of value or other microeconomic aspects of their paradigms, although we will refer to labour value concepts where relevant to the discussion.[3] For expositional convenience, we will often depict the classical-Marxian theories using models of long-run, steady-state equilibrium positions, as defined in Chapter 1. This procedure is at best a useful pedagogical device, however; we acknowledge that it could be argued to distort the dynamic visions of the classical authors – especially Marx, who emphasized the cyclical and long-term instability of capitalism. However, we agree with Cesaratto (2015, p. 179) when he writes: 'I also fully acknowledge the limitations of the investigation of (formally stable) normal accumulation paths in view of the instability of capitalism . . . Stylised models are, however, essential to fix our ideas and for policy purposes . . ..'

One important feature in the work of the classical economists and Marx is the idea of a close connection between the accumulation of capital, which is the main engine of systemic growth, and the distribution of income between the principal social classes. Smith and Ricardo assumed three social classes – landlords, capitalists and workers – corresponding to the class structure of British society in the late eighteenth and early nineteenth centuries. Following the industrial revolution, Marx and later growth theorists abandoned the emphasis on the landlord class, and even for the original classicals (Smith and Ricardo) many of their core ideas can be represented in a framework that only explicitly models the wages of labour and profits of capital. The important point, however, is that the distinctive behaviour of different social classes (especially in regard to their respective roles in production and saving) is a crucial aspect of the classical-Marxian framework. But before turning to the growth-and-distribution theories per se, we must first lay the foundations in regard to the classical accounting scheme, which is based on the classical approach to production discussed in Chapter 1.

## 2.2 Basic accounting framework and distributional relationships

The analysis in this chapter assumes a closed economy that produces a single good, which can be used for either consumption or investment (that is, accumulated as capital and used in future production). This is admittedly a heroic simplification, which has been subject to much criticism, especially during the Cambridge capital controversies of the 1960s and 1970s,[4] but it will be adopted here as a convenient foundation for macro-level analysis. This accounting framework will easily be seen as a highly simplified version of standard national income accounting.

Ignoring landlords and rents and assuming no government or taxation for simplicity, all of national income must be divided between profits of capital and wages of labour:

$$PY = WL + rPK \tag{2.1}$$

where $Y$ is output or income, $P$ is the aggregate price level, $W$ is the nominal wage rate (money units per worker or worker-hour), $L$ is the amount of labour employed (measured either in number of workers or worker-hours), $r$ is the profit rate and $K$ is the real stock of capital.[5] The capital stock is valued by the same price index as is used for output ($P$) on the assumption of a single good. We are also simplifying by assuming no depreciation of capital, so that there is no distinction between gross and net measures of output and profits. Since this is a closed economy, there is also no distinction between domestic output and national income.

As explained in Chapter 1, we will assume that output is produced using a fixed-coefficients or 'Leontief' production function,

$$Y = \min\left(\frac{L}{a_0}, \frac{K}{a_1}\right) \tag{2.2}$$

where $a_0 = L/Y$ is the labour–output coefficient and $a_1 = K/Y_K$ is the full-capacity capital–output coefficient. Employment is of course constrained by the available labour force ($L \leq N$), but full employment is not necessarily assumed (only in some versions of the classical-Marxian models) and labour supply is sometimes treated as endogenous, as will be explained below.

As noted in Chapter 1, the classical theory of production generally assumes that the capital constraint is binding, so that output is proportional to the capital stock, $Y = Y_K = K/a_1$, that is, output is at the potential level determined by the society's available capital stock. The level of employment, meanwhile, is determined by the amount of capital accumulated so that $L = a_0 Y = (a_0/a_1)K$. In other words, capital accumulation determines the level of output, which in turn (given the technical input–output coefficients) determines the level of employment. This approach assumes that output is produced in a constant proportion to the maximum output that can be produced with the society's capital stock, at some normal rate of capacity utilization ($u_n$), in a long-run (or 'long-period') steady-state equilibrium. We shall refer to this as the assumption of 'full utilization', although it really only requires a constant or normal rate of capacity utilization, but for purposes of simplification we will assume that $u_n = 1$ in the classical-Marxian models.[6] Under this simplifying assumption, we can say that $Y = Y_K$ and, in effect, we can write the capital coefficient as $a_1 = K/Y$.

Dividing both sides of equation (2.1) by $PY$ and rearranging, we obtain

$$1 = wa_0 + ra_1 \tag{2.3}$$

where $w = W/P$ is the real wage (measured in number of goods per worker or per hour). This equation shows how the *shares* of profits ($\pi = ra_1$) and wages ($1 - \pi = wa_0$) sum up to 100 per cent of total national income. Rearranging (2.3), we obtain the famous inverse relationship between wages and profits

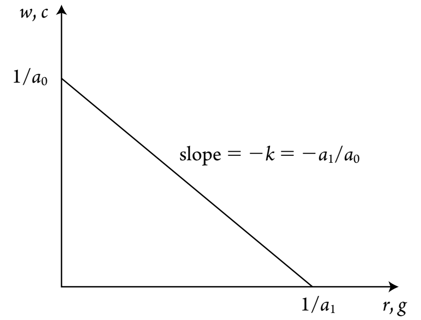



**Figure 2.1** The inverse wage–profit and consumption–growth relations

first clearly posited by Ricardo,[7] and later incorporated in subsequent models of growth and distribution in the classical, Marxian, neo-Keynesian and neoclassical traditions.[8] Given our simple specification of technology with fixed coefficients and a single good, this inverse relationship is linear:

$$w = \frac{1}{a_0} - \frac{a_1}{a_0} r \tag{2.4}$$

To have an economically meaningful solution with positive wages, we have to assume $r < 1/a_1$. Similarly, we must also have $w < 1/a_0$ in order for there to be positive profits. These, however, are only logical or outer limits to the possible configurations of wages and profits; more realistically, the real wage could not go, for example, below a level necessary to sustain and reproduce the labour force in the long run, and profits may not be reduced below a minimal rate deemed necessary for capital owners (capitalists) to be able to save and invest, that is, accumulate more capital. Some of these narrower, social limits on income distribution are addressed later in this chapter.

Equation (2.4) is graphed in Figure 2.1, where it can be seen that the vertical intercept (maximum possible real wage) is the output per worker or the productivity of labour ($Y/L = Q = 1/a_0$) while the horizontal intercept (maximum rate of profit) is the output–capital ratio or 'productivity of capital' ($Y/K = 1/a_1$). Also, note that the slope of this relationship is $-a_1/a_0$, the absolute value of which is the 'capital intensity' of production, that is, capital per worker or the capital–labour ratio ($k = K/L = a_1/a_0$). Thus, a steeper trade-off between wages and profits indicates a higher degree of capital intensity in production. Alternatively, either (2.3) or (2.4) can be solved for the profit rate:

$$r = \frac{1 - wa_0}{a_1} = \frac{\pi}{a_1} \tag{2.5}$$

which highlights that the profit rate is directly related to the profit share (written as $\pi = 1 - wa_0$) and inversely related to the capital coefficient $a_1$.

In addition, this system of production relations implies that there is an inverse relation or trade-off between consumption and growth, which is an exact 'dual' to the inverse wage–profit relation. Starting with the goods market equilibrium condition that aggregate demand (that is, the sum of consumption and investment) must equal output in an economy with no foreign trade or government:

$$PY = PC + PI \tag{2.6}$$

where $C$ is real consumption and $I = \Delta K$ is real investment in new capital; as noted earlier, we assume no depreciation of capital for simplicity, in which case there is no replacement investment. It is easily seen that (2.6) is equivalent to

$$c = \frac{1}{a_0} - \frac{a_1}{a_0} g \tag{2.7}$$

where $c = C/L$ is total consumption (of both workers and capitalists) per employed worker and $g = \Delta K/K = I/K$ is the rate of capital accumulation or 'growth rate'. Equation (2.7) has the exact same form as (2.4), with $w$ and $r$ replaced by $c$ and $g$ (respectively); hence, the graph is identical to Figure 2.1 except that $c$ replaces $w$ on the vertical axis and $g$ replaces $r$ on the horizontal axis.

Before leaving this accounting section, it is important to note that the existence of strict inverse relations or trade-offs between wages and profits on the one hand, and consumption and growth on the other hand, rests upon *two very strong assumptions*: a given technology and a constant or normal rate of capacity utilization. As we will see below, technological progress can shift these relationships outward (and possibly also alter the slopes, that is, the capital intensity), generally improving the trade-offs and potentially allowing one of each pair of variables ($r$ and $w$, $g$ and $c$) to increase without the other one decreasing (or possibly both to increase). An important exception, however, is the type of technological change emphasized by Marx, in which labour productivity increases but capital productivity falls, which as we will see can result in conflictive situations in which one class benefits at the expense of the other. Also, if there is excess capacity (output is below its potential level, or utilization is below its normal rate), then an economy will be operating *inside* the wage–profit and consumption–growth frontiers described by equations (2.4) and (2.7). In this situation, it is possible for both variables in each pair ($w$ and $r$, or $c$ and $g$) to increase simultaneously (or for one to rise without the other falling) if the economy increases its rate of utilization of capacity and moves closer to the frontier. The consequences of underutilization of capacity will be covered in Chapter 4, but in this chapter we confine ourselves to models of economies that function at a normal utilization rate as defined earlier.

## 2.3 Saving behaviour and distributional closures

The classical economists and Marx made a simple but powerful assumption about saving, namely that all saving was done by the capitalist class. They generally assumed that workers earned only enough wages to barely pay for their 'necessary' consumption, while the landed aristocracy spent all of its rental income on luxury consumption. Of course, the classicals recognized that this assumption could be too simple, especially because landlords might invest in improvements to the land and the erection of buildings (although, in Ricardo's theory, such improvements and buildings would count as capital that earned profits, regardless of whether they were owned by the same people who owned the land or by the capitalist farmers who rented the land from them). In any case, the classicals saw the capitalist class as consisting of those individuals who were frugal and parsimonious enough to accumulate capital and to invest it in purchasing raw materials and machinery and hiring productive labour to produce commodities for sale in order to make profits. Marx added the view that the capitalists were able to 'exploit' labour by virtue of their monopoly ownership of the 'means of production'. In this chapter we will assume that all savings come out of the profits received by capitalists; cases in which workers also save, or in which profits are divided into the interest income of 'rentiers' (bondholders or equity owners) and the net profits of firms or enterprises, are covered in later chapters.

An important feature of the classical-Marxian approach is that none of the classical economists or Marx clearly distinguished what we today call saving (devoting a portion of current income to the accumulation of financial assets instead of spending it on current consumption) and investment (the purchase of newly produced capital goods, such as machinery, equipment or structures). The classicals tended to use one word, 'accumulation', to mean both saving and investment. Among the classicals, only Malthus (1820 [1951]) struggled to break free from this conflation of saving and investment in his discussion of a 'general glut' of commodities, where he argued that an excessive 'passion for accumulation' (in the sense of saving) would imply a diminution of consumer demand. But because he could not conceptualize saving as distinct from investment, he failed to explain convincingly how too much 'accumulation' could lead to a shortfall in aggregate demand. Ricardo, following Smith (1776 [1976]) and Say (1803 [1971]), rebutted Malthus by arguing that accumulation (in the sense of investment) also contributed to the total demand for a society's output, and his argument won the day until Keynes (1936) sorted this out more than a century later (as we will discuss in the next chapter, in the course of introducing the neo-Keynesian approach to growth theory).

Marx (1867 [1976], 1894 [1981]) argued that excessive saving could lead to shortfalls of aggregate demand through his use of concepts such as 'hoarding', 'ruptures in the circuit of capital' and 'realization crises', and he ultimately recognized the emergence of 'joint-stock companies' (that is, publicly traded corporations) as well as financial intermediation between firms. Nevertheless, Marx also did not clearly distinguish saving from investment, and therefore he never developed a coherent model of how aggregate demand could limit or determine the level of output; nor did he fully incorporate his understanding of the evolution of industrial concentration and corporate finance into his core analytical framework. The best way to understand the classical-Marxian model of accumulation, therefore, is that it envisions a set of entrepreneurs who operate proprietary firms (with no public stock ownership and no separation of ownership from management), in a system with underdeveloped financial markets, in which the only way the capitalists (owners of the firms) can save a portion of their net earnings is by ploughing them back into their firms in the form of increased capital stock.

Given that our aim in this chapter is to represent the classicals and Marx, we will not have separate functions for saving and investment, as we will in later chapters. Instead, we will assume an 'accumulation function', in which for simplicity we postulate that capitalists accumulate (save *and* invest) a constant proportion $s_r > 0$ of their profits over and above a minimal profit rate that generates a floor level of consumption:[9]

$$g \equiv I/K \equiv S/K = s_r(r - r_{min}) \text{ for } r > r_{min};\ g = 0 \text{ otherwise} \qquad (2.8)$$

Here, $g$ is the rate of capital accumulation or growth rate of the capital stock, as defined previously, which is equal by definition to both the investment–capital and saving–capital ratios; $r_{min} \geq 0$ is the minimum profit rate at or below which capitalists spend all their profits on consumption, and above which they save (*and* invest) the fraction $0 < s_r < 1$ of any additional profits. We will refer to $s_r$ as the capitalists' 'propensity to save', not only to anticipate the Keynesian terminology used in other chapters, but also to emphasize that this is a behavioural propensity on the part of one social class and not the realized saving rate of the society (actual saving divided by capital), which in the models presented in this chapter equals $g$.

Combining equations (2.4), (2.7) and (2.8), we have a system of three equations in four endogenous variables or unknowns: the real wage $w$, profit rate $r$, growth (capital accumulation) rate $g$ and consumption per worker $c$. Although the system of equations is thus indeterminate or 'open', we can 'close' the system and make it determinate by introducing one additional, independent relationship. This means that a variety of alternative models can be represented by considering alternative assumptions for the fourth and final equation, a methodology that has become known as 'alternative closures'.[10] Furthermore, it is clear from the foregoing that the 'missing link' in the classical-Marxian system is an equation to pin down the distribution of income: a relationship that can determine either $w$ or $r$ (or some relationship between them) will suffice to make the system determinate.

This brings us to a considerable difficulty: neither the classical economists nor Marx had a single, unitary theory of income distribution. Each of them presented a number of different notions about what determined income distribution, which they expressed in different parts of their writings.[11] Overall, we can identify four distinct distributional hypotheses that can be found in the work of Smith, Ricardo, Marx and their followers and interpreters:

1) *An exogenously given real wage*, representing an ordinary standard of living for a working-class family. Ricardo (1821 [1951], pp. 96–7) called this the 'natural wage', and initially accepted that it was equivalent to the subsistence minimum of Malthus (1798 [1993]), but then stated that it could vary across countries or over time as it depended on the 'habits and customs of the people'. Marx (1867 [1976], p. 274) similarly referred to 'the means of subsistence necessary for the maintenance of' the worker, and then defined the 'value of labour-power' as the value in labour time of those means of subsistence, while making it clear that the necessary level of consumption was socially and historically determined and not biologically fixed. While social and cultural definitions of a worker's necessary standard of living could change gradually over time, it can be regarded as exogenously given in any historical period in a particular country, and hence can be represented as

$$w = \overline{w} \tag{2.9}$$

2) *An exogenously given wage share of national income*, representing a given balance of bargaining power between workers and firms in wage negotiations. The idea of wages as determined in a bargaining process dates back to Smith (1776 [1976]), who noted that the 'masters' usually had a stronger bargaining position than the 'workmen' due to the greater

financial resources of the former. The notion of wages as determined by the relative bargaining power of workers and employers was further developed by Marx in his concept of the 'class struggle'. Perhaps the place where this assumption was most explicitly adopted was in Marx's theory of the 'falling tendency of the rate of profit' (FTRP), in which he assumed a constant 'rate of surplus value' (also known as the 'rate of exploitation'). Since this rate was essentially the ratio of profits to wages (both measured in labour-value terms), constancy of this rate implies constant shares of national income (value added) for wages and profits. Later, we will see that this is indeed the only distributional hypothesis under which Marx's FTRP theory makes logical sense, but for now we simply represent it by the equation

$$w = (1 - \bar{\pi})/a_0 = (1 - \bar{\pi})Q \tag{2.10}$$

which says that the real wage must be a fixed share $(1 - \bar{\pi})$ of labour productivity ($Q = 1/a_0$). It should be noted that, for (2.10) to hold, the real wage must always increase at the same rate as labour productivity, $\hat{w} = q$, where $q = \hat{Q} = -\hat{a}_0$ is the rate of productivity growth, in the long run.[12]

3) *Full employment* or *a constant unemployment rate*. Smith, Ricardo and Marx all postulated that the real wage would vary over time depending on the balance between the growth of labour demand, which would tend to increase the wage, and the growth of labour supply, which would tend to decrease it.[13] Smith, for example, noted that rapid growth of labour demand was the one factor that could strengthen the otherwise inherently weak bargaining position of the 'workmen' by inducing employers to bid for their services. Ricardo made this the centrepiece of his analysis of the market wage, which he said could differ from the natural wage (hypothesis (1) above), and argued that the market wage could stay above the natural wage 'for an indefinite period' in a rapidly progressing society in which the demand for labour was continuously increasing (Ricardo, 1821 [1951], pp. 94–5). Marx stated very explicitly that 'accumulation is the independent, not the dependent variable; the rate of wages the dependent, not the independent variable' in his 'general law of capitalist accumulation' (Marx, 1867 [1976], p. 770).

Given a constant capital–labour ratio ($k = a_1/a_0$), an assumption we will relax later on, the growth rate of labour demand (employment) must equal the rate of accumulation of capital, or $l = \hat{L} = \hat{K} = g$, so this assumption can be represented by the differential equation

$$\hat{w} = \varphi(l - n) = \varphi(g - n), \varphi' > 0, \varphi(0) = 0 \tag{2.11}$$

where $n = \hat{N}$ represents the growth rate of the labour supply. Under this specification, a steady-state equilibrium for income distribution requires a constant real wage, $\hat{w} = 0$, which (assuming $q = \hat{a}_0 = 0$ or no labour-saving technical change) in turn implies[14]

$$g = n \tag{2.12}$$

Equation (2.12) represents a version of the *natural rate of growth* of Harrod (1939), at which there is a constant equilibrium unemployment rate (which may or may not coincide with 'full employment'), as discussed in Chapter 1. The differences here are that growth is represented by the rate of capital accumulation ($g$) instead of the growth rate of output ($y$), and that the growth of labour productivity is ignored for simplicity (so that $q = 0$).[15] Below, we will see that combining (2.12) with equations (2.4) and (2.8), and including any labour supply function in which $n$ depends on $w$, will imply a definite distribution of income along the $w - r$ trade-off line shown in Figure 2.1. Also note that since the wage adjusts gradually over time according to equation (2.11), it can be taken as given at any instant of time; hence, the model of an exogenously fixed real wage (equation 2.9) applies in the short run while the natural rate of growth (equation 2.12) is reached only in the long run.

4) *A given rate of profit, determined by financial market forces*. This possibility was proposed by Sraffa (1960, p. 33) when he wrote, 'The rate of profits, as a ratio, has a significance which is independent of any prices, and can well be "given" before the prices are fixed. It is accordingly susceptible of being determined from outside the system of production, in particular by the level of the money rates of interest.' This idea has acquired new importance in the twenty-first century, as it may be a useful modelling strategy for reflecting the increasing power of financial interests over industrial producers (see Panico et al., 2012). We can represent a simple version of this idea by postulating

$$r = i + \lambda \tag{2.13}$$

where $i$ is the interest rate on loans to firms (set, for example, by a markup or spread over the central bank's interest rate on overnight loans of reserves) and $\lambda$ is a risk premium. This distributional hypothesis has been less well developed in growth models, however, and caution must be taken in using it in analysing the comparative statics of the model for reasons that are discussed below.

In the next section, we will study the behaviour of our classical-Marxian growth model under these four alternative closures – each of them reflecting one of the distributional hypotheses just discussed – and discuss the conditions under which each of them might be observed and how they may be related to each other.

# 2.4 Model solutions under alternative closures

## 2.4.1 An exogenously given real wage

The first hypothesis, of a given real wage, could correspond to a 'dual economy' with a labour surplus in the sense of Lewis (1954) and Ranis and Fei (1961): an economy in which labourers migrate from a backward (traditional or subsistence) sector to a modern, capitalist sector, with an infinite elasticity of labour supply to the latter sector (up to some turning point) at a wage equal to the average product of labour in the backward sector plus a premium for migrating. Alternatively, this hypothesis can represent the short run of the dynamic model in distributional closure (3), in which the real wage is a state variable that is given in the short run and evolves over time according to equation (2.11) to reach the long-run equilibrium described by (2.12). We shall return to closure (3) below, but first we cover the (relatively simple) case of an exogenously fixed real wage in closure (1).

In this case, the four equations of the model are (2.4), (2.7), (2.8) and (2.9), and they can be solved recursively in the logical order of causality in the model. First, of course, (2.9) sets the equilibrium real wage as $w^* = \overline{w}$ (where '*' represents an equilibrium value or solution). Then, using (2.5), we obtain the equilibrium profit rate

$$r^* = \frac{1 - \overline{w}a_0}{a_1} \tag{2.14}$$

Next, the saving function (2.8) determines the equilibrium accumulation rate (growth rate of capital):

$$g^* = s_r(r^* - r_{min}) \tag{2.15}$$

where we assume that $r^* > r_{min}$ so that there is positive growth in equilibrium. Finally, the consumption–growth trade-off (2.7) determines equilibrium consumption per worker:

$$c^* = \frac{1}{a_0} - \frac{a_1}{a_0} g^* \tag{2.16}$$

Note that, in this scheme, both profits and consumption are essentially residuals: profits are whatever is left over after wages are paid at the fixed rate $\overline{w}$, and consumption is what's left over after the capitalists effectuate the accumulation (saving and investment) made possible by those profits. However, because the workers' consumption level is fixed at the same level as the real wage (recall that workers do not save), what really varies in determining total consumption is the level of capitalists' consumption. To see this, note that we can decompose total consumption (per employed worker) into workers' and capitalists' consumption (or rather, consumption out of wages and profits) as follows:

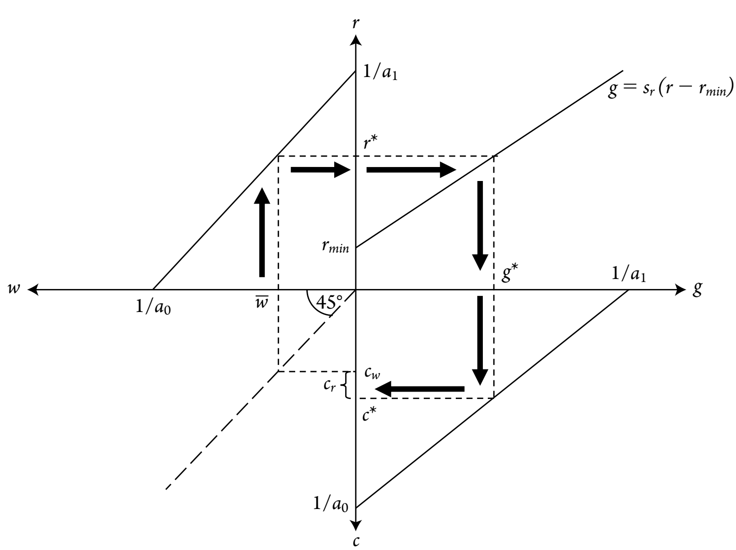
$$c = c_w + c_r = \overline{w} + c_r \tag{2.17}$$

which implies that the residual variations in equilibrium $c^*$ are actually (and entirely) variations in $c_r$.

This recursive model solution and its causal logic are displayed in the four-quadrant diagram in Figure 2.2, where the inverse $w - r$ and $c - g$ relations are shown separately in the second and fourth quadrants, respectively (with the axes rotated accordingly). Note also that the saving (accumulation) function (2.8) is drawn in this diagram, and all subsequent ones, with the dependent variable ($g$) on the horizontal axis and the independent variable ($r$) on the vertical axis (similar to a standard Marshallian supply-and-demand diagram, in which the quantity is a function of the price). The 45-degree line

Note: Bold arrows indicate direction of causality; in the fixed wage share case, $\overline{w}$ is replaced by $(1 - \overline{\pi})/a_0$.

**Figure 2.2** Model solution for a fixed real wage or wage share



in the third quadrant shows the correspondence between $w$ and $c_w$, while $c_r$ is the gap between $c$ and $c_w$. The bold arrows show the direction of causality.

### 2.4.2 An exogenously given wage share

This case could arise if there are powerful labour unions and supportive government policies that enable workers to claim a constant share of national income – which, as noted earlier, means that real wages increase proportionately to the productivity of labour. This was the case, for example, in the US economy during the so-called golden age of capitalism during the first two decades after World War II (roughly, the late 1940s to the late 1960s – see Marglin and Schor, 1990; Pollin, 2005). Such an outcome could also result from centralized labour-management bargaining in which wages are set to rise in nominal terms at the rate of labour productivity growth plus the expected rate of price inflation, as suggested by Hein and Stockhammer (2011a). However it comes about, an exogenously given wage share means that the real wage is determined by equation (2.10), and then the other variables in the system are determined recursively by equations (2.14) to (2.16). The graph is the same as Figure 2.2, except that in this case $w^* = (1 - \bar{\pi})/a_0$ per equation (2.10); otherwise, the causality flows in the same direction shown in Figure 2.2.

### 2.4.3 Natural rate of growth, or a constant unemployment rate

This closure arises, as discussed above, as the long-run, steady-state solution of a model in which the real wage adjusts to equilibrate the growth of labour demand and labour supply, and labour demand grows at the same rate at which capital accumulates ($g$). In the classical scheme, the growth rate of labour supply ($n$) was an endogenous variable that was determined by the demographic process of population growth. According to the classical economists, labour supply growth was an increasing function of the real wage relative to the 'natural' subsistence minimum $w_s$, as in the function $n = n(w - w_s)$, where $n' > 0$ and $n(0) = 0$. However, only Malthus (1798 [1993]) insisted that $w_s$ was a biological minimum level of consumption, below which workers would suffer malnutrition, reduced fertility and increased mortality.[16] Both Smith and Ricardo insisted that there was a historical and cultural element embedded in what they called the natural wage, and Ricardo in particular argued that the 'market wage' determined by equations (2.11) and (2.12) could remain indefinitely above $w_s$ in a rapidly growing economy (see section 2.7, below, for further discussion of this point).

Marx argued that, contrary to the Malthusian view, labour supply was socially determined within the capitalist economy rather than determined by a demographic or biological mechanism. According to Marx, labour supply could be augmented by bringing new sources of labour into the market system (for example, peasants, women or children), and later Marxists broadened this concept to include colonial labour, immigrants, guest workers, offshoring and so on.[17] Marx also believed that labour supply could be effectively augmented by labour-saving technological change, although that includes the possibility of varying the capital–labour ratio which we won't consider at this point. All of these mechanisms were, in Marx's view, endogenously induced by increases in the real wage insofar as these would cut into profitability (given the existence of an inverse $w - r$ relation), and thus his theory of labour supply can be represented by a similar function but omitting the subsistence minimum: $n = n(w), n' > 0$.

Regardless of whether we adopt the classical or Marxian view of labour supply, the analytics of the solution for this closure will be similar, although the interpretation of how labour supply adjusts will be substantively different (in terms of a natural versus a social mechanism). Here, we restrict ourselves to the long-run, steady-state equilibrium solution using (2.12) along with (2.4), (2.7) and (2.8); later, when we consider changes in the various model parameters, we will analyse the short-run equilibrium of this system and the dynamics of adjustment per (2.11). For the growth rate of labour supply, we will adopt a variant of the linearized neo-Marxian function used by Harris (1983):

$$n = n_0 + n_1 w, \qquad n_0, n_1 > 0 \tag{2.18}$$

Substituting equation (2.4) into (2.18) and then substituting the result together with equation (2.8) into (2.12), we obtain a single equation in one endogenous variable ($r$):

$$s_r(r - r_{min}) = n_0 + n_1\left(\frac{1 - a_1 r}{a_0}\right) \tag{2.19}$$

which solves for the long-run equilibrium profit rate[18]

$$r_{LR} = \frac{a_0(n_0 + s_r r_{min}) + n_1}{a_0 s_r + n_1 a_1} \tag{2.20}$$

We then obtain the simultaneous solution for the growth rates of capital and labour, which have to be equal, by substituting (2.20) back into either (2.8) or else (2.18) in combination with (2.4). Either way, the solution for the long-run equilibrium rate of growth is

Note: Bold arrows indicate the direction of causality in the long run; lighter arrowheads show the dynamics of real wage adjustment towards the long-run equilibrium.

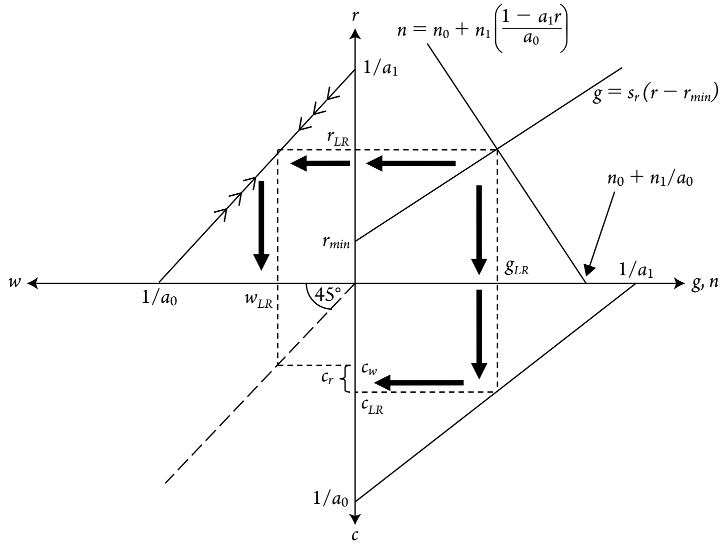



**Figure 2.3** Model solution assuming full employment or a constant unemployment rate

$$g_{LR} = n_{LR} = s_r\left(\frac{a_0 n_0 + n_1(1 - a_1 r_{min})}{s_r a_0 + n_1 a_1}\right) \tag{2.21}$$

where $1 - a_1 r_{min} > 0$ must hold for there to be positive wages at the minimum profit rate at which capitalists start to save and invest. Finally, substituting (2.20) into (2.4) yields the long-run equilibrium real wage ($w_{LR}$), while substituting (2.21) into (2.7) yields long-run equilibrium consumption per worker ($c_{LR}$). This solution is graphed in Figure 2.3, where again the bold arrows indicate the direction of causality. Note that since the labour supply function is increasing in the real wage $w$, it is decreasing in terms of the profit rate $r$. Also note that the causality radiates outward from the simultaneous solution for $r$, $g$ and $n$ in the first quadrant to determine the real wage and consumption per worker in the second and fourth quadrants.

The causality illustrated in Figure 2.3 pertains to the long-run, steady-state equilibrium configuration in which (2.12) holds. If the economy is out of long-run equilibrium, it would have a given real wage as in equation (2.9) in the short run, and the real wage would adjust towards its new, long-run equilibrium level according to the dynamics described by equation (2.11). This adjustment is shown by the light arrowheads along the inverse $w - r$ relation in Figure 2.3.

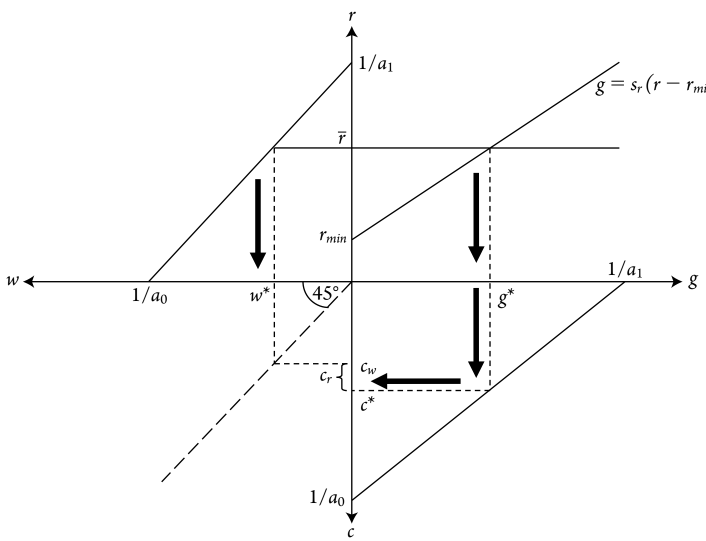
An important special case of this model closure is the one in which labour supply grows at an exogenously fixed rate $n_0$ (so $n_1 = 0$) and there is no minimum profit rate for positive saving (so $r_{min} = 0$). In this case, the solution (2.20) reduces to $r_{LR} = n_0/s_r$, which is the famous 'Cambridge equation' linking the profit rate to the ratio of the growth rate (here, $n_0$) to the propensity to save out of profits ($s_r$). Also in this special case, the long-run equilibrium growth rate (2.21) reduces to the exogenously given 'natural' rate $g_{LR} = n_0$, and the downward-sloping $n$ function in Figure 2.3 would be replaced by the vertical line $n = n_0$. However, it is more faithful to the classicals and Marx if the growth rate of the labour supply is allowed to be endogenous.

### 2.4.4 An exogenously given rate of profit

The last case we will consider is a profit rate that is determined by financial market forces independently of the real growth of the economy. In this case, with the profit rate determined by equation (2.13), the rest of the model is easily solved by substituting this profit rate into the inverse wage–profit relation (2.4) to get the real wage and into the accumulation function (2.8) to get the growth rate; finally, the latter is substituted into the inverse consumption–growth relation (2.7) to solve for consumption per worker. In this case, both income distribution and growth are driven by the dictates of the rates of return that firms find necessary to satisfy their financial investors (lenders); workers are powerless to influence their real wages unless they can win reforms of the financial sector that would lower the necessary profit rate. The logic of the solution is illustrated in Figure 2.4.

Note: Bold arrows indicate direction of causality.

**Figure 2.4** Model solution assuming an exogenously given profit rate



## 2.5 Effects of exogenous changes in saving propensities or income distribution

To better understand the logic and implications of each of the four distributional closures in a classical-Marxian model, we can study the effects of changes in the exogenous variables in each closure on the model equilibrium. Each variant (closure) of the model contains three types of exogenous parameters – parameters governing income distribution (whatever determines $r$ or $w$), saving-cum-investment (the propensity to accumulate out of profits, $s_r$) and technology (the input–output coefficients, $a_0$ and $a_1$). We will now consider the first two types of parameter shifts, for each of the alternative closures; technological changes will be covered in the following section. Of course, in reality more than one of these parameters may change at the same time, but for analytical purposes we treat each type of parameter shift separately, holding the other parameters constant.

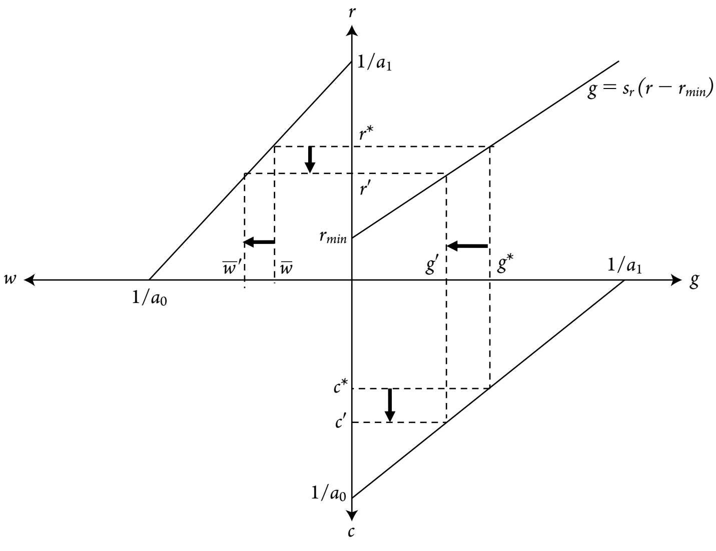
### 2.5.1 Redistribution of income

Recall that profits are the sole source of the saving and investment that lead to the accumulation of capital in the classical view, while profits and wages are inversely related. Hence, holding other factors constant, a shift in the distribution of income from wages to profits – essentially, a regressive redistribution that favours the wealthier class – is seen as necessary to promote more rapid growth in the classical vision. Or, to put it another way, the classical vision implies a potential trade-off or conflict between distributive equity and economic growth. To be sure, there were differences among the various classical authors and Marx in how they interpreted and qualified the inevitability of this trade-off, and we will elucidate some of those differences below, but first we will start by outlining the basic analytics of this view of a conflictive growth process. Each closure contains one exogenous variable that drives income distribution, and in this section we will study the effects of changes in that variable on the growth rate under classical-Marxian assumptions.

The cases of a fixed real wage and fixed wage share can be considered together, because they are similar in their causal logic and the same graph can be used to represent both. In Figure 2.5, where all shifted variables (including new equilibrium values) are indicated by a prime ($'$), an exogenous increase in the real wage $\overline{w}$ or wage share $(1-\overline{\pi})$ requires a decrease in the profit rate along the wage–profit inverse relation (second quadrant). Assuming that capitalists do not increase their saving-and-investment propensity $s_r$ in response, the lower profit rate in turn would reduce the funds available for capital accumulation and hence there would be a lower growth rate in the first quadrant.

Note: Bold arrows indicate shifts across equilibria; for a fixed wage share, $\overline{w}$ is replaced by $(1 - \overline{\pi})/a_0$.

**Figure 2.5** Effects of an increase in a fixed real wage or wage share



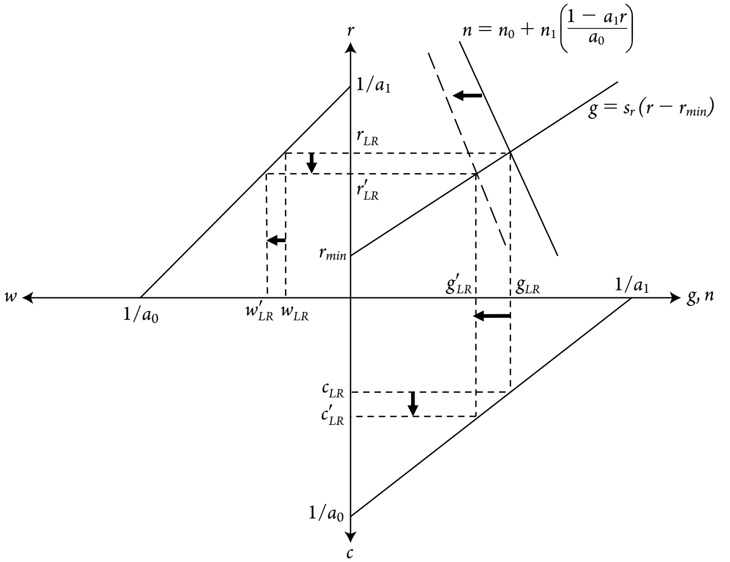

As the growth rate falls, total consumption per worker must rise according to the consumption–growth trade-off in the fourth quadrant. Although workers' and capitalists' consumption (both measured per worker) are not shown separately to avoid cluttering the graph, it is clear that workers' consumption (which equals the real wage) must rise while capitalists' consumption must fall, and it can be shown that the rise in workers' consumption must outweigh the fall in capitalists' consumption so that total consumption (also measured per worker) will end up higher.

Next, we turn to the natural rate closure, in which the distribution of income is governed by the dynamics of labour supply and demand in the long run. To focus on a shift that would produce a rise in the real wage (the case of a fall would be symmetrical), suppose that there is a reduction in the constant term $n_0$ in the labour supply function (2.18). This could occur, for example, as a result of a lower birth rate, a reduction in immigration inflows, or increased restrictions on child labour. As shown in Figure 2.6, the labour supply function would shift to the left, causing the equilibrium growth and profit rates to fall, and allowing the long-run equilibrium real wage and consumption per worker to rise.[19] Of course, this shows that the so-called natural rate of growth is not really natural or constant, as the growth rate that equilibrates labour supply and demand can vary when the social and institutional conditions underlying labour markets are altered.

Finally, if we assume an exogenously given profit rate, the real wage can only increase if this rate falls, perhaps because of a looser monetary policy

Note: Bold arrows indicate shifts across equilibria.

**Figure 2.6** Effects of an exogenous reduction in labour supply growth




that lowers long-term interest rates, a reduction in the concentration of the financial sector (which could lower banks' markups of their lending rates above the central bank rate) or a decrease in the risk premium. If the profit rate is influenced by the monopoly power of firms, then it could also fall if monopoly power is reduced, for example through stricter competition policies ('antitrust' actions of the government). Mechanically, the effects are quite straightforward, based on the solution outlined in section 2.4.4 and illustrated in Figure 2.4: the growth rate must fall and consumption per worker must rise, assuming that the capitalists' propensity to save and invest ($s_r$) remains unchanged. However, in this case it is especially problematic to assume that this propensity would remain unchanged, as changes in the conditions directly underlying capitalists' profitability might also affect their decision about how to allocate their profits between accumulation and consumption.

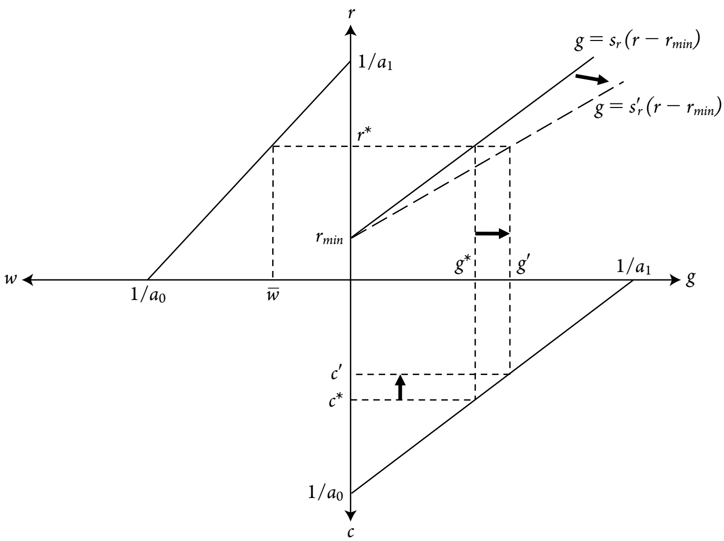
Overall, these exercises demonstrate the conflict between distributional equity and rapid growth that is implied by the classical framework, under given conditions of technology and saving behaviour, and assuming full or normal utilization of productive capacity as defined earlier. It remains to be seen how this conflict or trade-off can be either ameliorated or worsened, if those other conditions are allowed to vary.

## 2.5.2 An increased propensity to save

Now suppose that, taking the underlying distributional relationships as given, capitalists decide to increase their propensity to save – which, under classical

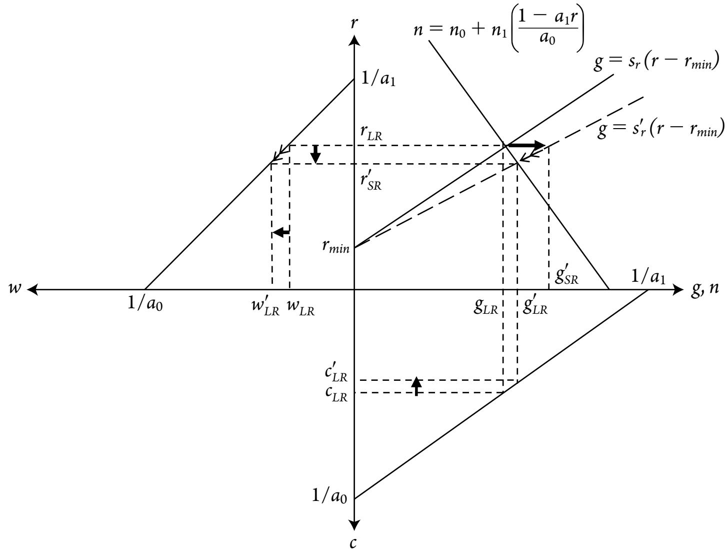
Note: Bold arrows indicate shifts across equilibria; holding the wage share or profit rate constant would be similar.

**Figure 2.7** Effects of an increase in the capitalists' propensity to save $s_r$, holding the real wage constant




assumptions, automatically implies greater investment in productive capital. If an act of saving is identically equivalent to an act of purchasing a new capital good, then no aggregate demand problem can result, because demand merely shifts from consumption to investment and its total level is not diminished. Under these strong assumptions, a rise in the capitalists' saving propensity $s_r$ almost always translates into a rise in the economy's growth rate, with one important exception that will be discussed below. However, the effects of the rise in the propensity to save on the equilibrium distribution of income vary among the different closures, as we will also see.

The three cases of a given real wage, a given wage share and a given profit rate are all qualitatively similar in regard to the impact of an increased propensity to save, and so will be treated together here. Each of these closures implies that the equilibrium profit rate is independent of the capitalists' saving (investment) propensity $s_r$, so that when the saving function rotates out to the right, the equilibrium growth rate increases in direct proportion to the rightward shift in that function (see Figure 2.7). Total consumption per worker has to be squeezed to make the increased saving and investment possible, but the entire reduction in consumption comes at the expense of capitalists' consumption (because the real wage and workers' consumption are unaffected in each of these closures) and there is no change in the distributional variables $w$ and $r$. In all of these closures, the central constraint on growth is the availability of savings out of profits to finance investment in new capital, so an increase in the propensity to save relaxes that constraint and permits more rapid growth without any change in income distribution.




Note: Light arrowheads show the adjustment from the new short-run equilibrium *after* the increase in the propensity to save to the new long-run equilibrium.

**Figure 2.8** Effects of an increase in the propensity to save, 'natural rate of growth' closure

Matters are more complicated when we consider the natural rate of growth closure. In this case, the long-run equilibrium profit rate depends on the capitalists' saving (accumulation) propensity per equation (2.20), so distribution will also be affected when propensity to save rises. Furthermore, we know that the move to a new long-run equilibrium in this closure must involve dynamic adjustments of the real wage per equation (2.11), so this adjustment also has to be modelled. In fact, what occurs is a type of 'overshooting' phenomenon, in which the long-run equilibrium growth rate rises, but in the short run the growth rate temporarily rises beyond its new long-run level, and then returns to the latter.

These dynamics can be understood with the aid of Figure 2.8, in which the accumulation line again rotates down and to the right (towards the *g*-axis). In the short run, the growth rate is determined by the new, higher propensity to save applied to the original profit rate $r_{LR}$, and therefore increases to $g'_{SR}$.[20] However, the increase in the growth rate in the short run makes labour demand increase faster than labour supply, thus causing the real wage to rise gradually according to equation (2.11). As the real wage rises, the profit rate gradually falls, until the economy reaches a new long-run equilibrium at which labour supply and labour demand are growing at the same rate ($g'_{LR}$).

If labour supply is endogenous, then the fall in profits and rise in wages induce faster growth in the labour supply (for example, via increased inflows of immigrants or guest workers), so that the new long-run equilibrium growth rate $g'_{LR}$ is greater than the *ex ante* growth rate $g_{LR}$ (even though $g'_{LR}$ is less than the temporary, short-run growth rate $g'_{SR}$), as shown in Figure 2.8. Here, for the first time we see a case in which the real wage and the growth rate can both increase simultaneously in the long run, *provided* that the growth rate of labour supply adjusts upward in response to the higher real wage as postulated in equation (2.18). In the special case of an exogenously growing labour supply ($n_1 = 0$, implying $n = n_0$), however, the long-run growth rate cannot rise (the labour supply curve in Figure 2.8 becomes a vertical line), so the growth rate returns to its initial level at $g_{LR} = n = n_0$ in the long run, but the profit rate still falls and the real wage still rises in the new long-run equilibrium (the interested reader should draw the corresponding diagram as an exercise).

It is important to understand the intuition for why the profit rate has to fall in the long run when the accumulation (saving) rate has increased. With a higher propensity to save out of profits, a higher percentage of each unit of profits is devoted to accumulation, so less profits are required to finance any given rate of investment (accumulation). If labour supply growth is exogenous and the long-run accumulation rate $g_{LR}$ does not change, then the story is relatively straightforward: less profits are required for that purpose, so the profit rate has to fall, and the rising real wage (induced by the temporarily higher growth of labour demand during the transition to the new long-run equilibrium) is the mechanism that ensures the necessary fall in profits. If labour supply growth is endogenous, however, then as the real wage rises, labour supply starts to grow more rapidly, so the long-run equilibrium growth rate increases to some extent, but the increase in the long-run equilibrium growth rate is less than proportional to the rise in the propensity to save (as shown by the fact that $g'_{LR} < g'_{SR}$), so a lower profit rate is still required to finance growth even at the somewhat higher long-run equilibrium accumulation rate.

## 2.6 Technological change

Technological change was an important theme for all the classical economists and especially for Marx. Adam Smith noted 'how much labour is facilitated and abridged by the application of proper machinery' (1776 [1976], p. 13) and argued that opportunities to increase what he called the 'division of labour' gave incentives to develop new or improved types of machinery. For Ricardo, technological improvements in agriculture and mining were essential to offset diminishing returns and prevent a society's decline into what he called the 'stationary state', and thereby to sustain high wages *and* profits (as discussed in section 2.7, below). However, Ricardo (1821 [1951], Chapter 31) also worried that the use of labour-saving machinery could create structural unemployment by making some workers redundant.

Marx saw contradictory aspects of technological change. On the one hand, he believed that it was the historical mission of capitalism to produce the greatest revolutions in technology in human history. As he and Frederick Engels wrote in the *Communist Manifesto*, 'The bourgeoisie, during its rule of scarce one hundred years, has created more massive and more colossal productive forces than have all preceding generations together' (Marx and Engels, 1968, pp. 39–40). Marx presciently foresaw the tendency of technological innovation to lead to ever-greater increases in the productivity of labour and the automation of production. On the other hand, Marx echoed and amplified Ricardo's preoccupation with the labour-displacing impact of mechanization; he also theorized that the adoption of labour-saving but capital-using technologies would ultimately unleash the FTRP, which could spell the doom of capitalism unless arrested by certain 'counteracting tendencies' (Marx, 1894 [1981], Chapters 13–14). In addition, Marx recognized the endogenous character of technological innovation, which he saw as responding to both the pressures of high wages that threatened profits and the scientific logic of technological evolution itself (see Rosenberg, 1976).

### 2.6.1 Types of technological change

In spite of their emphasis on the evolution of technology, the classical economists and Marx were not always clear in their conceptualizations of the nature of technological change, so it fell to later generations of economists to define more precisely the alternative types. Today, we distinguish four main types of process-oriented technological change, each named after the economist who is credited with identifying it (even though there may have been earlier anticipations). These types, which are illustrated in the four panels of Figure 2.9, are as follows:

1) *Harrod-neutral*: Pure labour-saving (or labour-augmenting) technological change, which raises labour productivity $Q = 1/a_0$ but leaves capital productivity $1/a_1$ unchanged (thus, capital intensity $k = a_1/a_0$ rises). One can think of this as the adoption of new and improved machines that don't produce more output per machine, but which can be operated by fewer workers. Named after twentieth-century British economist Sir Roy Harrod (whose other work was briefly discussed in Chapter 1, and will be covered in greater depth in Chapter 3).
2) *Hicks-neutral*: Factor-saving (augmenting) technological change, which raises the productivity of both labour and capital proportionately, so that $Q = 1/a_0$ and $1/a_1$ both rise by the same percentage and capital intensity $k = a_1/a_0$ remains unchanged. One can think of this as the adoption of new machines that are intrinsically more efficient (produce

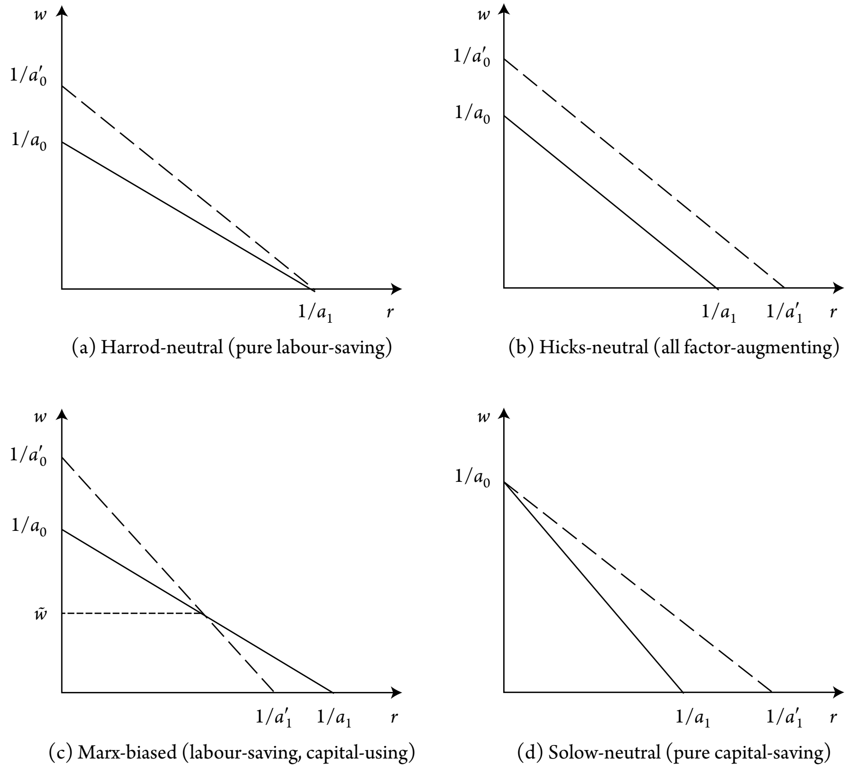



**Figure 2.9** Four types of technological change: (a) Harrod-neutral, (b) Hicks-neutral, (c) Marx-biased, (d) Solow-neutral (–-- and ′ indicate a new technique of production)

more output per machine) and also require less labour to operate them, in the same proportions. Named after another twentieth-century British economist, John (J.R.) Hicks.

3) *Marx-biased*: Technological change that is labour-saving but capital-using, so that the productivity of labour $Q = 1/a_0$ rises but the productivity of capital $1/a_1$ falls and capital intensity $k = a_1/a_0$ increases sharply. This refers to the replacement of labour by large-scale machinery that makes labour more productive, but which raises capital costs even while reducing labour costs. This type of technical change is named after Marx (1867 [1976]), who referred to it as an increase in the 'organic composition of capital'.[21]

4) *Solow-neutral*: Pure capital-saving technological change, which raises the productivity of capital $1/a_1$ but leaves the productivity of labour $1/a_0$ unchanged so that capital intensity $k = a_1/a_0$ falls. A contemporary example is the adoption of newer, cheaper and more efficient personal computers, which lower capital equipment costs for employers, but still require

the same number of workers to yield the same output (for example, to input a page of text). Although named after American economist Robert M. Solow, the progenitor of the 'old' NGT model covered in Chapter 1, the concept was anticipated by Marx in his recognition of capital-saving technological change as a 'counteracting tendency' to his FTRP.

In the rest of this section, we will examine how technological changes of these various types affect the distribution of income and the rate of growth, under the alternative closures of the classical-Marxian model introduced earlier. In reality, of course, technological change is very complex, differentiated across sectors, endogenous in response to economic incentives, and continuous (albeit at varying rates) over time. But for analytical simplicity, we will limit our discussion here to comparisons of long-run equilibria before and after the introduction of a single, one-time innovation of each type at the aggregate level, in order to clarify how the impact varies according to both the type of technological change and the nature of the underlying economy (especially in regard to the determination of income distribution). We will not cover all possible combinations of technological changes and model closures, but rather will focus on some of the most important cases and leave the others to the reader as exercises.

## 2.6.2 A note on the choice of technique

Before proceeding to analyse the effects of each type of technological change, it is important to specify the conditions under which firms will or will not adopt a given new technique. Most generally, a firm will adopt a new technique if doing so will increase the firm's expected future stream of profits (discounted to the present) compared with maintaining the existing technique of production. At such a broad level, the choice of technique involves many difficult questions, including uncertainty about likely future revenues and costs, and the need for the gains in reduced operating (marginal) costs to more than cover the extra fixed costs of acquiring and implementing a new technology. Here, we perform the heroic simplification of ignoring both uncertainty and fixed costs, and assume that firms adopt a new technique as long as it offers a higher rate of return (profit rate) at the existing price vector – which in the simplified, one-commodity model of this chapter, is simply the current real wage.

Under this criterion, the choice of technique is a simple matter for three of the four types of technological change. Harrod-neutral, Hicks-neutral and Solow-neutral innovations all push out the wage–profit relation to the right everywhere except possibly at one intercept point, so that at any feasible real wage ($0 < w < 1/a_0$), the new technique will unambiguously raise the profit rate. For Marx-biased innovations, however, the decision to adopt a new technique is non-trivial. As shown in panel (c) of Figure 2.9, the wage–profit relation for the new technique crosses the old one, and there is a unique 'switchpoint' real wage of $\widetilde{w}$ at which the two techniques (old and new) offer exactly the same profit rate. If the actual wage were $\widetilde{w}$, then firms would be indifferent between adopting the new technique and keeping the old one. For any real wage below $\widetilde{w}$, the old technique promises higher profits (essentially, because labour costs are so low that the savings in wages are more than offset by the higher capital costs), so firms will not switch to the new one. Only if the real wage is greater than $\widetilde{w}$ will the savings in labour costs exceed the increased capital costs so that the profit rate will be higher, and only then will firms be willing to adopt the new technique. In the latter case, we say that the new technique is 'viable' at the prevailing real wage. Thus, firms will be driven to adopt labour-saving and capital-using technological innovations only in a high-wage environment.[22]

As we will see below, the ultimate impact of a technological change in the new, long-run equilibrium depends on the extent to which the real wage eventually rises relative to the increase in the productivity of labour, and how this affects the equilibrium profit rate in the long run. In certain situations, the profit rate may fail to rise or even fall in the new long-run equilibrium, after all adjustments take place. This, of course, calls into question whether it is realistic to assume that firms do not attempt to form expectations about such adjustments, which could cause them not to adopt a new technique that would otherwise appear to increase profits at the initial, current real wage. We shall return to this issue when we come to such cases below.

### 2.6.3 Neutral technical changes under alternative distributional closures

In this subsection, we consider the three types of neutral technical shifts (Harrod, Hicks and Solow) together, as they involve common issues that can be resolved with similar modelling techniques. We will study the effects of such technical shifts on the long-run equilibrium pattern of growth and distribution in each of the three 'closures' in which the profit rate is an endogenous variable (we will not consider the closure with an exogenously given profit rate here; Marx-biased changes are covered in the following subsection).

If we assume the first distributional closure – a fixed real wage – then the outcome is trivially simple: the profit rate must increase for any of the three neutral types of technological change. As can clearly be seen in panels (a), (b) and (d) of Figure 2.9, if the real wage stays constant (at any given level

$0 < w < 1/a_0$), then the profit rate will rise upon the adoption of any one of these types of innovation and will remain higher. Less obviously, in two of these three cases the profit share of national income must also rise. Recall that the profit share can be written as $\pi = 1 - wa_0$, so if we substitute a given real wage $\overline{w}$, then $\pi = 1 - \overline{w}a_0$ must increase when the labour coefficient $a_0$ falls (and labour productivity $1/a_0$ rises) under Harrod- or Hicks-neutral technical change. This makes intuitive sense, because if labour becomes more productive but workers do not share in the fruits of their higher productivity to any degree, then the firms that employ them must be reaping a larger share of the income generated by their labour. However, if the technical change is Solow-neutral, then $\pi$ does not increase because the labour coefficient $a_0$ remains constant. The profit rate still rises, however, because of the increased productivity of the firms' capital. Recall that, by equation (2.5), the profit rate can be written as $r = \pi/a_1$, so a reduction in $a_1$ (rise in capital productivity, $1/a_1$) will increase $r$ even if the profit share $\pi$ stays constant. In this case, firms get higher profits relative to their capital because their capital costs are lower, even though labour costs have not fallen. As long as the profit rate increases and $r > r_{min}$, the economy's growth rate $g$ will also increase per equation (2.8) under any of these three types of technological change.

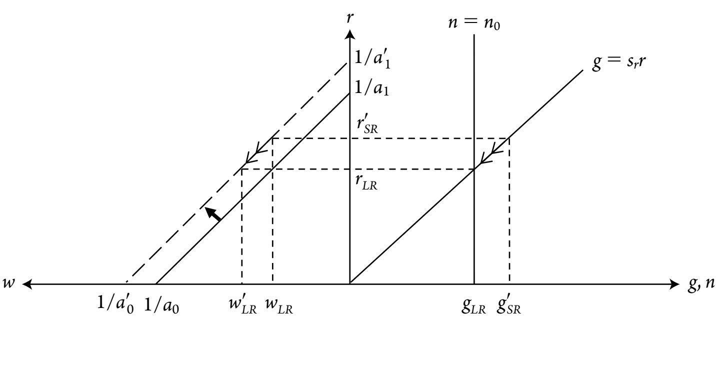
If the wage *share* is fixed instead of the real wage, then the outcome changes notably in each case. To see what happens, recall that in this case the real wage can be written as $w = (1-\overline{\pi})/a_0$. With the wage share $1 - \overline{\pi}$ held constant, the real wage will rise if and only if the labour coefficient $a_0$ falls (labour productivity $1/a_0$ rises), which will occur under either a Harrod- or Hicks-neutral technical shift. The real wage does not rise, however, under a Solow-neutral change, because the latter would leave the labour coefficient $a_0$ unchanged. In contrast, what happens to the profit rate in the long run depends *only* on what happens to the capital coefficient or capital productivity, since in this case the profit rate equals $r = \overline{\pi}/a_1$ and the profit share $\overline{\pi}$ is held fixed. Thus, the profit rate will rise as long as the technical change is Hicks- or Solow-neutral, but not if it is Harrod-neutral, since the latter leaves $a_1$ unchanged. Thus, the growth rate also increases under Hicks or Solow neutrality, but not under Harrod neutrality, in which case the relative shares of wages and profits remain constant.

Because a Hicks-neutral technical change raises the productivity of both labour and capital, workers and firms can both benefit and share proportionally in the productivity gains, when their relative shares of national income are fixed; the real wage and profit rate can both increase. This shows the possibility that technological progress can potentially benefit both social classes and relieve distributional conflict in a classical model, but only if the technical change is Hicks-neutral. In contrast, because a Harrod-neutral change is purely labour-saving, it tends to benefit workers more when they can achieve a constant share of national income (the real wage rises but the profit rate returns to its original level), whereas because a Solow-neutral shift is purely capital-saving, it tends to benefit firms more as their profit rate rises while the real wage remains constant.

The response of the economy to a technical shift is more complicated in the constant unemployment rate or natural rate of growth closure, so to fix ideas we will begin with a special, simplified case. That is, suppose that $n_1 = 0$ (labour supply is unresponsive to the real wage), so that the labour supply grows at the exogenously fixed rate $n = n_0$. Also assume that there is no minimum profit rate required for positive saving, so $r_{min} = 0$ and growth is directly proportional to the profit rate: $g = s_r r$. As shown previously, in this case the profit rate must equal $r_{LR} = n_0/s_r$ in long-run equilibrium, and the long-run equilibrium growth rate (2.21) is also constant at $g_{LR} = n_0$. In this special case, a technological change cannot affect either $r$ or $g$ in the long run! Therefore, any short-run increase in profits upon the adoption of a new technique must be transitory, and is eventually eroded by rising real wages induced by the temporary boost to growth. The wage only stops rising once profits and growth return to their long-run equilibrium rates.

These dynamics are illustrated in Figure 2.10 for the case of a Hicks-neutral technical change (Harrod- and Solow-neutral changes would only differ in the way the $w - r$ relation shifts in the left-hand quadrant, but the rest of the analysis would be similar and is left to the reader as an exercise). At the initial wage rate $w_{LR}$, the profit rate would immediately rise to $r'_{SR}$ after adoption of the new technique, and with more rapid capital accumulation enabled by higher profits, the growth rate would increase to $g'_{SR}$ in the short run. However, the faster growth would increase demand for labour, thereby pushing up the real wage and lowering the profit rate (as shown by the thin

**Figure 2.10** A Hicks-neutral technical change in the full-employment closure, special case of a fixed growth rate of labour supply and no minimum profit rate



arrowheads). Eventually, the profit rate returns to its initial level of $r_{LR}$, while the real wage remains permanently higher at $w'_{LR}$.

Although the real wage rises and the profit rate returns to its (constant) long-run equilibrium level for all three types of technical change, the impact on income distribution in the sense of relative shares differs among the three types (still in the simplified case of an exogenously growing labour supply). In this case, solving equation (2.5) for the profit *share* yields $\pi = r_{LR}a_1$. With the profit rate fixed at $r_{LR} = n_0/s_r$ in the long run, the profit share must change in the same direction as the capital coefficient $a_1$. Thus, either a Hicks- or Solow-neutral technical change lowers the long-run equilibrium profit share (because either type implies a fall in $a_1$), thus making income distribution more equal, while a Harrod-neutral change leaves the profit share unchanged (because $a_1$ remains constant) and therefore has no effect on distributional equity.

In the more general case of the constant unemployment rate/natural rate of growth closure, however, matters are more complicated because a change in technology alters the way that the growth of labour supply function (2.18), which is written in terms of the real wage, maps into the profit rate. In the general case, the labour supply function (as shown in the first quadrant of Figure 2.3) is

$$n = n_0 + n_1\left(\frac{1 - a_1 r}{a_0}\right) = \underbrace{\left(n_0 + \frac{n_1}{a_0}\right)}_{\text{intercept}} - \underbrace{n_1\left(\frac{a_1}{a_0}\right)}_{\text{slope}} r \qquad (2.22)$$

where the last part of this equation expresses the function in slope-intercept form. A Hicks-neutral shift (which lowers $a_0$ and $a_1$ proportionately) increases the intercept of the labour supply function on the *g*-axis without changing its slope, so the function shifts to the right in a parallel fashion. A Solow-neutral shift (which lowers $a_1$ but leaves $a_0$ unchanged) makes the labour supply function rotate to the right, pivoting on a constant intercept on the *g*-axis, as drawn. A Harrod-neutral shift will increase the intercept of the labour supply function and also rotate it to the left and make it flatter (given the way this line is drawn in Figure 2.3).

The long-run equilibrium values of *r* and *g* clearly increase in the first two cases (Hicks- and Solow-neutral technical changes), because in both cases the labour supply curve shifts or rotates out to the right; less obviously, it can be shown that *r* and *g* must also increase in the long run as a result of a Harrod-neutral shift (in other words, the shifted labour supply curve must, in spite of its flatter slope, intersect the capital accumulation curve above and to the right of the initial equilibrium in the first quadrant of Figure 2.3).[23] Intuitively, the induced increase in labour supply growth helps to relieve the upward pressure on the real wage resulting from more rapid capital accumulation, thereby allowing the profit rate to remain higher in the long run. But, correspondingly, the increase in the real wage should be smaller when the labour supply thus responds by bringing more workers into the economy to compete with existing ones.

### 2.6.4 Marx-biased technical change and the falling tendency of the rate of profit

Marx famously claimed that technological change under capitalism would tend to be of a labour-saving but capital-using variety, which would (under certain conditions) give rise to the FTRP. In this subsection, we investigate the logic of his claim, as well as his suggestions about possible 'counteracting tendencies'. In fact, more ink has probably been spilt on this topic than any other subject in this chapter, and we will not review all the major arguments here. However, our modelling framework does enable us to clearly identify the conditions under which the profit rate will or will not fall as a result of this type of technological change, at least in terms of the alternative distributional closures; whether most technological change is of this nature is beyond the scope of the present discussion.[24] Here, as in the previous subsection, we restrict our discussion to a one-time technological change for simplicity.[25]

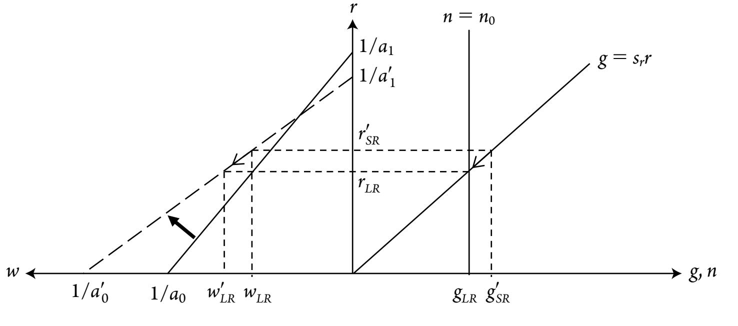
If the real wage is initially greater than the switchpoint level $\widetilde{w}$ shown in Figure 2.9(c) so that a Marx-biased technical change is viable, then evidently the profit rate must increase rather than decrease if the real wage remains constant after the change is adopted.[26] Although this may seem very straightforward, much controversy once surrounded this seemingly simple point. The difficulty arose from the use of multisectoral models (as opposed to the simple, single-sector model used here), in which it is not immediately obvious that the adoption of a viable, profit-increasing new technique by a single firm (or in a single industry) – based on the prevailing prices of inputs and outputs at the time of its adoption – will necessarily raise the economy-wide, equilibrium profit rate after all adjustments in prices have taken place. Okishio (1961) proved that in fact this was generally true: provided that the real wage remains constant, if a new technique raises the profit rate for a single producer, it cannot lower and in general will increase the new (equalized) equilibrium profit rate for the system as a whole.[27] Based on this type of logic, Roemer (1981) argued that Marx's FTRP theory was simply wrong: if anything, a Marx-biased technological change should raise rather than lower the rate of profit (assuming a constant real wage).

Roemer (1981, p. 134) also allowed that, 'if the real wage rises, however, [then] the equilibrium rate of profit may fall'. But that is exactly what *must* happen when the wage *share* is held constant instead of the real wage. In fact, Marx did *not* assume a fixed real wage in his theory of the FTRP. On the contrary, he assumed constancy of what he called the rate of surplus value or rate of exploitation, which is equivalent (if we translate Marx's labour-value concepts into modern national income accounting) to the ratio of profits to wages, and therefore he implicitly assumed constant shares of profits and wages in national income. Recall that, when the wage and profit shares are constant, the profit rate can be written as $r = \bar{\pi}/a_1$. Because the capital-using aspect of Marx-biased technical change implies a *rise* in $a_1$, it necessarily follows that the equilibrium profit rate *must* fall under this distributive closure. Furthermore, recalling that in this case the real wage can be written as $w = (1 - \bar{\pi})/a_0$, it is also clear that the real wage *must* rise because of the labour-saving aspect of the change, which implies a fall in $a_0$. Thus, Marx and Roemer are both right: with constant relative shares, a Marx-biased technical change does cause the profit rate to fall, and this is accompanied by a rise in the real wage (without which the wage and profit shares would not remain constant and the profit rate would not fall).[28] It is also evident that in this case the real wage must rise *in exact proportion* to the increase in labour productivity ($w$ has to rise by the same percentage as $a_0$ falls), since by definition $\bar{\pi} = 1 - a_0 w$.

However, this does not prove that there is in fact an ineluctable tendency for the profit rate to fall, even if technological change takes a Marx-biased form. Indeed, Marx (1894 [1981], Chapter 14) noted various counteracting tendencies in volume III of *Capital*. Most of these involve some means of suppressing wages, thereby lowering the wage share and raising the profit share, or else adopting capital-saving techniques that lower capital costs (thus anticipating what are now known as Solow-neutral technological shifts). What Marx did not acknowledge, however, is that some of these counteracting tendencies could be set in motion by the same forces unleashed by the labour-saving, capital-using technology in the first place. Consider what would happen if, initially, the wage share remains constant and the profit rate falls, as contemplated in the FTRP analysis. Because capital accumulation is a function of the profit rate per equation (2.8), the overall growth of the economy would slacken. But, this in turn would reduce the growth of labour demand and increase unemployment. In this situation, it is hard to see how workers could maintain the increase in their real wage needed to sustain a constant share of national income in the face of slower employment growth. Thus, the new equilibrium implied by the assumption of a constant wage share is likely to be unsustainable in this situation.

This brings us to the third distributional closure, in which the growth rates of labour demand and supply must be equal in the long run and the economy grows at the natural rate. With reference to Marx, this might better be called the 'industrial reserve army' closure, since it shows explicitly how excess growth of labour supply relative to demand can depress wages. As noted earlier, Marx's own comments on the rate of accumulation determining changes in the real wage are a key foundation for this model. And, as also noted earlier, the equilibrium condition $g = n$ can hold regardless of whether the equilibrium unemployment rate corresponds to 'full employment' or some other positive fraction of the labour force. Therefore, the $g = n$ closure can be considered to include the case of any constant unemployment rate (not just 'full employment').

For Marx-biased technical change, we will confine our exposition to the simple case of an exogenously growing labour supply, $n = n_0$. In this situation, which is depicted in Figure 2.11, neither the growth rate nor the profit rate can change in the long run, because the former must settle at $g_{LR} = n = n_0$ and the profit rate must end up at $r_{LR} = n_0/s_r$. It is easily seen that, if a Marx-biased technical change is viable (that is, if it would raise the profit rate at a given real wage), then it must also raise the real wage at a given profit rate. However, in this case the rise in the real wage must be less than proportional to the rise in labour productivity, because with capital productivity falling (the capital coefficient $a_1$ increasing), the profit share $\pi$ has to rise in order for the profit rate to remain constant at $r_{LR} = \pi/a_1 = n_0/s_r$. And, since the profit share can be written as $\pi = 1 - wa_0$, it can only rise if the increase in $w$ is less than proportional to the fall in $a_0$ (or rise in $1/a_0$). So, in this situation,




Note: Light arrowheads show adjustment from the new short-run equilibrium to the long-run equilibrium.

**Figure 2.11** Marx-biased technical change with a constant unemployment rate, special case of a fixed growth rate of labour supply and no minimum profit rate

Marx's FTRP does not hold – the profit rate does not fall in the long run but instead returns to $r_{LR}$.[29] However, in this situation another one of Marx's predictions, the relative 'immiseration of the proletariat', does come true: workers end up with a smaller share of national income $(1 - \pi)$, even though in absolute terms they get a higher real wage $w$.

Where Marx erred, then, was not in seeing the possibility of an FTRP under the conditions he assumed (essentially, a constant wage share), but rather in implying that both an FTRP and relative immiseration (which we are interpreting as a falling wage share) could be observed at the same time, as a result of labour-saving, capital-using technical change. If the wage share stays constant, then the profit rate falls but there is no immiseration (the real wage increases by as much as labour productivity); in order to get relative immiseration, the real wage must grow more slowly than labour productivity so that the wage share falls, but in this case there is no FTRP as the profit rate stays constant in the long run. Moreover, Marx did not see that if an FTRP resulted, even temporarily, it would be likely to cause an offsetting fall in the wage share – *not* as an *independent* counteracting tendency, but endogenously as a result of the fall in the growth of labour demand resulting from the same initial technological shift. Indeed, this is precisely what occurs if we invoke the logic of Marx's 'industrial reserve army' hypothesis and model the adjustment of labour demand and its impact on the wage explicitly.

## 2.7 The Ricardian stationary state

Ricardo (1821 [1951]) is well known for his pessimistic view of the long-run prospects for economic growth, enshrined in his famous theory of a stationary state in which growth would cease and wages would fall to a bare subsistence level. This state would allegedly be reached because, in the long run, continued economic growth would press on scarce supplies of natural resources (poetically rendered as the limited 'fertility of the soil'), leading to a simultaneous fall in the profit rate and the real wage. This theory thus provides an alternative to Marx's view of an FTRP, but ironically one in which technological change is the potential saviour rather than the villain. Properly understood, Ricardo's theory of a stationary state is more of a warning rather than a prediction, however, as his model allows for much more optimism than Malthus's (1798 [1993]) dire prophecy of perpetual misery for the working class. Ricardo's model can also provide a classical foundation for ecological economics, as we shall see.

We will formulate Ricardo's model of the stationary state based on the interpretation of Casarosa (1978), slightly reformulated to match some of the mathematical specifications used in this chapter.[30] First, we will use a simplified version of our saving (accumulation) function omitting the minimum rate of profit (equivalent to assuming $r_{min} = 0$ in equation 2.8):

$$g = s_r r \tag{2.23}$$

Substituting equation (2.5) for the profit rate, the growth (accumulation) rate can be expressed as a decreasing function of the real wage:

$$g = s_r\left(\frac{1 - wa_0}{a_1}\right) \tag{2.24}$$

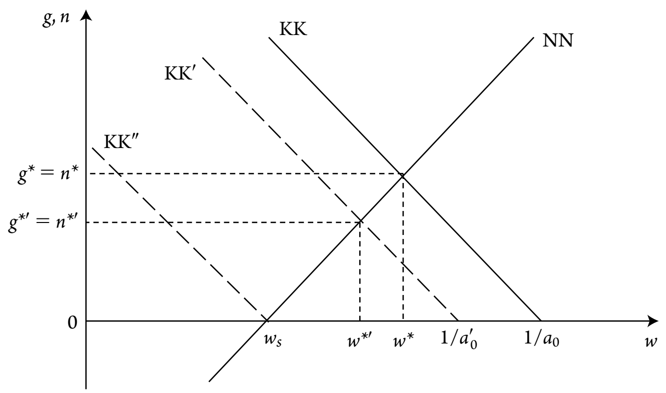
Following Casarosa (1978), we will reformulate our model of endogenous labour supply to portray the growth rate of the labour force as an increasing function of the proportionate gap between the actual real wage $w$ and the subsistence level $w_s$:[31]

$$n = n_1\left(\frac{w - w_s}{w_s}\right) \tag{2.25}$$

where $n_1 > 0$ is a positive constant and the absence of a constant term implies that $n = 0$ when $w = w_s$. Thus, population (labour force) growth ceases when the wage is at the subsistence level.

Equations (2.24) and (2.25) can be graphed as shown in Figure 2.12, where the two growth rates (of capital and labour) are plotted on the vertical axis and the real wage is on the horizontal axis. The curve KK represents equation (2.24) for capital accumulation, while NN represents (2.25) for labour supply growth. Note that KK has a horizontal intercept at $Q = 1/a_0$, which equals both the productivity of labour and the maximum feasible real wage, while NN has a horizontal intercept at the subsistence wage $w_s$. The real wage

**Figure 2.12** The Ricardian stationary state



adjusts according to equation (2.11), with the equilibrium condition $\hat{w} = 0$ or, equivalently, $g = n$ (equation 2.12). Given the technology and resources available at any point in time, the steady-state equilibrium thus corresponds to the point where KK and NN intersect. This yields an equilibrium real wage $w^*$ and corresponding equilibrium growth rates of capital and labour $g^* = n^*$.

In general, equilibrium growth will be positive and the equilibrium wage will be above the subsistence level. However, if rapid growth is sustained for a long period of time, the society will begin to strain its natural resources. The best farmlands and mineral deposits will have been used, forcing producers to rely on less fertile soils and less easily available minerals, or else inducing them to try to extract more crops and minerals from existing lands and mines,[32] all of which would imply increasing marginal costs or decreasing returns to combined inputs of labour and capital.[33] Such diminishing productivity of factors can be represented as equivalent to a Hicks-neutral technological retrogression, or a parallel inward shift of the wage–profit inverse relation (hence, the opposite direction of the shift shown in panel (b) of Figure 2.9). Mathematically, this means that $a_0$ and $a_1$ both increase by the same proportion. In Figure 2.12, such a Hicks-neutral reduction in productivity is represented by the shift of the KK curve down and to the left to KK′, with the equilibrium real wage falling to $w^{*\prime}$ and the equilibrium growth rate decreasing to $g^{*\prime} = n^{*\prime}$.[34]

As long as positive growth continues, and holding technology and resources constant, primary production (agriculture and mining) will eventually experience further diminishing returns, until KK shifts down and to the left to KK″, where the equilibrium real wage falls to the subsistence level ($w = w_s$) and growth of both capital and labour ceases ($g = n = 0$). *This is the Ricardian stationary state*. Although it may not be immediately apparent, the profit rate also falls in the process of reaching this stationary state. Since the profit rate is directly proportional to the growth rate, $r = (1/s_r)g$, $r$ must decline as $g$ decreases.[35] As Ricardo (1821 [1951], p. 120) put it, 'The natural tendency of profits then is to fall; for, in the progress of society and wealth, the additional quantity of food [and other primary products] required is obtained by the sacrifice of more and more labour.'[36]

Ricardo made it clear, however, that two forces could prevent the collapse of a society into the stationary state and permit positive growth (and above-subsistence wages) to persist indefinitely. One of these factors is technological progress, which by the logic of Ricardo's model would have to occur in the primary sectors (agriculture and mining) or in the efficiency of primary commodity use to prevent diminishing returns from occurring. In his own words, the FTRP 'is happily checked at repeated intervals by the improvements in machinery, connected with the production of necessaries, as well as by discoveries in the science of agriculture . . .' (Ricardo 1821 [1951], p. 120). Today, if we think of fossil fuels as one of the crucial primary sectors subject to diminishing returns, the development of renewable sources of energy such as wind and solar power to replace fossil fuels would be an equivalently beneficial type of innovation. Technological innovations in agriculture, mining and energy could shift KK up and to the right, or at least prevent it from shifting down and to the left in the first place. The other solution Ricardo saw was international trade: if a country could import food and raw materials instead of having to produce more of those goods at home, it could prevent diminishing returns from occurring in its own primary producing sectors, again keeping the KK curve from shifting inward and preserving an equilibrium with positive growth and high real wages.

Hence, as Ricardo wrote, 'If, therefore, by the extension of foreign trade, or by improvements in machinery, the food and necessaries of the workers [and other primary products, such as raw materials] can be brought to market at a reduced price, profits will rise' (Ricardo 1821 [1951], p. 132). Ricardo also reflected such optimism when he stated,

> Notwithstanding the tendency of wages to conform to their natural [subsistence] rate, their market [equilibrium] rate may, in an improving [growing] society, *for an indefinite period, be constantly above it*; for no sooner may the impulse, which an increased capital gives to a new demand for labour be obeyed, than another increase of capital may produce the same effect; and thus, if the increase of capital be gradual and constant, the demand for labour may give a continued stimulus to an increase of people. (Ricardo 1821 [1951], p. 95, emphasis added)

However, there are a few caveats concerning the interpretation and application of Ricardo's model of the FTRP and stationary state. First, the labour supply function relies on the discredited Malthusian law of population, according to which higher real wages induce faster population growth and more rapid increases in the labour force. In modern capitalist economies, higher real wages and living standards generally have the opposite effect, inducing reduced fertility (birth) rates and slower growth of both population and labour supply in the long run.[37] Second, the option of averting the FTRP and the stationary state via international trade is only available to those countries that have a comparative advantage in manufactures and would import primary products, such as England in Ricardo's time (hence, his strong advocacy of free trade for the UK).

For countries that have a comparative advantage in agriculture or minerals and would export them, free trade could actually cause their real wages and profit rates to fall, resulting in at least slower growth, if not an eventual stationary state.[38] This would occur because those countries would have to devote more of their own natural resources to agricultural or mineral production, which are subject to diminishing returns. Hence, the trade option solves the problem only for one group of countries, and it does so at the expense of widening global inequality by raising growth rates and living standards in one group of countries (manufacturing exporters) and reducing them in the other group (primary product exporters). Only technological innovation can avert the FTRP and stationary state in all countries and at a global level, and ironically, it is most important in the primary product exporting nations where diminishing returns would otherwise occur (or in the consumption of those products in the manufacturing nations, for example in their energy usage).

In spite of these reservations, Ricardo's model of the stationary state still has much to teach us about economics and ecology. To help us see the connection, recall that in Ricardo, the so-called natural or subsistence wage is not fixed at biological minimum; it is not Malthus's minimal level for human beings to avoid starvation and be able to reproduce. Rather, he emphasized that what he called 'the natural price of labour' (our $w_s$) 'varies at different times in the same country, and very materially differs in different countries. It essentially depends on the habits and customs of the people' (Ricardo 1821 [1951], pp. 96–7). If global society can achieve zero (or very low but sustainable) population growth at a reasonably high real wage, permitting a comfortable standard of living, while maintaining slow (not necessarily zero, but slow and ecologically sustainable) growth of capital and output, that could be a very desirable outcome for humanity that might help to avert the current threats of global warming and conflict over scarce resources that plague our planet today. Moreover, minimizing the need to exploit scarce natural resources, for example by switching to non-renewable energy sources, could reduce the high rents that contribute so much to global inequality and political tensions.

## 2.8 Neo-Marxian Goodwin cycles and the profit-squeeze hypothesis

Most of this chapter has focused on the classical and Marxian approaches considered as models of long-period growth. However, one part of Marx's analysis of capital accumulation – Chapter 25 in Volume I of *Capital* (Marx, 1867 [1976]) – can be interpreted as implying a model of business cycles or cyclical growth. In that chapter, Marx described how the real wage rises in an economic expansion (a period of high profitability and rapid accumulation, implying fast growth of demand for labour), but the resulting fall in the profit share (rate of exploitation) eventually diminishes the profit rate to the point where accumulation (growth) slows down, thus causing a downturn in the cycle (a recession, in more modern parlance). This mechanism has come to be known as the 'profit-squeeze' hypothesis about cyclical downturns. Then, as accumulation slows and unemployment rises in the downturn, workers' bargaining power in the class struggle is reduced, and the real wage starts to fall. Eventually, the fall in the real wage restores profitability, as the profit share and profit rate turn back upward, which induces accumulation to pick up steam and employment to increase again in a new phase of recovery and expansion. Marx saw such cyclical behaviour as a key instrument of capital (along with labour-saving technological change) for restraining labour costs and maintaining profitability.

Inspired by Marx's discussion, Goodwin (1967) produced a mathematical model of such a cyclical profit-squeeze mechanism.[39] To see how this model works, we need to start by modifying the equation for the rate of change in the real wage from what we specified previously. In equation (2.11) we assumed that the real wage changes in proportion to the *gap* between the *growth rates* of labour demand (assumed to equal the rate of capital accumulation, $l = g$) and labour supply (which grows at the rate $n$). For the Goodwin model, we need to assume instead that the rate of increase in the real wage is a function of the *level* of the *employment rate*, defined (as in Chapter 1) as the ratio of employed workers to the labour force, $e = L/N$:

$$\hat{w} = -\gamma + \xi e \tag{2.26}$$

where $\gamma$ and $\xi$ are positive constants. Equation (2.26) resembles a wage Phillips curve, except it is specified in terms of the change in the real wage rather than the nominal wage and it uses the employment rate rather than the unemployment rate (hence, it is upward sloping).

For Goodwin's model, it is more convenient to use the wage share $\psi$ rather than the profit share $\pi$, but they are monotonically inversely related to each other since $\psi = 1 - \pi$. By definition, $\psi = wa_0$, so the wage share changes according to $\hat{\psi} = \hat{w} + \hat{a}_0$. Then, using equation (2.26) for $\hat{w}$ and recalling that $q = \hat{Q} = -\hat{a}_0$ is the rate of labour productivity growth, we can obtain the following expression for the rate of change in the wage share:

$$\hat{\psi} = -(q + \gamma) + \xi e \tag{2.27}$$

The employment rate changes according to the difference between the growth rates of labour demand and labour supply, $\hat{e} = \hat{L} - \hat{N} = l - n$. Assuming for simplicity that all profits are saved and invested $(s_r = 1)$ with no minimum profit rate required for saving to occur, the rate of capital accumulation equals the profit rate $(g = r = \pi / a_1)$. Then, assuming that $a_1$ remains constant,[40] labour demand grows at the rate $l = g - q$, while labour supply is assumed to grow at the exogenous rate $n$. Using the fact that $\pi = 1 - \psi$ by definition, the rate of change in the employment rate is then

$$\hat{e} = \frac{1-\psi}{a_1} - (n + q) \tag{2.28}$$

where it may be noted that the employment rate rises if the capital accumulation rate $((1-\psi)/a_1)$ exceeds the natural rate of growth $(n + q)$ and falls if the latter exceeds the former.

Equations (2.27) and (2.28) constitute a system of two simultaneous, first-order differential equations in $\psi$ and $e$. The equilibrium solutions are easily found: setting $\hat{\psi} = 0$ implies

$$e^* = \frac{q + \gamma}{\xi} \tag{2.29}$$

while setting $\hat{e} = 0$ yields

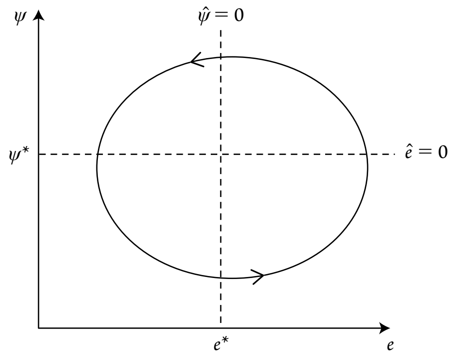
$$\psi^* = 1 - a_1(n + q) \tag{2.30}$$

Note that the equilibrium level of the wage share is a decreasing function of the natural rate of growth, that is, the growth rate of the effective labour supply (including workers made redundant by labour-saving technical change), whereas the equilibrium employment rate is an increasing function of the rate of productivity growth (because the latter enhances profitability and encourages accumulation).

The dynamics of the system (2.27) and (2.28) imply that the economy never reaches the equilibrium described by (2.29) and (2.30), but instead cycles perpetually around it, as shown in Figure 2.13.[41] To see why this occurs, first note the peculiar characteristic that in (2.27) and (2.28) the rates of change $\hat{\psi}$ and $\hat{e}$ are functions *only* of the level of the *other* variable; neither is a function of its own level. The Jacobian matrix of the first partial derivatives of (2.27) and (2.28) is

$$\mathbf{J} = \begin{bmatrix} 0 & \xi \\ -1/a_1 & 0 \end{bmatrix} \tag{2.31}$$

**Figure 2.13** A neo-Marxian Goodwin cycle in the employment rate ($e$) and wage share ($\psi$)




with trace $Tr(\mathbf{J}) = 0$ and determinant $Det(\mathbf{J}) = \xi / a_1 > 0$. This system describes a 'limit cycle' or closed orbit around the equilibrium point, where the cycles are neither damped nor explosive, as shown in Figure 2.13. The counterclockwise rotation of the employment rate and wage share represents a business cycle driven by increases and decreases in the wage share in response to increases and decreases in capital accumulation and employment – which is the profit-squeeze mechanism described above – while capital accumulation responds inversely to the wage share. We will call this a 'neo-Marxian' or 'original' Goodwin cycle, for contrast with the neo-Kaleckian Goodwin cycle (also called a 'neo-Goodwin cycle') developed by Barbosa-Filho and Taylor (2006), which will be covered in Chapter 5.[42] Recent studies that have estimated or extended neo-Marxian Goodwin cycle models include Harvie (2000), Veneziani and Mohun (2006) and Grasselli and Maheshwari (2017).

## 2.9 Conclusions

In the approach of the classical economists and Marx, economic growth is driven by the accumulation of capital by a class of business owners or capitalists, and therefore depends critically on the ability of the production process to generate enough of an economic surplus in the form of profits that the capitalists can save and invest in the expansion of the capital stocks of their firms. All else being equal, there is an inevitable trade-off or inverse relationship between the profits of capital and the wages of labour. For a given level of technology, with a given saving propensity of the capitalists and at a normal rate of utilization of the capital stock, it is only possible to devote more resources to accumulation and growth if income becomes more concentrated in the hands of the capitalists (firm owners). However, this is not the end of the classical or Marxian story, because all else is not always equal. Most obviously, if the business class increases its propensity to accumulate – devotes a greater share of its profits to saving and investment – the economy can grow faster without any change in the distribution of income between wages and profits.

More importantly, technological progress is an essential element for allowing the economy to potentially grow faster without having to repress workers' wages in the long run in the classical-Marxian vision (and for avoiding the stationary state in Ricardo's theory). For all of the neutral types of technological changes (Harrod/labour-saving, Solow/capital-saving and Hicks/all-factor-saving) it is possible for both social classes to benefit to some extent (or at least, for either wages or profits to rise, without the other one falling) even though relative shares might change in some cases. For one type of technological change (Marx-biased, or labour-saving and capital-using), however, conflicts can result if workers are able to maintain a constant share of national income, in which case the profit rate would fall and accumulation would decline. But we also saw that other outcomes are possible following a Marx-biased innovation. Especially, if the economy eventually reaches a steady-state equilibrium between the growth of labour supply and demand then the profit rate would not fall in the long run, but inequality would increase as real wages would rise by less than labour productivity and the profit share of national income would increase.

Most of the classical economists (including Smith, Ricardo and Say) rejected any role for demand-side factors in constraining or influencing long-run growth, or even in determining the level of output in the short run. Among the classicals, only Malthus (1820 [1951]) argued that an excessive 'passion for accumulation' could result in shortfalls of aggregate demand leading to involuntary unemployment of labour, but he was unable to demonstrate this point analytically because he lacked a distinction between the saving and investment aspects of accumulation and a corresponding vision of the role of money and other financial assets. Marx (1867 [1976], 1894 [1981]) was the first to explain clearly how the hoarding of money and other financial assets could lock up savings in forms that did not have to be spent on buying new productive assets (capital goods – plant, machinery and equipment), and he also argued that the restricted purchasing power of the working class could limit consumer demand and foster a 'realization crisis' (essentially, the economy producing below its potential and failing to generate the surplus or profits that it has the capacity to produce). Marx also offered a business cycle model, one captured in the Goodwin cycle formulation, which was driven by distributional shifts rather than aggregate demand.

Like his classical predecessors, however, Marx lacked an analytical distinction between saving and investment, and therefore he could not clearly conceptualize how output or profits could effectively be constrained by the level of effective demand or how demand factors could influence growth and distribution in the long run. This is the subject we will take up in the next two chapters, where we turn to neo-Keynesian and neo-Kaleckian growth models. But we will also see that classical-Marxian themes like the centrality of income distribution to the process of capital accumulation and economic growth reappear in those models and others covered in later chapters.

## ? STUDY QUESTIONS

**1)** What assumptions give rise to the idea of an inevitable inverse relationship or trade-off between wages and profits in classical-Marxian theory? What does this relationship imply for the relationship between income distribution and long-run growth, and why?

**2)** Under what circumstances does more rapid economic growth require a reduction in wages (either the real wage or wage share) and under what circumstances can more rapid growth be accompanied by a rise in wages (in either respect), within the classical-Marxian approach?

**3)** How is it possible that technological change can be the cause of a falling tendency of the rate of profit in Marx's theory, and the solution for avoiding such a tendency (and preventing a 'stationary state') in Ricardo's theory? Discuss with reference to the different assumptions about income distribution and technological change in each theory.

**4)** What are the theoretical consequences of the classical tendency to conflate saving with investment as equivalent forms of 'capital accumulation'? Specifically, how would the effects of an increase in the capitalists' propensity to save be different if such an increase was *not* automatically translated into an equivalent increase in investment expenditures?

**5)** In the 1960s, Nicholas Kaldor proposed a set of 'stylized facts' describing the capitalist growth process, which included (among other things) long-run constancy of the profit share and the capital–output ratio. Which versions of a classical-Marxian model, in terms of assumptions about income distribution and technological change, most resemble these stylized facts? Does empirical evidence from the late twentieth and early twenty-first centuries still support such a view?

## NOTES

**1** In the rest of this chapter, the term 'classicals' will refer mainly to Smith and Ricardo; lesser classical authors such as J.-B. Say and T.R. Malthus will be referred to when appropriate. Marx may be considered a classical economist in light of the analytical structure of his models, but he is often classified separately because he can be viewed as having had a distinctive vision of the capitalist economy and as having founded a separate school of thought. For later developments of Marxian theory see Sweezy (1942, 1981), Morishima (1973), Roemer (1981), Foley (1986) and Shaikh (2016), among many others.

**2** This exposition builds upon (and also summarizes more succinctly) certain earlier presentations and comparisons of classical and neo-Marxian approaches, including Harris (1978), Marglin (1984b), Dutt (1990) and Foley and Michl (1999).

**3** The classical (Ricardian) theory of value was developed in modern form by Sraffa (1960). See Pasinetti (1977), Steedman (1977), Harris (1978), Bharadwaj and Schefold (1990), Kurz and Salvadori (2003), Roncaglia (2009) and Sinha (2016), among many others, for later expositions and discussions of Ricardian and Marxian value theory in multisectoral frameworks.

**4** This set of debates is referred to as the 'Cambridge controversies' because the main protagonists were the critics of the neoclassical approach at the University of Cambridge in Cambridge, UK (led by Piero Sraffa and Joan Robinson) and the defenders of that approach at the Massachusetts Institute of Technology in Cambridge, MA, USA (led by Paul Samuelson and Robert Solow). In essence, the critics maintained that when capital goods are heterogeneous and are aggregated in price terms, there is not necessarily an inverse relationship between the total value of the capital stock and the profit rate, as there is in a one-good neoclassical model in which the quantity of capital can be unambiguously measured. Much of the debate revolved around phenomena such as 'reswitching' (the discontinuous adoption of the same technique at different levels of the profit rate) and 'capital reversals' (the adoption of a more capital-intensive technique at a higher rather than a lower profit rate). Ultimately, the neoclassical defenders conceded the logical validity of the critics' arguments, but believed that they were of little practical import and that the 'perverse' relationships would not be observed in optimal states. See Harcourt (1972) and Harris (1978) for accounts of these debates and citations to the original sources.

**5** The amount of labour can be expressed either as a number of workers (assuming a fixed number of hours per day per worker) or as worker-hours; we will often refer to this magnitude as 'workers' for brevity but one can substitute 'worker-hours' if one prefers a more exact measure. In empirical work, this distinction is non-trivial, because worker-hours per week or per year can vary.

**6** Models that assume variable utilization of capacity will be discussed in Chapters 4 and 5, while debates about whether the capacity utilization rate is variable in the long run are considered in Chapter 6. The classical theory implicitly assumes that the labour supply is endogenous, so that available labour does not normally become the binding constraint on production. The endogeneity of labour supply will be discussed more explicitly in sections 2.3 and 2.7 of this chapter.

**7** 'There can be no rise in the value of labour without a fall of profits' (Ricardo 1821 [1951], p. 35). Both Ricardo's text and later interpretations make it clear that this statement assumes a given technology and normal utilization of capacity – and, in Ricardo's own model, a given marginal productivity of labour in agriculture (see section 2.7 for what happens if the latter declines).

**8** See Chapters 3 and 4 for how this inverse relation is used in neo-Keynesian and neo-Kaleckian models. In the neoclassical (NGT) growth model of Solow (1956) and Swan (1956), covered in Chapter 1, an inverse relationship between $w$ and $r$ can be derived as the factor-cost 'dual' to the 'primal' aggregate production function written in terms of $K$ and $L$ inputs. Thus, if the production function $Y = F(K, L)$ obeys standard assumptions of twice differentiability, constant returns to scale and diminishing marginal productivity, it is possible to derive a unit cost function $uc = uc(w, r)$ which implies a downward-sloping relationship between $w$ and $r$ that is convex to the origin for any given level of unit costs (total costs per unit of output) $uc$.

**9** This notation ($s_r$) is intended to represent the propensity to save out of profit income, and it will be used this way throughout the next several chapters (even when investment is treated as distinct from saving), but in this chapter it can be thought of equivalently as the investment rate since the only way of saving in the classical scheme is to acquire additional capital.

**10** The methodology of comparing growth models via alternative closures of an open system of equations dates back to Sen (1963) and was later used by Marglin (1984b), Dutt (1990) and Foley and Michl (1999). Since most of them were comparing a wider range of models (including neoclassical and neo-Keynesian ones) than we are covering in this chapter, they generally started with only two basic equations (the wage–profit and consumption–growth relations) and the saving assumption was part of the 'closure' that varied across models.

**11** In particular, hypothesis (1) below is articulated in their theories of value, while hypothesis (3) is stated in their theories of distribution and growth. Hypothesis (2) is stated only by Marx and (4) was proposed by later authors, as noted below.

**12** Of course, if there is no labour-saving technical change, in which case $q = -\hat{a}_0 = 0$, then a constant wage share implies a constant real wage and equation (2.10) is equivalent to (2.9).

**13** Harris (1983, 1986) argued that an equilibrium between the growth of labour demand and labour supply was implicit in the classical and Marxian theories of growth and distribution. See also Casarosa's (1978) interpretation of Ricardo, developed in more depth in section 2.7 below.

**14** Note that when there is ongoing labour-saving technological change, equation (2.12) becomes $g = n + q$, where $q = -\hat{a}_0$ is the growth rate of labour productivity.

**15** This is the same assumption that was built into the neoclassical growth model of Solow (1956) and Swan (1956), as discussed in Chapter 1, and the neo-Keynesian growth models of Kaldor (1955–56) and Pasinetti (1962), discussed in Chapter 3.

**16** One strand in modern growth theory sees the Malthusian model as applying to the pre-modern or precapitalist era of low and stagnant productivity. See, for example, Galor (2011). Since our focus here is on the modern, capitalist era, we do not pursue the Malthusian conception further.

**17** Note that these sorts of responses of labour supply to higher real wages may result from business-supported policies designed to obtain cheaper labour (for example, weakened child labour laws or loosened immigration restrictions), or alternatively may result from workers' own responses to the incentives of higher wages (for example, increased immigration from countries or regions where wages are relatively lower). Also note that it is possible that the labour supply curve can become 'backward bending' and respond inversely to the real wage, for example if women enter the labour force more when the wages of male family members are reduced (see Blecker and Seguino, 2002, 2007). We do not address this possibility here, but it should be kept in mind as it may be important in some circumstances.

**18** The subscript '*LR*' is used here to indicate a long-run equilibrium level that may diverge from a short-run equilibrium level of the same variable. In such cases, '*SR*' is used to denote the short-run equilibrium value and the * is omitted to avoid notational clutter.

**19** These conclusions would have to be modified if firms responded to the higher wages by introducing labour-saving technical change that would reduce $a_0$, as discussed below.

**20** Recall that the short run in this model is essentially the same as the fixed real wage closure, since we assume that it takes time for the real wage to adjust to the long-run equilibrium level consistent with $g = n$.

**21** Generations of Marxian economists have debated the correct interpretation of what Marx originally meant by this concept. In some interpretations, it refers to the ratio of the dead labour embodied in the means of production (Marx's 'constant capital') to the labour value of the means of subsistence of the live labour (his 'variable capital'), while in other interpretations the denominator of the ratio is total live labour time (that is, value added, which is the sum of variable capital plus surplus value in Marx's terms). We adopt the phrase 'Marx-biased' from Foley and Michl (1999), who (aside from differences in notation) define it the same way as we do here, without invoking the labour theory of value.

**22** In fact, Marx argued that, more than this, firms would be driven to seek out and develop labour-saving techniques in a high-wage situation, even if this required higher capital costs for the new, often larger-scale machinery and equipment. We abstract from such endogenous innovation here and focus on the impact of a given new technique, once it is adopted, and we require only that it raises the profit rate at the initial, *ex ante* real wage.

**23** Mathematical solutions for the changes in long-run equilibrium $r$ and $g$ in all of these cases are left to the reader as an exercise.

**24** For other perspectives on Marx's FTRP theory, see the sources cited in note 1 to this chapter and Harris (1983). Wolff (2003) showed that technological change at the industry level in the US economy in the late 1980s and 1990s was mostly of a Marx-biased nature, but average capital productivity was rising rather than falling because of a shift in the composition of industries towards less capital-intensive sectors. For more recent evidence on the profit rate in the US economy, see Basu and Vasudevan (2013).

**25** See Foley and Michl (1999, Chapter 7) for an analysis of continuous Marx-biased technical change, and section 2.7 for the alternative view of an FTRP found in Ricardo's theory of the stationary state.

**26** If the real wage is initially below $\widetilde{w}$, a Marx-biased innovation would lower the profit rate, but for this very reason the innovation would be unviable and would not be adopted.

**27** In order for the equilibrium profit rate to rise, the innovation must take place in a 'basic' industry in the sense of Sraffa (1960): an industry that produces a good that is used directly or indirectly in the production of all other goods. In contrast, if the innovation takes place in a non-basic industry (for example, a luxury sector), it has no effect on the equilibrium profit rate in the basic industries or the economy as a whole.

**28** Roemer (1981, p. 132) states somewhat tendentiously that, 'If the rate of profit falls in such a changing real wage model, it is a consequence of the class struggle that follows technical innovation, not because of the innovation itself.' This statement can be logically reconciled with the interpretation adopted here if we think of a certain degree of worker bargaining power in the class struggle as being necessary to maintain a constant wage share.

**29** However, the profit rate does rise to $r'_{SR}$ in the short run following the introduction of a Marx-biased technical change, before falling back to $r_{LR}$ (see Figure 2.11). Therefore, as noted by Harris (1983, 1986), a series of such innovations could lead to *cyclical* increases and declines in the profit rate, but such a cyclical pattern would not constitute a long-term FTRP. However, Harris also notes that the *maximum* profit rate

($1/a_1$ in our notation) would be decreasing in the long run, as long as technical change continues to be Marx-biased. See also Foley et al. (2019, Chapter 8) on Marx-biased technical change.

**30** We reject the earlier interpretation of Ricardo's model of growth and distribution by Pasinetti (1960), which assumes that the real wage is fixed at the subsistence level at all times even in periods of rapid accumulation. As Casarosa forcefully shows through extensive textual references (a few of which are cited here), this formulation represents a serious misreading of Ricardo's views on wages.

**31** In Ricardo's (1821 [1951]) original terminology, these two wages rates would be the 'market' and 'natural' prices of labour, respectively.

**32** Using new lands or mines that require more capital and labour to produce the same amount of output is referred to as the 'extensive margin', while employing additional capital and labour on existing lands or mineral deposits with diminishing marginal productivity is referred to as the 'intensive margin'.

**33** In the classical view, capital is the fund that hires labour, so capital and labour are not really separate inputs but rather the first employs the second.

**34** In Ricardo's (1821 [1951], p. 67) original formulation, this process also involves two other elements that cannot be seen in our simplified presentation here. First, as the marginal product of labour in food production falls, the cost of labour (for any given real wage in terms of 'corn') rises, so that capitalists perceive rising wage costs even as the real wage for workers diminishes. Second, the falling marginal product in primary production also implies an increase in rents paid to landlords in exchange for the 'original and indestructible powers of the soil'. Thus, from the capitalists' point of view, they are squeezed by rising labour costs on the one hand and increasing rents on the other, until their profits fall to zero. However, since the profit rate is determined on the marginal unit of land (the parcel last brought into cultivation) on which no rent is paid, we can conduct our analysis for that unit in terms of wages and profits alone. In the stationary state, rents absorb the entire surplus over the necessary cost of labour (at the minimum subsistence wage) and profits go to zero.

**35** If $r_{min} > 0$ is allowed, then the stationary state would be reached earlier, in the sense that accumulation would cease as soon as profits fall to this minimal rate instead of to zero.

**36** It should also be noted that landlords' rents would absorb a rising share of total output in the primary sector as an economy declined into the stationary state, since Ricardian 'differential rent' equals the difference between the productivity of each production unit and the productivity of the last or 'marginal' unit brought into production (in agriculture and mining). We don't see the rising proportion of rents in our depiction of this model because we focus on the marginal production unit.

**37** This is in the long run; in the short run, there may be a temporary bulge in population growth if mortality (death) rates fall before fertility (birth) rates come down. This phenomenon, which is known as the 'demographic transition', has been experienced in many countries historically.

**38** For extended Ricardian trade models that encompass uneven development between industrialized and developing countries, see Maneschi (1983) and Burgstaller (1985).

**39** Our presentation will follow the mathematics used by Harvie (2000) and Grasselli and Maheshwari (2017). See also Taylor (2004).

**40** Note that the Goodwin model thus incorporates pure labour-saving (Harrod-neutral) technological change, but not the labour-saving and capital-using (Marx-biased) variety.

**41** Mathematically, this is a special case of a system of Lotka–Volterra cycles, named after the 'predator–prey' model of mathematicians Alfred Lotka and Vito Volterra, in which the wage share is the predator and the employment rate is the prey (since the wage share is increasing in the employment rate, whereas the employment rate is decreasing in the wage share). The cycles have a periodicity of

$$T = \frac{2\pi}{((q+\gamma)\,[(1/a_1)-(n+q)])^{1/2}}$$

time units $T$, where $\pi$ is the mathematical expression pi (approximately 3.14159) and not the profit share as used elsewhere in this book.

**42** The theme of cyclical behaviour will be continued in Chapters 6 and 7, which will cover alternative cycle models that may exhibit Goodwin-like properties in spite of being driven by forces other than labour costs and a profit-squeeze.

# 3 Neo-Keynesian models

## 3.1 Introduction

John Maynard Keynes published his *General Theory of Employment, Interest and Money* during the depths of the Great Depression in 1936. His book revolutionized economic thinking and led to the foundation of the separate branch of macroeconomics within the economics profession. Most importantly, Keynes showed how aggregate demand could determine output and employment, and how a collapse of aggregate demand could lead to a sustained equilibrium with involuntary unemployment of the type observed in the depression. Later generations of post-Keynesian economists sought to extend Keynes's thinking about the role of demand in the macroeconomy to the analysis of long-run growth. This chapter will cover some of the early efforts to incorporate demand into models of long-run growth that focused on domestic sources of demand, principally investment; models that focus more on consumption (especially workers' consumption) are considered in Chapters 4 and 7, while models that focus on external sources of demand (exports) will be covered in Chapters 8–10.

In the classical-Marxian models of the previous chapter, economic growth was constrained by the supply of saving, which in turn was fundamentally determined by the technology of production, the distribution of income and the propensity of the business owners (capitalists) to accumulate (save and invest) their profits in the form of new capital. Those models did not allow any role for demand-side factors to influence the path of long-run growth, because they did not distinguish between saving (that is, refraining from consuming out of current income, or accumulating financial assets) and investment (that is, purchasing stocks of newly produced capital goods, or accumulating productive assets). Thus, they effectively assumed Say's law, which implies that all savings are invested (in the sense of being spent on newly produced capital goods), so that diverting income from consumption to saving does not result in any diminution of aggregate demand.

By the time Keynes wrote the *General Theory*, the distinction between saving and investment (and the concept of financial intermediation between the two) was better appreciated. However, neoclassical theorists such as Wicksell (1898 [1936]) and Pigou (1933) had portrayed the interest rate as the sole variable equilibrating between saving and investment, with no impact on the level of economic activity (output or employment). Keynes's great analytical insight was to recognize that national income (and with it the employment required to produce the corresponding output) could be the adjusting variable, so that a depressed desire to invest by firms would translate into reduced output and lost jobs.[1] It was this contribution of Keynes that finally showed the fallacy in Say's law, although Malthus and Marx had tried (with only partial success) to refute it earlier.

Furthermore, Keynes reversed the direction of causality between saving and investment from the neoclassical models: instead of investment being constrained by available saving (as was implicitly assumed by the classicals as well as the neoclassicals), in Keynes's view the equilibrium amount of saving would adjust to the investment spending of firms through variations in the income flows generated by the multiplier impact of the investment, which in turn would provide the corresponding saving. This insight led to Keynes's famous 'paradox of thrift', in which he showed that an effort by society to save a higher proportion of its income would fail to boost either investment or output, and could in fact diminish both and worsen unemployment because of the associated loss of consumption demand. One of Keynes's colleagues at the University of Cambridge (who later moved to Oxford), Sir Roy Harrod, was a pioneer in the development of an analysis of long-run growth in which aggregate demand mattered and a long-run paradox of thrift could be observed (Harrod, 1939).[2]

Keynes and Harrod, however, had relatively little to say about the distribution of income, which had been a major focus of the classicals and Marx. Keynes did observe that the 'marginal propensity to consume' was likely to be inversely related to household income levels, but he did not systematically incorporate income distribution into his analysis. In the decades following the publication of the *General Theory*, economists working in the post-Keynesian tradition were influenced by two sources who (in different ways) revived the classical-Marxian focus on income distribution. On the one hand, Italian economist Piero Sraffa – who moved to Cambridge in the late 1920s to escape from Italian fascism – was editing the collected works of David Ricardo and working on a neo-Ricardian approach to the theory of value and distribution.[3] Sraffa (1960) showed how the classical inverse relationship between profits and wages emerges from a multisectoral analysis using input–output analysis. On the other hand, Polish economist Michał Kalecki – who moved to the Oxford Institute of Statistics to escape the Nazi occupation of Poland during World War II – also interacted with Keynes's circle at Cambridge. Kalecki's modelling approach (Kalecki, 1990, 1991), which will be covered in more depth in the next chapter, incorporated different saving propensities out of wage and profit income (with all savings coming out of profits in his simpler models) and made the relative shares of wages and profits in national income a key determinant of aggregate economic behaviour. As a result of these influences, when Keynes's followers (after Harrod) began to apply his new paradigm to long-run growth, many of the resulting 'neo-Keynesian' models (especially those of Kaldor, 1955–56; Pasinetti, 1962; Robinson, 1956, 1962) placed the functional (wage–profit) distribution of income and differential saving propensities out of wages and profits at the centre of the analysis. This neo-Keynesian approach (sometimes also referred to as the 'Cambridge school') is one of the chief foundations for all the post-Keynesian approaches to growth and distribution covered throughout the remainder of this book.

The rest of this chapter is organized as follows. Section 3.2 develops Harrod's foundational model of unstable growth. Section 3.3 discusses Kaldor's (1955–56) model of growth and income distribution, along with providing a brief account of the critique by Pasinetti (1962) and response of Kaldor (1966b); more details are given in Appendix 3.1. Section 3.4 then presents Robinson's (1956, 1962) pioneering model of long-run growth with an independent investment function; the text presents a 'neo-Robinsonian' version, while Robinson's original formulation is discussed in Appendix 3.2. Section 3.5 covers Marglin's (1984a, 1984b) effort to synthesize neo-Marxian and neo-Keynesian growth models, which can be seen as formalizing Robinson's idea of real wage resistance or an inflation barrier. Section 3.6 concludes by connecting Robinson's growth theory to other post-Keynesian approaches.

## 3.2 Harrod's model of unstable growth

Theoretical work on long-run growth took a long hiatus during the period between Marx's analysis of the dynamics of capitalism in the 1860s and the Great Depression of the 1930s. During this period, the mainstream economics of the time (the early neoclassical school) was focused primarily on static models of efficiency in resource allocation, although neoclassical work on capital theory and intertemporal choice (for example, Ramsey, 1928; Fisher, 1930) laid the foundations for later neoclassical growth models. In the neo-Keynesian tradition, Harrod (1939) was the first to extrapolate Keynes's short-run framework to the analysis of long-run growth; his work laid the foundations for large literatures among both mainstream and heterodox economists for the next several decades. Moreover, Harrod invented several key concepts that are still cornerstones of growth theory today, so we begin our discussion here with an account of his original approach. The contemporary neo-Harrodian approach to macrodynamics will be covered in Chapter 6.

### 3.2.1 Three growth rates

Rather than focusing on a single (equilibrium) growth rate, Harrod's macrodynamics emerge from the interaction of three distinct growth rates: the *actual* rate of growth, which, as its name suggests, is the rate of growth actually observed in the economy; the *warranted* rate of growth, which is the rate of growth consistent with equilibrium between investment and savings; and the *natural* rate of growth which, as noted in Chapter 1, ensures either full employment of labour or a constant employment rate.[4] Recall that the natural rate of growth is essentially an upper limit on the growth rate – more precisely, the maximum rate of growth that the economy can achieve in the long run, given the availability (and productivity) of labour resources. As previously demonstrated in Chapter 1, the natural rate of growth can be written as

$$y_N = q + n \tag{3.1}$$

where, as before, $q$ is the rate of growth of labour productivity and $n$ is the rate of growth of the labour force or (assuming a constant labour force participation rate in the long run) the rate of growth of the population.

In order to derive the value of the warranted rate of growth, consider the following investment function:

$$I_t = a_1(Y_t^e - Y_{t-1}) \tag{3.2}$$

Equation (3.2) describes an accelerator relationship relating investment spending in the current period to the expected expansion of output (the difference between expected output in the current period, $Y_t^e$ and the actual level of output in the previous period). Note that in equation (3.2) we assume, for simplicity, that there is no depreciation. As such, all new investment constitutes a net addition to the capital stock, since no investment is required simply to replace capital that is lost to depreciation. On this assumption, the accelerator coefficient linking the expected expansion of output to planned investment spending in equation (3.2) is equal to $a_1$, the value of the full-capacity capital–output ratio. As determined by the current state of technology, $a_1$ describes the quantity of capital required to produce any given level of output and, at the margin, the additional quantity of capital needed to produce any additional output.

The use of $a_1$ as the accelerator coefficient in (3.2) is no accident. To see why, first note that from equation (3.2), we have

$$I = \Delta K = a_1 \Delta Y^e \tag{3.3}$$

where $\Delta$ means the change in a variable in discrete time. Now recall from Chapter 1 (equation 1.11) that $K_u = uK$ represents the amount of the available capital stock, $K$, that is utilized to produce the actual level of output, $Y$, at any point in time (where $u = Y/Y_K$ is the utilization rate), and that based on this definition,

$$\frac{K_u}{Y} = a_1$$

It follows that

$$K_u = a_1 Y$$

$$\Rightarrow \Delta K_u = a_1 \Delta Y \tag{3.4}$$

Whereas equation (3.3) describes the expansion of the capital stock available for production (based on firms' investment behaviour stemming from the accelerator relationship), equation (3.4) describes expansion of the amount of capital actually used in the production process, based on the realized (or actual) expansion of the economy.

Now consider a situation where $\Delta Y = \Delta Y^e$ so that expectations are realized. Under these conditions, we can equate the right-hand sides of equations (3.3) and (3.4), which reveals that:

$$I = \Delta K = a_1 \Delta Y = \Delta K_u$$

In other words, the investment undertaken by firms creates exactly the amount of additional capital capacity ($\Delta K$) that is needed to meet the capital requirements of the actual expansion of output ($\Delta K_u$). This is *precisely* the behavioural rationale for the accelerator relationship, according to which firms are attempting to expand productive capacity in order to keep pace with the expansion of the economy (given the current state of technology).

In short, the accelerator coefficient is not – and should not be – any arbitrary constant, but instead conforms to the known capital requirements of the production process given by the value of $a_1$.

Returning now to our quest to identify the value of the warranted rate of growth, suppose we introduce a simple proportional savings function of the form

$$S_t = sY_t \tag{3.5}$$

Equation (3.5) assumes that all households have the same marginal propensity to save, $s$, with $0 < s < 1$. Now consider both equations (3.2) and (3.5), describing investment by firms and savings by households, respectively. Conditions of equilibrium require that

$$S_t = I_t$$

(equalization of savings and investment), and

$$Y_t = Y_t^e$$

(realization of expectations). Imposing these equilibrium conditions on equations (3.2) and (3.3) and combining the resulting expressions, we get

$$sY_t = a_1(Y_t - Y_{t-1})$$

$$\Rightarrow \frac{s}{a_1} = \frac{Y_t - Y_{t-1}}{Y_t}$$

The term on the right-hand side of this expression describes nothing more than the proportional rate of expansion of actual output – that is, the actual rate of growth of output, $y$. What the expression then tells us is that under conditions of *equilibrium*, the actual rate of growth takes on the particular value that appears on the left-hand side of this equation. In other words, thanks to the analysis above, we have now succeeded in uncovering the value of the equilibrium or (to use Harrod's parlance) *warranted* rate of growth, $y_w$. In accordance with this result, we can therefore state

$$y_w = \frac{s}{a_1} \tag{3.6}$$

### 3.2.2 The first and second Harrod problems

Having identified the actual, natural, and warranted rates of growth, we are now in a position to study the growth outcomes that, according to Harrod, are characteristic of a capitalist economy. This is done by considering the interaction of the actual, natural, and warranted rates of growth, which interaction is summarized in the form of two famous problems.

We have already encountered the first Harrod problem in Chapter 1, as a general property of heterodox growth models. To see the nature of this problem in the current (Harrodian) context, consider the equation:

$$y = y_w = y_N$$

which, appealing to (3.1) and (3.6) (and taking both the rate of growth of productivity and the rate of growth of the population as given), yields:

$$y = \frac{s}{a_1} = \overline{q} + \overline{n} \tag{3.7}$$

The first equality in (3.7) states that the actual rate of growth equals the warranted rate, while the second states that the warranted rate of growth equals the natural rate. Together, these constitute the conditions necessary for the observation of equilibrium growth consistent with the natural rate. The problem that is drawn to attention by the expression in (3.7) is that the determinants of the warranted and natural rates of growth are independent of one another. In other words, equilibrium growth consistent with the natural rate is *possible* (nothing rules out the equality in equation 3.7) but *not likely*: there is no obvious mechanism that would coordinate household saving decisions, the state of technology (and its rate of progress) and the growth rate of the population so as to bring about the equality of $y_w$ and $y_N$. The importance of this result lies in its suggestion that a capitalist economy is unlikely to experience growth consistent with a constant rate of unemployment, much less growth consistent with *full* employment, which, in addition to the equality in (3.7), would also require initial equality of the actual and potential levels of output.

To understand the second Harrod problem, consider again the saving equation (3.5). If we assume that $S_t = I_t$, then substituting into (3.5) and solving for $Y_t$ yields:

$$Y_t = \frac{1}{s} I_t \tag{3.8}$$

Equation (3.8) is a simple Keynesian multiplier relationship between the level of output and the level of investment spending. Together with equation (3.2), this gives us a simple multiplier–accelerator model of the sort first contemplated by Harrod (1936, p. 70).[5]

Substituting equation (3.2) into (3.8) gives us

$$Y_t = \frac{a_1}{s}(Y_t^e - Y_{t-1})$$
$$\Rightarrow \frac{Y_t}{Y_t^e} = \frac{a_1}{s} \cdot \frac{(Y_t^e - Y_{t-1})}{Y_t^e} \qquad (3.9)$$

Now define the expected rate of growth, $y^e$, as

$$y^e = \frac{Y_t^e - Y_{t-1}}{Y_t^e}$$

and recall from equation (3.6) that $y_w = s/a_1$. Substituting into (3.9) yields

$$\frac{Y_t}{Y_t^e} = \frac{y^e}{y_w} \qquad (3.10)$$

Rather than furnishing an expression for 'the' rate of growth, equation (3.10) enables us to explore the interaction between the actual, expected and warranted rates of growth and, in the process, demonstrate the substance of the second Harrod problem.

We are already aware that the warranted rate of growth is an equilibrium consistent with (among other things) realized expectations. Note, then, that if $y^e = y_w$ initially, it follows from (3.10) that

$$\frac{Y_t}{Y_t^e} = 1$$
$$\Rightarrow \frac{1}{Y_t^e} = \frac{1}{Y_t}$$
$$\Rightarrow -\frac{Y_{t-1}}{Y_t^e} = -\frac{Y_{t-1}}{Y_t}$$
$$\Rightarrow 1 - \frac{Y_{t-1}}{Y_t^e} = 1 - \frac{Y_{t-1}}{Y_t}$$
$$\Rightarrow y^e = y$$

In other words, expectations consistent with the warranted rate $(y^e = y_w)$ will bring about conditions of equilibrium $(y = y^e = y_w)$. But suppose that expectations are *not* consistent with the warranted rate initially. For example, assume that $y^e > y_w \Rightarrow y^e/y_w > 1$. Then it follows from (3.10) that

$$\frac{Y_t}{Y_t^e} > 1$$

$$\Rightarrow \frac{1}{Y_t^e} > \frac{1}{Y_t}$$

$$\Rightarrow -\frac{Y_{t-1}}{Y_t^e} < -\frac{Y_{t-1}}{Y_t}$$

$$\Rightarrow 1 - \frac{Y_{t-1}}{Y_t^e} < 1 - \frac{Y_{t-1}}{Y_t}$$

$$\Rightarrow y^e < y$$

We now have $y > y^e > y_w$, describing conditions of disequilibrium where the actual rate of growth differs from the expected rate, which differs again from the warranted rate. The critical question is, what happens next? Are there forces at work that will correct the disequilibrium by pushing the actual rate towards the warranted rate, or will the actual rate diverge further away from this equilibrium value?

Suppose that, in the first instance, firms respond to the expectational disappointment in accordance with the following general behavioural principle (Fazzari, 1985, p. 77)

$$y_t > y_t^e \Rightarrow y_{t+1}^e > y_t^e$$

$$y_t < y_t^e \Rightarrow y_{t+1}^e < y_t^e$$

This suggests that growth expectations are adjusted upward if the actual rate of growth exceeds expectations in any given period, and vice versa. The behavioural principle at work here is satisfied by any form of adaptive expectations mechanism, where expectations are revised in response to realized events in the past. For example, if firms set their expectations as

$$y_t^e = y_{t-1}$$

then the behavioural principle outlined above will be satisfied.

Given that we are concerned with specific conditions where $y > y^e$, what all this suggests is that firms will now revise their growth expectations upward.

Note, however, that we *began* with $y^e > y_w$. It appears that the development we are now contemplating – an increase in $y^e$ – will *exacerbate* rather than resolve the conditions of disequilibrium we have uncovered. To see that this is, in fact, the case, note that since by definition

$$y = \frac{Y_t - Y_{t-1}}{Y_t}$$

we can write:

$$y = \frac{Y_t - Y_{t-1}}{Y_t} \cdot \frac{Y_t^e}{Y_t^e}$$

$$\Rightarrow y = \frac{Y_t^e - Y_{t-1} + Y_t - Y_t^e}{Y_t^e} \cdot \frac{Y_t^e}{Y_t}$$

which, by appealing to the result in equation (3.10), can be written as:[6]

$$y = \left(y^e + \frac{y^e}{y_w} - 1\right) \cdot \frac{y_w}{y^e}$$

$$\Rightarrow y = 1 - \frac{y_w}{y^e} + y_w \qquad (3.11)$$

It is clear from (3.11) that the actual rate of growth is increasing in the expected rate, so that an increase in $y^e$ in response to $y > y^e$ will produce an increase in $y$, pushing the actual rate of growth further away from the warranted rate. To see this formally, note that it follows from (3.11) that:

$$\frac{dy}{dy^e} = \frac{y_w}{(y^e)^2} > 0$$

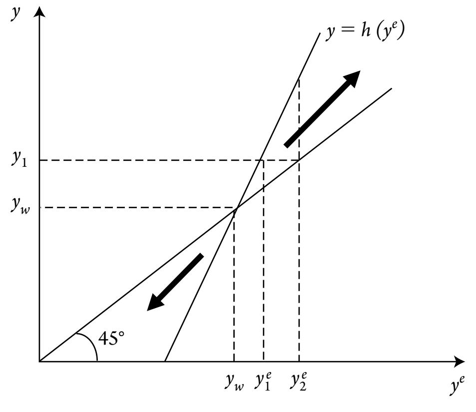
This is the second Harrod problem: the warranted rate of growth – the equilibrium of the system – is unstable. Any initial divergence of the actual rate of growth from the warranted rate will be self-reinforcing rather than self-correcting, producing still greater divergence subsequently.[7]

Since $y_t^e = y_{t-1}$ by hypothesis, so that $dy_t^e/dy_{t-1} = 1$, strictly speaking we require $dy_t/dy_t^e \geq 1$ in order to observe continual self-reinforcing increases in $y$ and $y^e$ following $y > y^e > y_w$ initially (see also Fazzari et al., 2013, p. 7). In other words, in addition to the hypotheses entertained thus far, a necessary condition must be satisfied in order for Harrodian instability to materialize. Referring again to equation (3.11), we can see that

$$\frac{dy}{dy^e} = \frac{y_w}{(y^e)^2} > 1$$

if

**Figure 3.1** Instability of the warranted rate of growth




$$y_w > (y^e)^2$$

Since $y^e$ is a proportional rate of growth (so that $y^e < 1$), then even with $y^e > y_w$ it is possible for this condition to be satisfied locally, so that $y_w$ will, indeed, display the instability ascribed to it by Harrod. Figure 3.1, where the relationship $y = h\ (y^e)$ has been linearized for the sake of simplicity, illustrates these principles. Starting with $y_1^e > y_w$, we obtain $y_1 > y_1^e > y_w$. Expectations then adjust to $y_2^e$, as a result of which we get $y_2 > y_2^e$, and so on – the actual and expected rates of growth continually increasing in a self-reinforcing fashion as indicated by the upward-pointing bold arrow in Figure 3.1.

Some critiques of Harrod's macrodynamics focus on the instability of the warranted rate as a weakness of the analysis. The basis of this claim is the commonplace observation that there is no obvious tendency for growth rates in capitalist economies to systematically increase or decrease over time, least of all in the long run. But as Kregel (1980) argues, the model developed so far to demonstrate the potential emergence of the two Harrod problems is designed only to capture an instantaneous rate of growth: a full model explaining growth in the long term would require additional adjustment mechanisms. Were this full model to be consistent with Harrod's original vision, it would likely produce something resembling a cyclical expansion path about the warranted rate of growth (see also Fazzari and Greenberg, 2015). Indeed, interpretations of Harrod's dynamics along these lines led to a whole generation of business cycle models based on interactions between the investment accelerator and the Keynesian output multiplier with various sorts of ceilings and floors in the 1940s and 1950s.[8] These contributions are now understood as the progenitors of the modern neo-Harrodian analysis discussed in Chapter 6.

### 3.2.3 The importance of expectations for the second Harrod problem

The result just derived demonstrates that the multiplier–accelerator model as originally specified in equations (3.2) and (3.8) contains the *seeds* of instability, but this instability only *materializes* if expectations respond in a particular way to the experience of disequilibrium. Hence if $y^e > y_w$ results in $y > y^e$ as described above, but if this outcome, in turn, provokes no response on the part of expectations, the divergence of $y$ from the warranted rate of growth will not occur. As noted by Fazzari (1985, p. 78), 'if firms used the most rigid of expectational rules – i.e., pick an expected growth rate and never change it – the explosion [Harrodian instability] would not occur'. At first this behaviour may seem unreasonable: why *wouldn't* expectations change in response to their observed falsification? On second thought, however, if the decision-making environment is one of uncertainty where behaviour is guided (to a substantial extent) by conventions, and if one such norm is to regard some magnitude of expectational error as 'normal', then the disparity between $y$, $y^e$ and $y_w$ may be insufficient to provoke a reaction by firms.[9] In this case Harrodian instability will not materialize: 'the explosion [Harrodian instability] is created by the incorporation of new information into expectational structures' (Fazzari, 1985, p. 78); absent such a revision of expectations, there will be no tendency for the economy to diverge from the warranted rate.

The importance of expectations for the dynamics associated with Harrodian instability is further demonstrated by the fact that these dynamics will fail to materialize in the presence of even relatively weak forms of rational expectations, such as myopic perfect foresight. Myopic perfect foresight is sometimes used in heterodox macroeconomics as a means of entertaining the possibility of model-consistent expectations without going to the extremes of omniscience regularly postulated by advocates of the rational expectations hypothesis in mainstream macroeconomics (Flaschel et al., 1997). The distinction between mainstream dynamic optimizers and decision makers with myopic perfect foresight can be illustrated by the following example. Consider a group of people who find themselves in the cluttered basement of a friend's house when the lights suddenly go out. As mainstream dynamic optimizers, the group would find, to their delight, that they are in possession of an arc lamp. Turning on this arc lamp would illuminate the entire basement, allowing the group to see clearly the path that will lead them from their current position to the basement door, even before they take a single step. Seeing their destination and the path they need to take towards it in advance, they can set off with surety and – barring unforeseeable mishaps (such as a startled reaction to the appearance of a mouse) – make their way out of the basement none the worse off.

Decision makers with myopic perfect foresight, however, would find themselves equipped only with an old and rather feeble torch (or 'flashlight'). Emitting only a dim beam of light, this would allow them to see one step ahead so as to put one foot in front of the other without banging their shins against any of the discarded ephemera cluttering the darkened basement. But most of the terrain that surrounds them would remain shrouded in inky darkness, and the group would have no inkling as to either the exact location of the basement door, or a suitable path they might take towards it. Although sure of their next footfall thanks to the torch, as the group moves forward by putting one foot in front of the other, they would remain unsure as to exactly where their journey is taking them.

As this (admittedly fanciful) tale demonstrates, myopic perfect foresight allows heterodox model builders to postulate model-consistent expectations without indulging the worst extremes of the rational expectations hypothesis, while simultaneously remaining faithful to the heterodox notion that the economy is a historical construct that 'makes itself up as it goes along' (as opposed to advancing in accordance with a predestined dynamic programme, thwarted only by unforeseen and unforeseeable exogenous shocks). It turns out, however, that even myopic perfect foresight is destructive of the second Harrod problem. To see this, recall that the basis of our analysis is the multiplier–accelerator model in equations (3.2) and (3.8), combination of which equations yields:

$$Y_t = \frac{a_1}{s}(Y_t^e - Y_{t-1})$$

If we now assume myopic perfect foresight, so that $Y_t^e = Y_t$, and rearrange the expression above, we get

$$\frac{Y_t - Y_{t-1}}{Y_t} = \frac{s}{a_1}$$
$$\Rightarrow y = y_w$$

This tells us that with the exercise of myopic perfect foresight (and in the absence of exogenous shocks), the economy never strays from its warranted (equilibrium) path: the disequilibrium conditions that sow the seeds of Harrodian instability never materialize.[10] On reflection, this is but another demonstration of the basic principle elucidated by Fazzari (1985, p. 78), as cited earlier, that Harrodian instability 'is created by the incorporation of new information into expectational structures', so that absent such revision there will be no divergence from the warranted rate. In this case, there is no failure on the part of firms to incorporate new information into their expectations: expectations are correct, and so, absent their falsification (and holding all else constant), there is no 'new information' to incorporate.

There are possible caveats to this result. As noted by Asimakopulos (1991, pp. 156–8), in an environment of uncertainty where behaviour is affected not just by decision makers' most probable forecasts of future events but also, given the self-acknowledged incomplete informational basis of this forecast, by their animal spirits, the mere realization of expectations in one period is *not* necessarily sufficient to ensure that a forecast will remain unchanged over time. The crux of the argument has to do with the one sector, 'representative agent' representation of the economy in the model above, which abstracts from the commonplace observation that the economy is, in fact, made up of a multiplicity of heterogeneous agents. For Asimakopulos, this heterogeneity extends to expectations formation, so that the 'one size fits all' assumption of myopic perfect foresight (and the assumed common investment response to these identical expectations) is inappropriate. Instead:

> [aggregate] investment decisions are made by a very large number of entrepreneurs producing different types of goods, and in a world characterized by uncertainty it would be a fluke if these individual decisions added up to the investment corresponding to that required by a warranted growth. *It would be even more remarkable if they continued to do so*. (Asimakopulos, 1991, p. 156; emphasis added)

The failure of expectations (and hence the actual rate of growth) to remain unchanged even if the warranted rate of growth is achieved might also result from cumulative experience of the warranted rate itself. Assume once again that the economy is made up of a multiplicity of heterogeneous firms, and assume further that the form of competition between these firms is of the strategic type associated by Hall and Hitch (1939 [1988]) with conditions of oligopoly, or what Shaikh (2016, p. 14) calls 'real competition'. These are not the anonymous, quasi-cooperative conditions of perfect (or even imperfect) competition in neoclassical theory, but are instead analogous to the conditions of combat, where known adversaries interact with one another, understanding that their competitive actions will elicit reactions from their adversaries (without their being able to perfectly anticipate these reactions thanks to conditions of uncertainty), and where part of the process of competition involves striving to get ahead by inflicting damage on (or preferably eliminating) one's opponents.[11] In a competitive environment of this sort, the cumulative experience of any one decision maker of the same, tranquil equilibrium conditions may excite the feeling that things are 'too quiet', and so elicit behavioural change designed to 'stay one step ahead' of adversaries – even if a sufficiently large information set, containing the intentions of all other firms, would reveal that no such behavioural response is required. There would then be a (seemingly spontaneous) change in expectations in response to nothing more than the *cumulative realization of expectations in the recent past*, in much the same way that experience of the same outcome may induce boredom and thus provoke behavioural change. This behaviour at the microeconomic level would create conditions of disequilibrium in the aggregate, and so give rise to the potential onset of Harrodian instability.

Note, though, that once dislodged from the warranted rate, expectations must continue to change in order to propel the economy further from its equilibrium. Hence even if the basis for maintaining constant expectations (and thus avoiding divergence from the warranted rate altogether) is contestable, the fact remains that absent continual expectational change (and an appropriate behavioural explanation of why this *does* occur), continual divergence from the warranted rate will not be observed.

In short, Harrodian macrodynamics are sensitive to both the formulation and reformulation of expectations. Expectations must be formulated in an essentially adaptive manner, reacting to past realized outcomes (including, possibly, cumulative experience of the same outcomes). Furthermore, expectations must be actively reformulated in response to their disappointment in order for the instability that is latent in conditions of disequilibrium to become manifest. While it is important to be aware of these issues, it should be noted that the assumptions we must therefore make about the formulation and reformulation of expectations in order for the second Harrod problem to materialize are behaviourally plausible in the context of a Keynesian decision-making environment of fundamental uncertainty. Indeed, a stronger argument that most contemporary post-Keynesians would entertain is that these assumptions are not just plausible but highly likely to obtain. In other words, even the relatively weak variant of model-consistent expectations that is myopic perfect foresight is likely well beyond the capacities for foresight of the great majority of decision makers.

### 3.2.4 Another look at Harrodian instability

Suppose we return to the conditions shown above to give rise to the possibility of Harrodian instability, where $y^e > y_w \Rightarrow y^e/y_w > 1$. It follows from equation (3.10) that

$$\frac{Y_t}{Y_t^e} > 1$$
$$\Rightarrow Y_t > Y_t^e$$
$$\Rightarrow Y_t - Y_{t-1} > Y_t^e - Y_{t-1}$$
$$\Rightarrow a_1(Y_t - Y_{t-1}) > a_1(Y_t^e - Y_{t-1}) \quad (3.12)$$

Now recall that, as previously demonstrated,

$$\Delta K = a_1 \Delta Y^e = a_1(Y_t^e - Y_{t-1}) \quad (3.3)$$

and

$$\Delta K_u = a_1 \Delta Y = a_1(Y_t - Y_{t-1}) \quad (3.4)$$

Using (3.3) and (3.4) in conjunction with the inequality in (3.12), it follows that

$$\Delta K_u > \Delta K \quad (3.13)$$

What this inequality tells us is that in disequilibrium – specifically, conditions where $y^e > y_w$ so that $y > y^e$ – firms are creating (through their investment spending in equation 3.3) too little new capital ($\Delta K$ on the right-hand side of 3.13), given the rate at which the economy is expanding and the capital requirements of this expansion ($\Delta K_u$ on the left-hand side of 3.13). At first this claim may seem absurd. How can the economy expand faster than the rate facilitated by the new productive capacity that firms are creating? The answer, of course, is through an increase in the rate of capacity utilization, $u$. Suppose that, prior to the onset of the disequilibrium conditions $y^e > y_w$, the economy was operating at its normal rate of capacity utilization, $u_n$.[12] The increase in capacity utilization necessary to accommodate the onset of Harrodian instability would therefore result in a situation where $u > u_n$. This observation is valuable for two reasons.

The first is technical. Recasting the workings of the second Harrod problem to draw to attention its implications for the proximity of the actual and normal rates of capacity utilization is important because as mentioned in Chapter 1 (and as will be discussed in more detail in Chapter 6), the relationship between these variables – specifically, the potential for the actual rate of utilization to depart from its normal rate in the long run – is an important source of debate in heterodox growth theory *writ large*. Relating the second Harrod problem to the proximity of the actual and normal rates of capacity utilization therefore helps situate the phenomenon of Harrodian instability within the broader corpus of heterodox growth theory.

Second, the analysis above gives us additional insight into the processes at work when the economy experiences Harrodian instability. So far, we have seen that the formulation and reformulation of expectations are essential to the dynamics of Harrodian stability. What the inequality (3.13) now draws to attention is another aspect of the behavioural foundations of the second Harrod problem. In (3.13), the experience of disequilibrium conditions sends a signal to firms to increase their investment spending in order to correct a supply-side imbalance – namely, that they are creating too little additional productive capacity, as a result of which they are ultimately forced to increase the utilization rate of installed capacity over and above its preferred normal rate. But because the environment in which firms operate is one in which both the level and rate of growth of activity are demand-led, investment spending plays a dual role. It is both a source of new productive capacity on the supply side *and* a source of demand for goods (and hence a stimulus to aggregate demand) on the demand side – where its effects are amplified by the operation of the expenditure multiplier. The result is that as firms attempt to adjust their capital stocks in order to make up for the shortfall of capacity signalled by equation (3.13), the macroeconomic consequence of their increased investment spending is a further increase in aggregate demand that exacerbates the original problem (of excessive capacity utilization). This will provoke a further attempt to adjust capacity by raising investment spending, and so on.

In a nutshell, the second Harrod problem is now seen to work through the dual role of investment, as both a source of productive capacity and a source of aggregate demand. The essence of the problem is that in initial conditions of disequilibrium where, for example, demand is excessive relative to available capacity, resulting in the outcome in inequality (3.13), the efforts of firms to adjust by changing their investment plans will have a greater effect on the demand side than the supply side, thereby worsening the original problem. Disequilibrium conditions are once again revealed as self-reinforcing rather than self-correcting.

### 3.2.5 Interaction of the first and second Harrod problems and the properties of Harrodian growth outcomes

Having studied the first and second Harrod problems individually, we are now in a position to consider their *interaction*. Now – and this is where Harrod's analysis becomes difficult and confusing – once we combine the instability of the warranted growth rate with the likely discrepancies between that rate and the natural growth rate, some of the results are quite counter-intuitive. Start with the case in which $y_w < y_N$, so that *if* the economy grew at the warranted rate it would face perpetually increasing unemployment of labour. However, given the instability of the warranted rate, the economy would *not* likely grow at $y_w$, and in fact Harrod considered this situation to be one in which the economy would be prone to boom-bust behaviour: small increases in actual growth above the warranted rate would lead to economic booms (episodes of ever-rising actual growth); such booms would be limited by the fact that the economy would eventually grow too fast relative to the natural rate, at which point the labour force would be exhausted and inflation would likely result from upward pressure on wages and prices. Thus, this situation might result in chronic 'overheating' of the economy with rapid growth regularly pushing up against labour supply limitations resulting in recurrent inflationary pressures.

In contrast, suppose instead that $y_w > y_N$. Of course, in this situation the economy cannot actually grow at $y_w$ in the long run, because it would eventually hit upon a labour supply constraint. But in Harrod's scheme, the likely outcome in this scenario is in fact a chronically depressed economy, rather than a booming or inflationary one. To see this, note that the slow growth of the labour force implies that the economy would normally grow more slowly than the warranted rate in this situation (since in the long run we must have $y \leq y_N < y_w$). Given the instability of the warranted rate, actual growth would tend to fall further and further below its warranted (*and* natural) rate, until the economy hit upon some kind of floor (for example, where the need to replace depreciated capital or government stimulus policies started to kick in).

This analysis reveals why Harrod's equation for warranted growth should not be applied mechanically for policy analysis, as it is sometimes in the emphasis on 'mobilizing domestic saving' in many economic development textbooks (for example, Todaro and Smith, 2012). It might be thought from inspection of $y_w = s/a_1$ that policies to increase the saving propensity $s$ would be helpful for boosting long-run growth, but in Harrod's view an increase in $s$ would only increase the *warranted* growth rate – *not* necessarily the *actual* growth rate.

Indeed, such policies would likely be counterproductive and backfire, due to the instability of the warranted rate. Even if $y = y_w$ initially (which of course could only happen by accident, per the second Harrod problem), any increase in $y_w$ not accompanied by a sufficient boost to aggregate demand would immediately create a situation where $y < y_w$, thereby inducing a downward spiral of falling actual growth due to the instability of $y_w$. Furthermore, if a rise in $s$ increased $y_w$ to the point where $y_w > y_N$, the economy would also become chronically depressed for the reason explained above.

This result is the Harrodian equivalent of Keynes's famous 'paradox of thrift': increasing a society's propensity to save would, paradoxically, lower (rather than raise) the actual rate of growth. Fundamentally, the paradox results from the fact that increased saving means reduced consumption demand, and in Harrod's analysis there is no guarantee that investment demand would rise to offset this and prevent aggregate demand from falling overall. Indeed, just the opposite is likely to occur: investment is likely to fall as a result of accelerator effects from falling actual output growth.[13]

### 3.2.6 Are the Harrod problems more apparent than real?

It is commonplace in mainstream accounts of growth theory to be dismissive of the two Harrod problems. The first Harrod problem is said to be created by the assumed constancy of the full-capacity capital–output ratio, $a_1$, which renders the warranted rate constant and hence unlikely to take the exact value of the (independently determined) natural rate of growth. According to this view, the first Harrod problem was effectively solved by Solow's (1956) neoclassical growth theory based on a continuous production function, along which the capital–output ratio is variable, as capital and labour are substituted for one another in response to changes in relative factor prices.[14] As discussed in Chapter 1, this allows $y_w$ to adjust towards $y_N$ through variations in $a_1$. Much can be said about the different theories of capital and production we studied in Chapter 1, and their implications for the treatment of $a_1$, in response to this view.[15]

The empirical significance of the neoclassical position is also open to debate, given that constancy of the capital–output ratio is one of Kaldor's (1961 [1989]) celebrated stylized facts. But these points notwithstanding, the claim that the Solow model 'solves' the first Harrod problem by relaxing the constancy of the capital–output ratio is simply false.[16] Instead, and as was discussed in Chapter 1, the critical feature of the Solow model that establishes the natural rate of growth as the equilibrium rate of growth towards which the economy automatically gravitates is its incorporation of Say's law: the assumption that demand automatically and passively adjusts to accommodate potential output as determined, at any point in time, by the availability and productivity of factors of production.[17] The variability of the capital–output ratio is a *facilitating* not a *causal* factor in the automatic adjustment of the Solow model towards the natural rate. This difference is akin to that between oil and petrol (gasoline) in a motor vehicle. It is petrol, not oil, that makes a vehicle move, following its ignition in the process of internal combustion. This does not, however, mean that removing the oil from the vehicle's engine would be of no consequence. On the contrary, oil plays an important facilitating role, insofar as the engine would seize up without it. Nevertheless, simply adding oil to the engine of a vehicle will not make it go, and, in like fashion, simply imbuing an economy with a flexible capital–output ratio will not eliminate post-Keynesian effective demand problems.

The second Harrod problem, meanwhile, is rejected as an unrealistic 'knife-edge' property, predicting ever-increasing (or decreasing) growth rates in response to any slight deviation from equilibrium conditions of a type that are nowhere observed in the capitalist growth record. However, as first discussed in section 3.2.2, this is an unsophisticated criticism based on a superficial understanding of Harrod's macrodynamics as articulated up to this point. The alleged 'knife-edge' property of the second Harrod problem was attributed to Harrodian dynamics by Solow (1956), and is not consistent with Harrod's own interpretation of the phenomenon.[18] Hence, as noted by Asimakopulos (1991, p. 161), even Harrod's earliest work on growth theory (Harrod, 1939) made it clear that the operation of the instability principle involves the *reaction time* (a period of about six months in his reckoning) required for firms to respond to discrepancies between actual and expected events. Fleeting departures from the warranted rate will not trigger the sort of changes in investment behaviour required to propel Harrodian instability. In his later work, Harrod expanded on this theme. Harrod (1970, 1973) argues that the size (as well as the duration) of departures from the warranted rate play a crucial role in determining whether or not the second Harrod problem will materialize, ultimately likening the instability principle to a 'shallow dome' rather than a 'knife edge', from which position the economy would require 'a much larger push . . . to set it moving' (Harrod, 1970, p. 740). As stated in his final book:

> if they [deviations of the actual from the warranted rate] are of moderate dimensions, I would not suppose that they would bring the instability principle into operation. That is why I so much object to the knife-edge idea. It requires a fairly large deviation . . . to bring the instability into play. (Harrod, 1973, p. 33)

Other criticisms of Harrod's macrodynamics have focused on the fact that Harrod framed his analysis in terms of growth rates rather than tracking the growth paths of the *levels* of output, capacity and so on. This was part of the motivation for the later work of Domar (1946), who further developed Harrod's approach by seeking to track the evolution of full capacity or potential output in the growth process and generally further mathematize the model, thus leading to what became known as the 'Harrod–Domar model'. It should also be noted that this 'problem' is, once again, a matter that has been addressed in neo-Harrodian models, to which we will turn our attention in Chapter 6. In the meantime, it seems appropriate to conclude at this stage that dismissive criticisms of Harrod's dynamics are – if the reader will pardon the pun – unwarranted.

## 3.3 The early Kaldorian model of growth and distribution

After World War II ended in 1945, initial fears that the western economies would return to the depressed conditions of the 1930s were not realized. On the contrary, by the 1950s the US and many other nations were enjoying what was later dubbed (for example, by Marglin and Schor, 1990) 'the golden age of capitalism': an era of relatively rapid and stable growth with only minor recessions. During this prolonged boom, which lasted roughly from the late 1940s to the late 1960s, there was normally (except for brief and mild recessions) something close to full employment in most national labour markets, at least in the advanced industrialized nations. Consequently, the next generation of growth theorists after Harrod and Domar became interested in explaining, contrary to what the latter had analysed, how countries could grow *at* their 'natural rates', maintaining full employment in the long run having first achieved it in the short run.[19] But, drawing upon the analytics of the Harrod–Domar model, this required providing a theoretical account for what forces could reconcile the warranted and natural rates of growth.

In the 1950s, economists of all stripes tended to take the natural rate of growth as exogenously given (that is, they assumed that both labour supply and labour productivity grow at exogenously fixed rates), so they focused on how the warranted rate could adjust to equal a given natural rate. In the neoclassical tradition, as discussed above and in Chapter 1, Solow (1956) and Swan (1956) argued that equality between $y_w$ and $y_N$ could be achieved if the capital–output ratio $a_1$ was flexible. The early work of neo-Keynesian economist Nicholas Kaldor took a different approach. (We emphasize that this was his early work because he later changed his views on long-run growth analysis considerably, as will be discussed in Chapter 8.) Kaldor (1955–56) regarded the parameters $n$, $q$ and $a_1$ as exogenously given, and analysed how the saving propensity $s$ could adjust endogenously to bring about full-employment growth in the long run. Thus, Kaldor was concerned only with the first Harrod problem (disequilibrium between the warranted and natural rates of growth), not with the second (instability of the actual rate around the warranted rate). It should be noted, however, that Kaldor's main emphasis in his early work was to construct a macro-level theory of income distribution that could serve as an alternative to the neoclassical theory based on diminishing marginal productivity of factor inputs, not to address stability issues.

For this purpose, Kaldor borrowed the idea from the classicals, Marx and Kalecki that there were different saving propensities out of profit and wage income. More like Kalecki than Marx or the classicals, Kaldor allowed for the possibility of positive saving out of wages, but with a lower propensity than saving out of profits: $0 \leq s_w < s_r < 1$, where the subscripts $w$ and $r$ indicate wage and profit income respectively. In this case, total saving is a weighted average of saving out of wage and profit income, where the weights are the shares of wages $(1 - \pi)$ and profits $(\pi)$ in total national income (as defined in Chapter 2):

$$S = [(1 - \pi)s_w + \pi s_r]Y \tag{3.14}$$

For expositional purposes, we shall abstract from labour-saving technological progress here and assume that $q = 0$, implying that the natural rate of growth is simply $n$. Also, following the practice of all the economists covered in this chapter except Harrod, we switch from the growth rate of income ($y = \Delta Y/Y$) to the rate of capital accumulation ($g = \Delta K/K$) as our benchmark measure of economic growth; the latter is also more consistent with the way we modelled the classical-Marxian theories in the previous chapter. Also, as in the classical-Marxian theories in Chapter 2, we will assume here that the economy operates at a 'normal' or 'full' rate of capacity utilization in the long run, and for convenience we will normalize this rate to unity: $u = u_n = 1$. Assuming also an exogenously given capital–output ratio $a_1$, and with $q = 0$ implying constancy of $a_0$, the capital–labour ratio ($K/L = a_1/a_0$) must then remain constant. In this case, the capital stock must grow at the same rate as the labour force in order to maintain full employment (or a constant employment rate) in the long run

$$g = \frac{I}{K} = n \tag{3.15}$$

On the other hand, saving–investment equilibrium requires that the capital stock must grow at the rate made possible by saving out of current income

$$g = \frac{S}{K} = [(1 - \pi)s_w + \pi s_r]\frac{Y}{K} = \frac{(s_r - s_w)\pi + s_w}{a_1} \tag{3.16}$$

Setting equations (3.15) and (3.16) equal to each other, which is equivalent to equalizing investment and saving (or assuming that the natural and warranted growth rates must be equal), we can solve for the equilibrium profit share that is necessary for the economy to grow at the natural rate $g_N = n$[20]

$$\pi^* = \frac{a_1 n - s_w}{s_r - s_w} \tag{3.17}$$

where evidently it must be assumed that $s_w < a_1 n$ for the profit share to be positive. In addition, since the profit share must be less than unity in order for wages to be positive, we can easily see that $\pi^* < 1$ implies $s_r > a_1 n$.[21]

This pair of inequalities can be given an economic interpretation. If we recall that $n = g$ in long-run equilibrium, then $a_1 n = a_1 g = (K/Y)\cdot(\Delta K/K) = \Delta K/Y = I/Y$, in other words, $a_1 n$ equals the investment share of output when the economy is growing at the natural rate. Thus, the Kaldor model only works under the condition that

$$s_w < a_1 n < s_r \tag{3.18}$$

or, equivalently (in a long-run equilibrium with the economy growing at the natural rate),

$$s_w < I/Y < s_r \tag{3.19}$$

Of course, this makes perfect intuitive sense, since if the average saving propensity is going to equal the investment rate (both measured as shares of output), and the average saving propensity is a weighted average of the saving propensities out of wage and profit income, then the investment rate must lie in-between those two saving propensities.

In effect, Kaldor's model implies that the distribution of income – the shares of national income going to wages and profits – must adjust so that the weighted average saving propensity for the economy as a whole reaches exactly the right level to make Harrod's warranted rate equal the natural rate (here denoted in terms of the growth rate of the capital stock rather than that of output).[22] To see this, note that drawing on (3.15) and (3.17), the equilibrium solution of the Kaldor model can be written as

$$g^* = n = \frac{(1 - \pi^*)s_w + \pi^* s_r}{a_1} = \frac{(s_r - s_w)\pi^* + s_w}{a_1} = \frac{s^*}{a_1} \tag{3.20}$$

where the last two expressions are alternative ways of writing the warranted rate of growth, $s^*$ is the equilibrium (endogenous) average saving propensity and $n$ is the natural rate of growth (recall we are assuming here that $q = 0$).

Thus, contrary to Solow–Swan, in the Kaldor model it is the saving rate rather than the capital–output ratio that adjusts to bring about full-employment growth in the long run, and the adjustments in the average saving propensity are brought about by shifts in income distribution. Moreover, Kaldor reversed the causal logic of most of the classical and Marxian models: instead of taking a distributional parameter like the real wage or wage share as given and then deducing the rates of profit and growth, Kaldor identified the condition for growth to occur at the natural rate (thus solving the first Harrod problem) and deduced what the distribution of income (profit share and corresponding profit rate and real wage) would have to be in order to achieve that outcome.[23]

For this solution to be meaningful, two additional (and related) conditions are required. First, some short-run adjustment mechanism is required to solve the second Harrod problem and ensure that output and capital actually grow at the warranted rate in the long run, rather than deviating from it. This could be achieved, for example, through government stabilization policies (fiscal and/or monetary policies) that intervene to stabilize output growth. Second, the profit and wage shares of national income must be flexible, which in turn requires that nominal wages and prices are also flexible in the long run and adjust to the 'right' levels that yield the equilibrium *real* wage $w^* = (1 - \pi^*)/a_0$. In effect, Kaldor assumed flexibility of wages and prices in the long run, which would guarantee the achievement of an equilibrium distribution of income that would ensure growth with full employment (or a constant employment rate).

Kaldor's early model has the disturbing implication that, for an economy to grow faster (assuming given technological conditions and saving propensities), it must allow the distribution of income to become more unequal, because the share of income going to the more wealthy owners of capital must rise at the expense of the share going to the less wealthy workers. Of course, this makes intuitive sense, since a higher natural rate of growth $g_N = n$ implies faster growth of the labour force, and one would expect this (all else being equal) to lower the real wage.[24] Thus, Kaldor's model is often cited as implying the need for greater inequality (a higher profit share) in order to allow for more rapid growth in the early stages of economic development, for example in relation to Kuznets's (1955) famous inverted-U hypothesis about the relationship between inequality and growth (in which countries must allegedly pass through a phase of rising inequality as they increase their per capita incomes up to some point, after which inequality supposedly diminishes – an empirical hypothesis that has been convincingly refuted by Piketty, 2014).

Indeed, the conclusion that a higher profit share is necessary to promote more rapid growth (in the absence of technological change) is a conclusion that we saw in most of the classical-Marxian models that we covered in Chapter 2. To see the similarity most clearly, note that the classical-Marxian model with an exogenously given natural rate of growth is essentially the same as Kaldor's model with zero saving out of wages. If $s_w = 0$ in the Kaldor model, then the equilibrium profit share is simply

$$\pi^* = a_1 n / s_r \tag{3.21}$$

As we saw in Chapter 2, the profit rate in a classical model can be written as $r = \pi / a_1$, so the equilibrium solution of the Kaldor model with $s_w = 0$ can be expressed alternatively in terms of an equilibrium profit rate as $r^* = n/s_r$, or

$$g^* = n = s_r r^* \tag{3.22}$$

Equation (3.22) is exactly the same solution as we obtained in subsection 2.4.3 of Chapter 2 for the classical-Marxian model with a natural rate of growth (full employment), in the special case where the growth rate of labour supply is exogenous (at $n = n_0$) and there is no minimum profit rate for positive savings ($r_{min} = 0$).

Equation (3.22) is the famous 'Cambridge equation', which in this context should be interpreted as implying that the equilibrium profit rate is determined by the natural rate of growth and the saving propensity out of profits ($r^* = n/s_r$). Observe that, written in this form, the Cambridge equation reveals that for any $s_r < 1$, we will observe $r > g$ – the result made famous recently by Piketty (2014). Recall, however, that according to equation (3.21), this outcome is consistent with a constant equilibrium profit share, $\pi^*$. Hence according to the Kaldor model, and contrary to Piketty (2014), $r > g$ need not be associated with increasing income inequality, at least in terms of the functional distribution between capital and labour income.

What the similarity between equation (3.22) and the solution for the classical-Marxian model with a natural rate of growth (covered in subsection 2.4.3 of Chapter 2) draws attention to is that, except for allowing for positive saving out of wages, Kaldor's early model is in many respects more classical than Keynesian. Aggregate demand plays no role in this model, as growth seamlessly adjusts to the supply-side-determined natural rate, while flexibility of wages and prices is assumed (contrary to what was argued by Keynes in the *General Theory*) to ensure a situation of sustained full employment. These observations would be modified if stabilization policies are required to adjust $g$ towards $n$. But since Kaldor did not explicitly include demand forces in his early model, it is therefore vulnerable to criticism for being un-Keynesian.

### 3.3.1 The Pasinetti and neo-Pasinetti theorems

In evaluating Kaldor's model of income distribution, it is important to briefly note the important criticism of Pasinetti (1962). Recall that we have carefully defined $s_w$ and $s_r$ as the saving propensities out of wage and profit *income*, not the saving propensities of workers and capitalists as social classes. This distinction is vital because, as Pasinetti pointed out, if workers save then in the long run they must come to own a portion of the capital stock and therefore they must also receive a portion of the profits (for example, through interest on pension funds that hold corporate stock). However, *if* the profits received by workers are saved at the lower saving propensity $s_w$, then Kaldor's model is incorrect because it assumes that *all* profits are saved at the higher rate $s_r$.

Pasinetti's critique is thus based on the view that the relevant saving propensities pertain to social classes, not to the functional categories of income they receive. In Pasinetti's view, the justification for assuming a higher saving propensity for capitalists (that is, those who only own capital and receive profits, and do not perform labour or receive wages) is that they are generally in much higher income brackets than workers, and hence would be expected to save a significantly higher percentage of their income than workers do. This assumption can be viewed as following Keynes (1936), who argued that the marginal propensity to consume would be a decreasing function of household income levels. On this assumption, Pasinetti (1962, 1974) showed that if $s_w$ and $s_r$ are interpreted as the saving propensities of workers and capitalists, respectively, and if workers save their portion of profits at the lower rate $s_w$, then – in a long-run steady state with full employment – the equilibrium condition turns out to be the same as in the original Kaldor model when $s_w = 0$, that is, $g^* = n = s_r r^* = s_r \pi^*/a_1$ (per equations 3.20 and 3.22 above). In other words, the Cambridge equation applies, and the workers' saving propensity does not matter to the long-run equilibrium distribution of income between wages and profits, even if workers' savings are positive.

For detailed proofs of the Pasinetti theorem and discussion of neoclassical responses, one should refer to his original work (Pasinetti, 1962, 1974) and various later texts (for example, Harris, 1978; Marglin, 1984b). Here, we sketch out a simple intuitive explanation, mostly following Hein (2014). In a long-run, steady-state equilibrium in which workers own part of the capital stock, the workers' and capitalists' shares of the capital stock must be constant, which means that the workers' and capitalists' capital stocks must both grow at the same rate as total capital. At the same time, since they both own the same asset (capital), both workers and capitalists must earn the same rate of return: $r_c = r_w = r$. The capitalists' capital (denoted by $K_c$) grows at the rate $g_c = \Delta K_c/K_c = s_r r K_c/K_c = s_r r$ (since all savings out of profits are invested in new capital) and this must be equal to the overall growth rate of the capital stock in order to maintain constant shares, so $g = g_c = s_r r$.

Thus, the workers' saving propensity does not enter into the determination of the overall growth rate of the economy. Rather, it can be shown that what the workers' saving propensity does determine is the workers' steady-state share of the total capital stock and total profits – not the total amount, share or rate of profit. Furthermore, it can be shown that (in the steady state) the workers' saving out of their wages exactly equals their 'extra' consumption out of their part of the profits, that is, the portion of the profits that would otherwise have gone to capitalists' saving if the capitalists owned all of the capital. Hence, the total amount of saving is not affected by the workers' saving at the lower propensity $s_w < s_r$. However, if $s_w < s_r$, then in order for the Pasinetti theorem to hold it must be true that $s_w < \pi s_r$, as shown by Samuelson and Modigliani (1966) – see Appendix 3.1 for a proof. Since $\pi < 1$, this stricter condition on $s_w$ may not hold even if $s_w < s_r$ – indeed, Samuelson and Modigliani (1966, p. 274) expressed some scepticism that it would. Modern stylized facts suggest otherwise, however. With a profit share of 30–35 per cent and saving propensities of 20–25 per cent for the top 1 per cent of wealth owners and 3 per cent for the bottom 90 per cent (Saez and Zucman, 2016), the US economy comfortably satisfies the condition $s_w < \pi s_r$ required for the Pasinetti theorem to hold.

Kaldor (1966b) issued a rejoinder in which he defended and extended his original approach. Kaldor justified the assumption of a higher saving propensity out of profits not because of the higher incomes of capitalists compared with workers, but by reference to corporate savings. That is, firms retain a portion of their profits as retained earnings, which are counted as corporate saving in the national income accounts.[25] In this view, even if all households (workers who may own some part of the capital and pure capitalists who only own capital) save at the same rate (call it $s_h$), the saving propensity out of profits will be higher than out of wages because a portion ($s_c$) of the profits are saved (retained) by the corporations before the remainder is paid out (as dividends or interest) to the households who own equity in the firms. Then the saving propensity out of profit income still exceeds the saving propensity out of wage income: $s_r = s_c + (1 - s_c)s_h = (1 - s_h)s_c + s_h > s_w = s_h$. In this case, Kaldor's original model would be valid.

However, Kaldor (1966b) went further and developed a more explicit alternative to the Pasinetti model in which household saving also plays no role in the long-run equilibrium. This so-called neo-Pasinetti theorem of Kaldor's can be sketched out briefly as follows (in a simplified version, ignoring the role of capital gains). Suppose that households accumulate all their savings by purchasing shares of equity in corporations. Let $\chi$ be the share of investment ($I = \Delta K$) that is financed by issuing new equity to households, while $1 - \chi$ is the share financed internally out of profits retained by corporations (again assuming a corporate saving or retention propensity of $s_c$). Then investment would be equal to

$$I = \Delta K = (1 - \chi)\Delta K + \chi\Delta K = s_c rK + \chi\Delta K \tag{3.23}$$

Dividing both sides by $K$, we get

$$g = I/K = s_c r + \chi g \tag{3.24}$$

which solves for the profit rate

$$r = (1 - \chi)g/s_c \tag{3.25}$$

And when growth occurs at the natural rate ($g = g_N = n$), the equilibrium profit rate is

$$r^* = (1 - \chi)n/s_c \tag{3.26}$$

and the corresponding profit share is $\pi^* = a_1 r^* = (1 - \chi)a_1 n/s_c$. Thus, Kaldor's neo-Pasinetti theorem arrives at a similar substantive conclusion to the original Pasinetti theorem – that the savings decisions of working-class households do not affect the long-run rate of growth or distribution of income in the economy – via a different route, with emphasis on the saving and financing behaviour of corporations rather than the saving behaviour of wealthy capitalist households. However, Kaldor's result depends on several strong assumptions, including that the only financial asset in which workers' households can hold their savings are shares of corporate equity, all households (workers and capitalists) have the same propensity to save out of personal income, and workers' capital receives the same rate of return as capitalists' capital.

The various models of Kaldor and Pasinetti from the 1950s and 1960s have proved to be more enduring than their authors' own support for them. In his later work, Kaldor abandoned his earlier focus on identifying the distributional conditions for full-employment growth and fundamentally shifted his thinking about long-run growth in new directions. Especially, Kaldor (1961 [1989], 1966a [1989]) began to focus on the endogeneity of technological progress, increasing returns to scale in production, the role of exports in the growth process, and feedbacks from aggregate demand to aggregate supply conditions, as will be discussed in Chapter 8. Pasinetti always maintained that his 1962 theorem was only intended to correct a logical flaw in the original Kaldor model, and he too turned his attention in other directions, ultimately leading to his later work on structural change (Pasinetti, 1981, 1993). Nevertheless, the early models of Kaldor and Pasinetti are still an important part of the intellectual heritage of macroeconomic theory, especially because they highlight the centrality of alternative specifications of saving behaviour and corporate finance in macro modelling. We will return to the importance of different assumptions about saving in later chapters, but here we turn instead to the approach of Robinson, who sought to escape from the straightjacket of analysing the conditions for growth with full employment and to inject more of a Keynesian focus on aggregate demand into models of long-run growth.

## 3.4 The neo-Robinsonian model of profits and growth

The model presented in this section is inspired mainly by the work of Joan Robinson (1956, 1962). However, what is presented here is not her original version, which is discussed in Harris (1978) and summarized in Appendix 3.2. Rather, we will present a neo-Robinsonian model of the sort developed in later works including Marglin (1984a, 1984b), Dutt (1990) and Hein (2008, 2014). The key aspect of Robinson's model is her effort to depict a growth process that is driven by the desired investment spending of business firms and which may diverge persistently from a full-employment growth path. Robinson was the first neo-Keynesian growth theorist who explicitly introduced an independent investment function (that is, an investment function that is separate and distinct from the saving function) into a long-run growth model, rather than assuming that investment would necessarily adjust to yield growth at the natural rate (as in Kaldor) or would generally follow an unstable adjustment process (as in Harrod). In fact, Robinson created a model of growth and distribution that allowed for both stable and unstable equilibria, as we shall see below.

### 3.4.1 Saving, investment and the equilibrium growth rate

For Robinson's model, we return to the classical assumption that profits are the sole source of a society's savings. Thus, her saving function is the same as the classical-Marxian saving function (2.8) from the previous chapter, except there is no minimum level of profits required for positive savings ($r_{min} = 0$); this is also equivalent to Kaldor's saving function (equation 3.16) assuming $s_w = 0$. Because Robinson so clearly distinguished saving from investment, we will also introduce a notational distinction here by using $\sigma \equiv S/K$ to represent the ratio of saving to capital or 'saving rate',[26] while $g \equiv I/K = \Delta K/K$ will continue to represent the investment–capital ratio; the latter should be considered as the 'growth rate' of the economy since it is the rate of increase in the capital stock. Thus, our neo-Robinsonian saving function is

$$\sigma = s_r r \tag{3.27}$$

where $s_r$ $(0 < s_r < 1)$ is the marginal propensity to save out of profits and $r$ is the profit rate. If we continue to assume that the economy operates at a normal rate of capacity utilization in the long run and normalize that rate to unity $(u = u_n = 1)$, the profit rate is given by $r = \pi / a_1$.[27] To clarify terminology, we will refer to $s_r$ as the saving *propensity* out of profits (which we take as exogenously given), while we will refer to $\sigma$ as the realized saving *rate* (which we treat as endogenous).

On the investment side, we will assume that the desired rate of investment depends on expected future profits, in a relationship determined by what Keynes (1936) called the 'animal spirits' of entrepreneurs, or their degree of optimism (bullishness) or pessimism (bearishness) about the future returns to current investment. In this view, it is not possible to determine a strictly optimal level of investment based on an exact calculation of expected, discounted future returns, because of fundamental uncertainty. No one can know the future, and as Keynes pointed out, we do not even know the probability distributions of future events ('states of the world') that would be needed to make precise calculations of expected returns. Therefore, a neo-Robinsonian investment demand or desired investment function can be written in very general form as

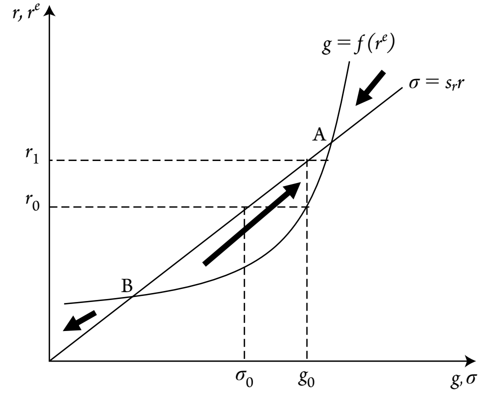
$$g = f(r^e) \tag{3.28}$$

where $r^e$ is the expected profit rate and we continue to assume (as previously) that there is no depreciation so that all investment is net investment and hence $g = \Delta K/K$. Technically speaking (and in the classical language favoured by Robinson), $g$ is really the 'rate of capital accumulation' (rate of increase in the capital stock), but we will refer to it as 'the growth rate' or 'investment rate' (depending on the context) in what follows.[28]

Initially, we follow Robinson in assuming a positive but diminishing marginal impact of expected profits on desired investment, that is, we assume $f' > 0$ and $f'' < 0$. The assumption of diminishing responsiveness of investment to expected profits can be justified on the basis that expectations of higher future profit rates are likely to be held with less certainty as the level of expected profit rises, and hence firms would be more cautious in boosting investment demand in response to increases in expected profits when expected profits are already higher to begin with.[29] Both the shape of the function $f(\cdot)$, that is, how much investment responds to given expectations about future profits, and the profit expectations themselves depend fundamentally on firms' animal spirits, and hence the demand for investment cannot be mechanically predicted by traditional variables such as interest rates or the 'cost of capital'. To be clear, Robinson did not deny that interest rates and other variables could influence investment to some degree, but in contrast with conventional models she emphasized the expected profit rate (subjective anticipations of the real returns to additional capital goods purchased) rather than the interest rate (cost of borrowed funds, or opportunity cost of using internal funds) as the main driver of investment expenditures.

The saving and investment functions (3.27) and (3.28) are combined in what has become known affectionately as 'Robinson's banana diagram', shown in Figure 3.2. Using the old British style (similar to a Marshallian

**Figure 3.2** The Robinson 'banana diagram' with multiple equilibria



supply-and-demand diagram), the dependent variable ($\sigma$ or $g$) is placed on the horizontal axis, while the independent variable ($r$) is on the vertical axis, so the slopes are inverted from what we would normally think (thus, the slopes of the curves as drawn are $1/s_r$ and $1/f'$, so a higher $s_r$ or $f'$ implies a flatter slope). In a long-run equilibrium, expected and actual profits must coincide ($r^e = r$) and – assuming there is no foreign trade or government budget for simplicity – saving and investment have to be equal ($\sigma = g$), so the economy must be at a point such as A or B in Figure 3.2.

It is easily seen that point A, where the investment function cuts the saving function from below, or the saving propensity exceeds the responsiveness of investment to profits ($s_r > f'$), is a stable equilibrium, while point B, where the investment function cuts the saving function from above, or the response of investment to profits exceeds the saving propensity ($f' > s_r$), is unstable. One way to see this is mathematically. In standard fashion, we analyse the stability of the goods market by finding the condition for changes in the adjusting variable (here, the profit rate $r$) to eliminate any excess demand for (or supply of) goods. In this simple model with no government or international trade, excess demand for goods (normalized by the capital stock), evaluated in the neighbourhood of an equilibrium point at which $r^e = r$, can be written as[30]

$$EDG = g - \sigma = f(r) - s_r r \tag{3.29}$$

Then, the equilibrium is stable if $r$ rises when $EDG > 0$ and falls when $EDG < 0$, and if these adjustments in $r$ would tend to eliminate excess demand by affecting saving more than they affect investment. Mathematically, this requires

$$\frac{dEDG}{dr} = f'(r) - s_r < 0 \tag{3.30}$$

or $s_r > f'$, where the derivative $f'$ is evaluated in the neighbourhood of an equilibrium where $r = r^e$. Given the inverted way in which the diagram is drawn, this must occur at a point at which the saving function is flatter than the investment function.

Alternatively, one can reason geometrically from the diagram itself. Starting from any given profit rate such as $r_0$, if the initial desired investment rate $g_0$ is greater than the corresponding saving–capital ratio $\sigma_0$, the resulting rate of investment will push up the actual profit rate to a higher level ($r_1$), and assuming that firms then revise their profit expectations upward, this will in turn induce a still-higher rate of investment, and so on. Thus, starting from any point between A and B, the profit and growth rates will both move upward, away from B and towards A. Conversely, if $g$ is less than $\sigma$ initially, which would be the case if the initial profit rate were above A or below B, the profit rate would fall leading to a contraction of investment, followed by a further fall in the profit rate and contraction of investment. From points above A this would be a move towards the equilibrium at A while for points below B it would be a move away from this equilibrium and towards the origin, that is, an economic collapse. Hence, we conclude that point A is a stable equilibrium while B is an unstable one.

Although the analytics are not exactly the same, the unstable equilibrium at point B bears a strong resemblance to the second Harrod problem of instability of the actual growth rate relative to the warranted rate, which was illustrated in Figure 3.1. Since Harrod's 'warranted' rate is defined as the growth rate that equilibrates saving and investment, the growth rates at both points A and B in Figure 3.2 can be considered warranted rates. Also note that expectations would be fulfilled $(r = r^e)$ at each point, A and B. Robinson's analysis thus shows that, in the presence of nonlinear investment responses, there may be multiple warranted equilibria for growth, some of which may be stable and others unstable. In Robinson's model, instability arises from an overly exuberant response of investment to a deviation of actual from expected profits, since $f' > s_r$ at point B. However, Robinson effectively 'tames' this instability by assuming a diminishing degree of responsiveness of investment to profit expectations, so that $f'$ falls as $r$ rises and eventually an equilibrium where $f' < s_r$ is reached at the stable equilibrium point A; this outcome depends on the desired investment curve having sufficient curvature to eventually become steeper than the saving line at a feasible level of the profit rate.

## 3.4.2 Connecting accumulation back to distribution

If we adopt the classical inverse wage–profit relation (2.4) from Chapter 2 (assuming, as before, a given technology with a single good, fixed coefficients, and full or normal utilization of capacity at the normalized rate $u = u_n = 1$), we can see that the equilibrium profit rate at a stable point like A in Figure 3.2 must correspond to a unique equilibrium real wage. Figure 3.3 combines the inverse wage–profit relation[31] on the left-hand side with a neo-Robinsonian saving–investment diagram on the right, except that the nonlinear investment function (3.28) shown in Figure 3.2 is replaced with a linear version for analytical convenience (and a stable case is assumed).

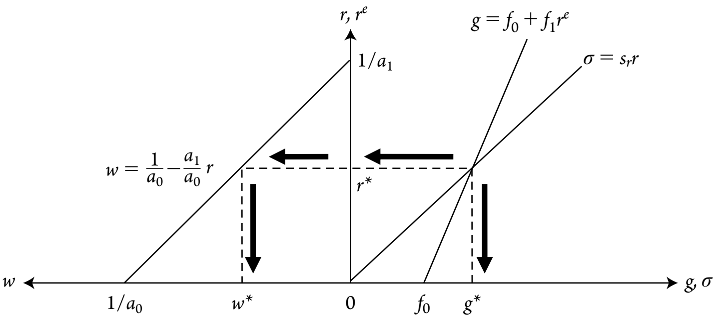



**Figure 3.3** Neo-Robinsonian growth model with a linearized investment function combined with a classical wage–profit inverse relation (bold arrows show the direction of causality)

More specifically, the linear neo-Robinsonian investment function can be written as

$$g = f_0 + f_1 r^e \tag{3.31}$$

where $f_0, f_1 > 0$. In this linear specification, we will usually think of the intercept $f_0$ as a 'shift factor' representing the firms' animal spirits or the degree of business confidence, although in principle animal spirits could also be infused into the slope parameter $f_1$, which shows the degree to which investment spending responds to given profit expectations. Again, given that $r = r^e$ and $g = \sigma$ must hold in a long-run equilibrium, causality in the model begins with the simultaneous solution of the saving and investment functions (equations 3.27 and 3.31) for $r$ and $g$, which occurs at the point of intersection in the right-hand quadrant, and then flows to the left to determine the real wage as shown by the bold arrows in Figure 3.3.

Mathematically, using the linear investment function (3.31), we can easily obtain an explicit, reduced form solution of the model. We start with the equilibrium condition

$$\sigma = g \tag{3.32}$$

into which we substitute equations (3.27) and (3.31), while assuming also that expectations are fulfilled so that $r^e = r$ in the long run. The simultaneous solutions for the long-run equilibrium profit and growth rates are as follows

$$r^* = \frac{f_0}{s_r - f_1} \tag{3.33}$$

and

$$g^* = \frac{s_r f_0}{s_r - f_1} \tag{3.34}$$

Now – and this is a property of this model that we will see repeated many times below – satisfaction of the stability condition can be used to sign the denominators (and to help sign the derivatives) of these solutions. Using the method shown earlier, it is easily seen that stability of the goods market now requires $s_r > f_1$ or $s_r - f_1 > 0$. Then, substituting (3.33) into the wage–profit relation (2.4) from Chapter 2 yields the solution for the long-run, equilibrium real wage[32]

$$w^* = \frac{1}{a_0} - \frac{a_1}{a_0}\left(\frac{f_0}{s_r - f_1}\right) \tag{3.35}$$

Although it is not shown on the graph, the model also solves for the equilibrium profit share of national income, which must be $\pi^* = a_1 r^*$.

For this solution to be meaningful, the real wage would have to be flexible in the long run, so that it would gravitate to (and workers would accept) whatever real remuneration is dictated by equation (3.35). Similarly, the wage share must also be flexible and adjust to the equilibrium level $1 - \pi^*$. Robinson and other neo-Keynesians were well aware that such flexibility of the real wage and relative shares might not prevail in reality, and we will discuss critical perspectives on (and alternative approaches to) the issue of wage determination below. But first, we explore the implications of the solution in equations (3.33) through (3.35) for the effects of changes in certain key parameters, which will provide greater insight into the properties of the model and also the plausibility of its distributional assumptions.

### 3.4.3 The long-run paradox of thrift and the underlying price mechanism

In this model, a deliberate effort to increase the society's saving propensity, without changing the firms' willingness to invest as reflected in a given investment function (3.28 or 3.31), will necessarily backfire; such an increase in the saving propensity not only will fail to boost growth, but actually will reduce the equilibrium rates of profit and growth $(r^*, g^*)$ – and the equilibrium saving rate $\sigma^* = g^*$ as well! This can be seen graphically in Figure 3.4, where an increase in the saving propensity out of profits (recall that the saving propensity out of wages is implicitly zero here) rotates the saving line downward (towards the $g$-axis) and causes the equilibrium point to move down and to the left (from A to A′) along a given investment function. The intuition for this result is clear: a rise in the proportion of profits saved requires a diminution in the consumption spending of the profit recipients (owners of the firms), which reduces aggregate demand.

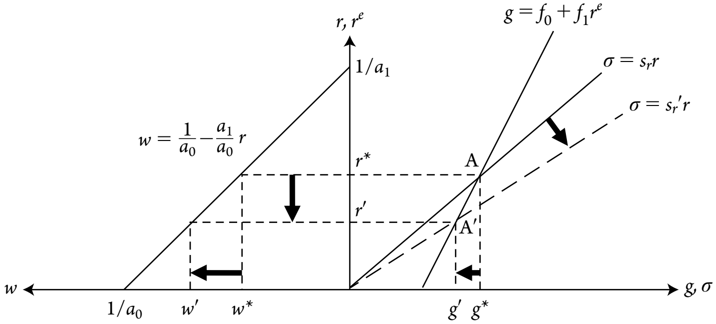



**Figure 3.4** The paradox of thrift: effects of a rise in the saving propensity out of profits (bold arrows show shifts from one equilibrium state to another)

As demand falls, the amount of profit realized from the existing investment expenditures also falls. However, once the actual profit rate starts to fall and firms revise their profit expectations downward, the incentives to invest are reduced, so investment also begins to decline (the reduction in investment is, however, a move down along the investment function, not a shift in that function). The decline in investment then leads to further reductions in aggregate demand and realized profits, which in turn further dampen expectations and further reduce investment, and so on. This process continues until, assuming the system is stable, it converges to new, lower equilibrium rates of profit and growth at the point A′ in Figure 3.4. The reductions in equilibrium $r$ and $g$ in this case constitute the dynamic equivalent of Keynes's short-run paradox of thrift. Moreover, the equilibrium saving rate $\sigma^*$ is also lower at point A′.

Mathematically, these results can be obtained in one of several ways. Given the reduced form solutions (3.33) to (3.35), one can obtain the partial derivatives with respect to the saving propensity $s_r$, and one can see that $\partial r^*/\partial s_r < 0$, $\partial g^*/\partial s_r < 0$, and $\partial w^*/\partial s_r > 0$, assuming that the stability condition $s_r - f_1 > 0$ holds. Alternatively, one can substitute equations (3.27) and (3.31) into the equilibrium condition (3.32) and totally differentiate with respect to $r$ (assuming $r^e = r$) and $s_r$, and solve the result for $\partial r^*/\partial s_r$; then the chain rule can be applied to derive $\partial g^*/\partial s_r$ and $\partial w^*/\partial s_r$ from equations (3.31) and (2.4), respectively.

Although the intuition for the negative effects on the profit and growth rates is clear and sensible, the intuition for the predicted rise in the equilibrium real wage is harder to see. Mechanically, of course, if there is a strictly inverse relationship between wages and profits, then wages must rise if profits fall, but what force propels the real wage to increase in this circumstance? Since growth is reduced, one would not think that the demand for labour is rising faster (quite the contrary). However, one should not think in terms of a classical-Marxian labour market in which the real wage is determined by the (growth of) supply and demand for labour in the context of the neo-Keynesian model. Rather, the nominal wage $\overline{W}$ is fixed exogenously in the bargaining process between labour and management, and the real wage $w = \overline{W}/P$ then adjusts endogenously and depends on what happens to the price level $P$. If aggregate demand is depressed because of the reduction in capitalists' consumption and the induced reduction in investment, even if prices might be 'sticky' (downwardly rigid) in the short run, we would expect them to be flexible and fall in the long run, so eventually $P$ must decrease to such a degree that the real wage rises as predicted. Thus, this model predicts that the real wage should vary inversely with the growth of the economy in the long run, similar to the early Kaldorian model considered in section 3.3. However, the assumptions that lead to this result may be questioned: nominal wages may not remain constant in spite of a falling price level and slackening labour demand, and prices may not be freely flexible downward even in the long run.[33]

### 3.4.4 Animal spirits and the 'widow's cruse'

In contrast to a rise in the saving propensity, an improvement in the level of business confidence leading to an outward shift in the investment function will cause a sustained increase in the profit and growth rates, at the expense of the real wage, in the neo-Robinsonian model. To see this, suppose that business firms become more optimistic, so that their expected rate of profit $r^e$ increases relative to the actual profit rate $r$ (in other words, expected profits become higher for any given level of current profits). This would effectively change how the actual profit rate maps into desired investment, and so could be considered as an increase in the intercept term $f_0$ in the investment function (3.31) written in terms of $r$ rather than $r^e$. Thus, firms would be willing to invest more at any given current profit rate, and hence the investment function would shift out and to the right in $g \times r$ space as shown on the right-hand side of Figure 3.5.

Mathematically, it is quite easy to show that $\partial r^*/\partial f_0 > 0$, $\partial g^*/\partial f_0 > 0$ and $\partial w^*/\partial f_0 < 0$, assuming again that the stability condition $s_r - f_1 > 0$ holds.[34] Intuitively, as the firms start spending more on investment at the initial rate of profit $r^*$ (at point A′ in Figure 3.5), they pump up demand in the economy, leading them to reap more profits from their increased sales (this would occur at a point vertically above A′ along the saving line $\sigma = s_r r$). As the actual or realized profit rate increases, this gives a further stimulus to investment, which then raises realized profits further, and the whole process continues until the economy converges to a new long-run equilibrium at

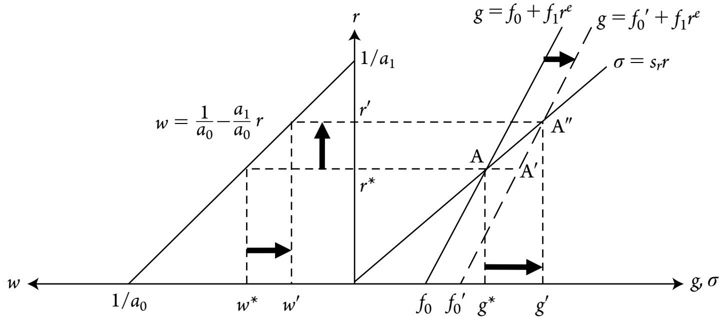



**Figure 3.5** The widow's cruse: effects of a rise in firms' animal spirits and desire to invest (bold arrows show shifts from one equilibrium state to another)

point A″ (this process will be convergent if the equilibrium is stable, which again requires $s_r - f_1 > 0$ so that the investment function is steeper than the saving line as drawn).

To understand why the real wage falls, we must once again think in terms of how the price level adjusts, assuming a given nominal wage that is fixed in labour contracts. With greater investment leading to higher growth, we would expect prices to gradually increase, and a rising price level $P$ relative to a given nominal wage rate $\overline{W}$ will lead to a fall in the real wage $w = \overline{W}/P$. The fall in the real wage is of course necessary to permit the profit rate to rise, assuming that the technology is unchanged. However, if the investment brings with it a technological improvement (say, of the Harrod-neutral or Hicks-neutral variety), then the real wage might not need to fall or might not need to fall as much as in the case of a stagnant technology (and the cost reductions brought about by the improved technology could ameliorate the pressures for the price level to rise).

Under the assumptions of the neo-Robinsonian model, the rise in investment eventually becomes completely self-financing, so there is no need for the saving function to shift outward. On the contrary, the saving function can remain stationary, but as long as it is upward sloping, then more profits are generated thereby creating the increased saving rate ($\sigma$) that is needed *ex post* to finance the increased investment rate ($g$). This occurs through a version of the standard Keynesian multiplier mechanism: the initial rise in investment spending leads to increased profits, which induce additional investment spending plus extra consumption spending by the owners of capital, which in turn boost profits further, and so on ad infinitum, but – if the stability condition holds – the increases in profits and investment become smaller and smaller until both variables converge to their new equilibrium levels. The difference is that, in the conventional multiplier story of macro textbooks, it is total output and income that rise and generate additional rounds of consumption spending (by all social classes equally, assuming a uniform marginal propensity to consume), while investment is assumed to be exogenously given. In Chapter 4, we will see another kind of model (neo-Kaleckian) in which both profits and output rise in the multiplier process.

In the neo-Robinsonian model shown here, one could say that the equilibrium amount of saving rises with the equilibrium level of investment (up to the point at which $\sigma = g$ again holds); the *ex ante* level of saving does *not* constrain investment and there is no need for any shift in the saving relationship itself, only a redistribution of income towards profits sufficient to allow for the increased equilibrium savings to occur. This result can be called the 'widow's cruse', based on an analogy that Keynes made in his *Treatise on Money* (1930) to the Biblical story of a jar or vessel used to hold oil that would never be empty, no matter how much the contents were poured out or consumed (I Kings, 17:10–16; II Kings, 4:1–7).[35] Just like the proverbial widow's cruse, the vessel of investment finance is never empty, because the very act of investing boosts the profits that are the source of the funds required to finance it. Whereas, if society attempts to increase its growth by raising the saving propensity, without any change in animal spirits or the desire to invest (at any given expected profit rate), the society not only will fail to boost growth, but will actually reduce its equilibrium growth rate $g^*$ (*and* reduce the equilibrium saving rate $\sigma^* = g^*$) by shrinking the profits that are the main stimulus to investment, as shown earlier.

## 3.5 Real wage resistance and inflation

As we have repeatedly noted, the distributional implications of the basic neo-Robinsonian model are both strong and potentially implausible. One particularly important example is that workers might not be willing to accept the reduction in their real wage shown in Figure 3.5 following an increase in investment demand, especially if (as predicted) such an increase would stimulate more rapid growth of output and employment. More generally, workers might not be satisfied with the equilibrium real wage that emerges from any given constellation of saving and investment behaviour on the part of capitalists (as in Figure 3.3). However, as Keynes (1936) emphasized, workers can seek a higher real wage only by trying to raise the nominal wage, since it is the latter and not the former that (under normal circumstances) is determined in labour contracts.[36] If workers are powerful enough to win higher nominal wages (for example, because of strong unions or support from a pro-labour government), firms could raise prices further in response, and a 'cost-push' inflationary spiral could result. Profits would then not increase as much as predicted in the simple model, and firms' animal spirits could be dampened so that the initial upward shift in investment demand might be reversed. Moreover, if profits don't rise as much as necessary to finance the prospective increase in investment, that increase will not occur as planned and growth will not increase as much as expected.

Robinson was in fact well aware of these possibilities and how they could alter the outcome predicted by her basic model. She anticipated this particular problem about rising nominal wages when she wrote about what she called a 'bastard golden age':[37]

> A bastard golden age sets in at a fairly high level of real wages when organised labour has the power to oppose any fall in the real-wage rate. Any attempt to increase the rate of accumulation, unless it is accompanied by a sufficient reduction in consumption of profits, is then frustrated by an inflationary rise in money-wage rates. In such a situation, the rate of accumulation is limited by the 'inflation barrier'. (Robinson, 1962, pp. 58–9)

Since Robinson's time, this concept of an inflationary increase in nominal wages preventing firms from realizing the profits necessary to finance an investment-led boom has more commonly become known as the case of 'real wage resistance'. We turn next to the effort of Marglin (1984a) to construct a formal model of real wage resistance by combining elements from the neo-Marxian and neo-Keynesian approaches (using the neo-Robinsonian version of the latter).

## 3.5.1 Marglin's neo-Marxian/neo-Keynesian synthesis

To provide the rationale for Marglin's approach, let us consider what happens if we attempt to combine a classical-Marxian model assuming a fixed real wage, as covered in Chapter 2, with a neo-Robinsonian model of capital accumulation driven by firms' animal spirits. The combined model includes the following equations[38]

$$w = \frac{1}{a_0} - \frac{a_1}{a_0} r \tag{2.4}$$

$$w = \overline{w} \tag{2.9}$$

$$\sigma = s_r r \tag{3.27}$$

$$g = f_0 + f_1 r \tag{3.31'}$$

$$\sigma = g \tag{3.32}$$

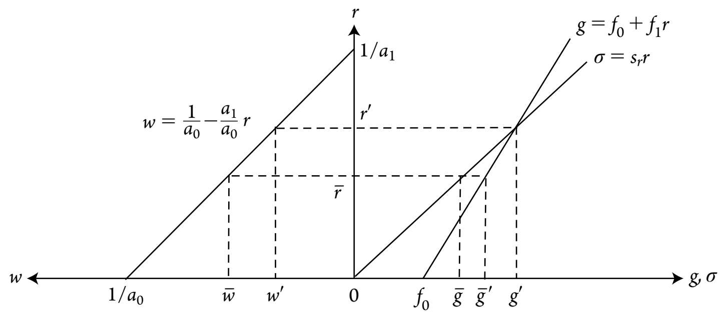



**Figure 3.6** Neo-Marxian and neo-Keynesian (Robinsonian) models combined: overdetermined solution

where we use the linearized version of the neo-Robinsonian investment function assuming $r^e = r$ for simplicity in (3.31′). It is easily seen that this model consists of five equations in only four endogenous variables ($w$, $r$, $g$ and $\sigma$), implying that the model is overdetermined – that is, it has no unique solution.

This theoretical dilemma is illustrated in Figure 3.6, which combines elements from Figures 2.2 and 3.3. The neo-Marxian part of the model consists of the four equations excluding the desired investment function (3.31′), which together imply a profit rate of $\bar{r} = (1 - a_0\bar{w})/a_1$ and a growth rate of $\bar{g} = s_r\bar{r}$. Essentially, this is a classical-style solution in which the rate of profit is determined by the exogenously given technology and real wage, and the growth rate depends on the supply of saving out of the profits thus generated. However, this solution represents a disequilibrium situation in terms of the neo-Robinsonian model, because (based on the last three of these equations) the desired investment (growth) rate is $\bar{g}' = f_0 + f_1\bar{r} > \bar{g}$ at the profit rate $\bar{r}$. In other words, there is excess demand at the profit rate $\bar{r}$, which should induce increases in realized investment and saving until the economy reaches the neo-Robinsonian solution at the point ($r'$, $g'$), defined by $f_0 + f_1 r' = s_r r'$. However, the economy cannot reach the latter point because it would require the real wage to fall to $w'$, which is below the conventional or necessary real wage of $\bar{w}$. Thus, the two solutions are mutually incompatible, and there is no unique equilibrium to the model described by these five equations (except in the special case where the functions happen to align so that $\bar{r} = r'$, which could occur only by accident).

Marglin's solution to this theoretical inconsistency is to borrow another key idea from Keynes, namely the proposition that it is the nominal rather than the real wage that is determined in the bargaining between workers and firms, and set in labour contracts. As Marx emphasized, what workers care about is their real wage, that is, their standard of living, but as Keynes stressed, the only instrument workers have for seeking to achieve their objectives for the real wage is to negotiate over the nominal wage; what this implies for the real wage then depends on what happens to the price level.[39] Or, in Marglin's words, 'The wage *bargain* may be formulated in terms of money, but in the long run *bargaining* takes place in real terms' (Marglin, 1984a, p. 129, emphasis in original). Based on this insight, Marglin suggested that inflation (that is, continuous increases in nominal wages and prices) could be the equilibrating force that reconciles the neo-Marxian and neo-Keynesian features of the synthetic model. In fact, Marglin went so far as to suggest that equality between the rates of inflation in wages and prices could be used as the long-run equilibrium condition in order to arrive at a unique, determinate solution to a hybrid neo-Marxian/neo-Keynesian model.

To see how Marglin did this,[40] we keep equation (2.4), but we combine (3.27) and (3.32) into a more classical-Marxian saving function written as

$$g = s_r r \tag{3.36}$$

thus eliminating $\sigma$ as a separate variable and ensuring that actual, realized investment is equal to saving in a long-run equilibrium. Then, we replace the other equations with the following

$$g^d = f_0 + f_1 r \tag{3.37}$$

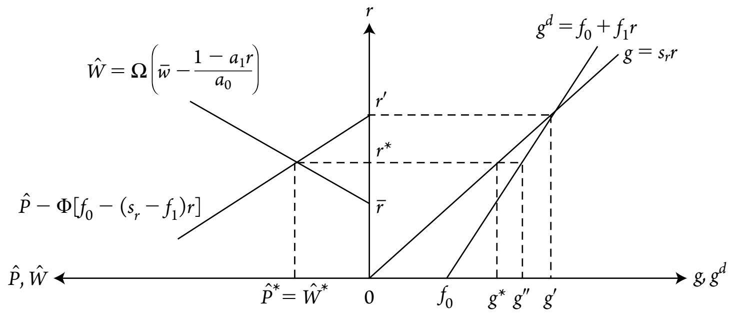
$$\hat{W} = \Omega(\overline{w} - w) \tag{3.38}$$

$$\hat{P} = \Phi(g^d - g) \tag{3.39}$$

$$\hat{P} = \hat{W} \tag{3.40}$$

where $g^d$ is desired (as opposed to actual or realized) investment, $W$ and $P$ are the nominal wage and price level (respectively), $\Omega$ and $\Phi$ are positive constants representing speeds of adjustment and a circumflex over a variable indicates its instantaneous rate of change. Equation (3.38) represents labour bargaining: it shows that the nominal wage adjusts upward in proportion to the degree to which the actual real wage $w$ falls short of the workers' target $\overline{w}$. Equation (3.39) formalizes the assumption in all the neo-Keynesian models, as discussed earlier, that price inflation is driven by excess demand in the goods market, or the gap between desired investment and saving (where the latter equals actual or realized investment, which is constrained by available saving). Finally, equation (3.40) is Marglin's equilibrium condition of equality between the rates of increase in prices and wages, which implies that

**Figure 3.7** Marglin's neo-Marxian/ neo-Keynesian (Robinsonian) synthesis model




the real wage $w = W/P$ is constant in the long-run equilibrium (implicitly, labour productivity $1/a_0$ is held constant in Marglin's model).

Thus, Marglin's hybrid model effectively consists of six equations (2.4 and 3.36–3.40) in the six endogenous variables $w$, $r$, $\hat{W}$, $\hat{P}$, $g$ and $g^d$. The solution to this model is illustrated in Figure 3.7, where we replace the wage–profit inverse relation in the left-hand quadrant with the functions representing wage and price inflation; we also substitute equation (2.4) into (3.38) and (3.36) and (3.37) into (3.39) in order to express $\hat{W}$ and $\hat{P}$ as functions of $r$. In this diagram, $\bar{r}$ and $r'$ have the same meanings as in Figure 3.6, that is, they represent the solutions for the profit rate in the neo-Marxian and neo-Keynesian (Robinsonian) models, respectively. Thus, $\hat{P}$ has an intercept of $r'$ on the $r$ axis, because there is only positive price inflation for profit rates below $r'$, while $\hat{W}$ has an intercept of $\bar{r}$ because there is only positive wage inflation for profit rates above $\bar{r}$ (at which $w < \bar{w}$).

The solution to the synthesis model starts with the equilibrium condition (3.40), shown in the left-hand quadrant of Figure 3.7 as the point where $\hat{P} = \hat{W}$. Making all necessary substitutions and rearranging, the equilibrium profit and inflation rates are

$$r^* = \frac{1}{\Lambda}\left[\Phi f_0 + \Omega\left(\frac{1}{a_0} - \bar{w}\right)\right] \tag{3.41}$$

$$\hat{P}^* = \hat{W}^* = \frac{\Phi\Omega}{\Lambda}\left[f_0\left(\frac{a_1}{a_0}\right) - (s_r - f_1)\left(\frac{1}{a_0} - \bar{w}\right)\right] \tag{3.42}$$

where $\Lambda = \Phi(s_r - f_1) + \Omega(a_1/a_0) > 0$ (recall that $s_r - f_1 > 0$ by the neo-Robinsonian stability condition, and note that $(1/a_0) - \bar{w} > 0$ assuming that the workers' target real wage is less than the maximum real wage given by the productivity of labour). The equilibrium solutions for the real wage $w^*$ and growth rate $g^*$ are then obtained using equations (2.4) and (3.36), respectively. Note that the solution for $g^*$ lies on the realized-investment-equals-saving line $g = s_r r$, not on the desired investment function $(g^d)$; essentially, this means that desired investment is not brought into alignment with equilibrium investment even in the long run in Marglin's model (as shown by the gap between $g''$ and $g^*$ in Figure 3.7).

In this long-run solution, both social classes are partially frustrated in the pursuit of their respective economic targets. On the one hand, the capitalists (firms) are constrained to having their capital grow at the rate $g^*$ made feasible by available saving, and cannot achieve the desired growth rate $g'$ (which they could reach if the real wage was perfectly flexible, as in the neo-Robinsonian model), or even the lower desired growth rate $g''$ (at the equilibrium profit rate $r^*$ in the hybrid model). Saving-equals-investment equilibrium is achieved only because investment is constrained to equal saving, while the firms' excess desired investment (relative to the supply of saving) drives inflation to be positive. On the other hand, the workers cannot attain their desired or target real wage of $\overline{w}$ (which would require the profit rate to be $\overline{r}$) as their ambitions are also partially frustrated by the inflation that results when $r^* > \overline{r}$.

The solutions (3.41) and (3.42) imply some important comparative dynamic results. A rise in the workers' target for the real wage $\overline{w}$ leads to a reduction in the equilibrium profit rate $r^*$ and equilibrium growth rate $g^*$, a rise in the equilibrium real wage $w^*$ and an increase in equilibrium inflation $\hat{P}^* = \hat{W}^*$. Graphically, this would occur because the $\hat{W}$ line would shift down and to the left in Figure 3.7. Thus, the model behaves in a classical-Marxian fashion for an increase in the target real wage, except for the neo-Keynesian element of increased inflation. An increase in firms' animal spirits (a rise in $f_0$) would also lead to higher equilibrium inflation, but the equilibrium profit and growth rates would rise and the equilibrium real wage would fall, which is more of a neo-Keynesian outcome (essentially, the widow's cruse continues to refill itself). Graphically, this would occur because $g^d$ would shift to the right, leading to an increase in $r'$ and an upward shift in $\hat{P}$.

Finally, Marglin's synthesis model implies what he calls an 'ambiguous paradox of thrift' (1984a, p. 137): an increase in the saving propensity $s_r$ definitely lowers the equilibrium profit rate $r^*$ and inflation rate $\hat{P}^* = \hat{W}^*$, as can be seen from equations (3.41) and (3.42), but has an ambiguous impact on the equilibrium growth rate since $s_r$ rises while $r^*$ falls in $g^* = s_r r^*$. Graphically, the $g = s_r r$ line rotates down and to the right while $\hat{P}$ shifts downward in Figure 3.7; as a result, $g^*$ can end up either higher or lower. This ambiguity arises from the mix of classical-Marxian and neo-Robinsonian elements in the model: on the one hand, savings are higher for any given profit rate and growth is saving constrained, but on the other hand, a higher saving propensity also depresses the realized profit rate, so that the net impact on realized saving (and hence on equilibrium growth) is ambiguous.

The Marglin model has the advantage of making explicit some of the arguments about inflation that, as discussed above, were only implicit (or discussed in qualitative terms) in much of the earlier neo-Keynesian literature (by Kaldor, Robinson and others). However, there are several problems with Marglin's synthesis. First, inflation in his model is *always* driven by excess demand in the goods market; inflation cannot occur unless there is a persistent excess of desired investment over realized saving. In other words, inflation in his hybrid model is always 'demand-pull' rather than 'cost-push', to use the terminology of mainstream Keynesian macroeconomics in the 1960s and 1970s. Inflation due to firms passing through increases in their direct costs (unit labour costs or raw materials costs) cannot appear in Marglin's model; at best, it is a model of inflation in an 'overheated' economy. We will study more general models of 'conflicting claims' of workers and firms in which cost-push inflation can occur (but aggregate demand pressures are also incorporated) in Chapter 5.

Second, it is not clear why firms would continue to desire to invest at a rate that they never achieve in a 'long-run' equilibrium; it would make more sense that firms would eventually ratchet their expectations downward until they desire to invest at the rate made feasible by available saving. However, in this case, a plausible adjustment mechanism suggests that the inflation would eventually disappear and the Marglin synthesis model would degenerate into the neo-Marxian case with a long-run solution of $(\overline{w}, \bar{r}, \bar{g})$. To see this, suppose that the shift factor $f_0$ in the investment function falls when $g^d = f_0 + f_1 r^* > s_r r^*$. Then the $g^d$ line would shift down over time, causing $r^*$ to fall, which in turn would make $f_0$ decrease further, until eventually $r^* = \bar{r}$.

Third, and perhaps most importantly, this supposedly 'neo-Keynesian' synthesis allows at most a very limited role for aggregate demand in the long run (Nell, 1985; Dutt, 1987). The widow's cruse result does imply that greater investment demand would be expansionary in the long run, albeit at the expense of higher inflation. However, the model does not allow for increased wages to be expansionary via their impact on consumer demand. Since all wages are spent while part of profits is saved, a redistribution of income towards wages is sure to boost consumer spending. In principle, such a boost to consumption not only could offset any possible negative effects of lower profitability on investment, but also could possibly stimulate investment through strong accelerator effects (as larger consumer purchases would require increased output, which in turn would induce firms to invest in more capital equipment regardless of a lower profit share). As a result of certain restrictive assumptions, however, Marglin's model does not allow such a boost to consumption to be the dominant effect of a wage increase; his model only permits what might be called 'profit-led' rather than 'wage-led' growth.[41]

In order to see the possibility of a wage-led growth regime, one approach is to relax the assumption that capacity utilization is fixed at a 'full' or 'normal' rate and allow for the utilization rate $u$ to be an endogenous variable. In addition, we will need to consider various alternatives to the neo-Robinsonian investment function (equation 3.28 or 3.31), which can potentially allow for stronger accelerator effects. Neo-Kaleckian models that allow for variable capacity utilization will be covered in Chapter 4, while recent efforts at reviving the modelling approach of Harrod (with goods market instability bounded by ceilings and floors) and debates over the validity of allowing utilization to vary in the long run will be covered in Chapter 6. The supermultiplier approach, in which distribution and demand can affect the level of output but not the utilization or growth rate in the long run, will be addressed in Chapter 7.

But, the basic point is this: in all of the theories (except Harrod's) considered in this chapter and the previous one, output can change over time only as capital grows more rapidly or more slowly, since output is rigidly tied to the capital stock by the relation $Y = K/a_1$ assuming $u = u_n = 1$. In the neo-Kaleckian models covered in Chapter 4 and some of the other alternative models covered in Chapters 5–7, output can be determined by demand-side forces, and in some of these models the growth of the capital stock and potential (full-capacity) output adjust to actual output growth in the long run, rather than the other way around. Moreover, the 'normal' rate of capacity utilization itself can become an endogenous variable, as we shall see in Chapter 6.

## 3.6 Conclusions

Robinson (1956, 1962) emphatically did *not* believe that her basic model depicted a long-run, steady-state equilibrium that would be likely to persist forever in the absence of exogenous shocks or external disturbances. For Robinson, the model presented in section 3.4 above was only a starting point for an analysis that would take other forces into account, including endogenous feedbacks of the growth rate and income distribution determined in this model on the future evolution of the economy. Robinson encapsulated this approach in her concept of a 'golden age' equilibrium and her analysis of all the reasons why it might not be sustained over time. Thus, her concept of a bastard golden age, or the real wage resistance case, was just one of the many alternative scenarios that she outlined (but which, in many cases, she left for future scholars to investigate in greater depth). In these conclusions, we will address a few of the other concerns that Robinson identified and which later economists have also noted, with some indications of where they are addressed either later in this book or in other literature (in addition to the real wage resistance issue, covered in the previous section).

One point that should be obvious after our discussion of the classicals, Marx, Harrod and the early Kaldor is that the basic neo-Robinsonian model does not address the first Harrod problem: it does not take into account the need for long-run growth (the equilibrium rate of capital accumulation $g^*$) to match the natural rate of growth – the sum of the growth rates of labour supply and labour productivity, $g_N = n + q$ (which can also be thought of as the growth rate of the effective labour supply, taking labour-saving technical progress into account). If the equilibrium growth rate based on saving and investment behaviour exceeds the natural rate ($g^* > n + q$), the economy would eventually run out of workers unless it could relieve this constraint, for example by allowing more immigration, hiring foreign guest workers or modifying gender relations (all of which relate to how elastic $n$ is), or else by accelerating labour-saving technological progress (increasing $q$). If the effective labour supply cannot be increased sufficiently, the real wage would rise instead of falling, and the economy would be unable to grow at the rate $g^*$ (as Robinson recognized in her case of a 'restrained golden age').

In the contrary case in which $g^* < n + q$, the unemployment rate would be chronically increasing, and various adjustment mechanisms could be set into motion. For example, the real wage might fall, possibly leading to a reduction in population growth (à la Malthus) or labour force participation in the modern capitalist sector (à la Marx, Lewis or Cornwall), in which case $n$ might fall (as envisioned in Robinson's 'leaden age'). Alternatively, the low rate of capital accumulation could lead to a slowdown in the rate of technological innovation (especially, the lower real wage would lessen incentives for labour-saving innovations), in which case $q$ could be reduced. Overall, it is important to understand the adjustment mechanisms in regard to labour supply, wage-setting and technological progress that might occur in response to a disequilibrium between the growth rate implied by saving–investment behaviour and the rate necessary to keep pace with the growth of the effective labour force and maintain a constant unemployment rate. Models that address some of these issues – particularly inflationary outcomes and endogenous innovation – are covered later in this book, especially in Chapters 5 and 8.

Another deficiency in all of the neo-Keynesian models reviewed so far is that they tend to ignore the crucial role of money and financial institutions in financing investment.[42] Because an endogenous increase in saving 'finances' increased investment only in an *ex post* sense, the finance required to pay for the investment expenditures *ex ante* must be obtained *before* those additional savings (and the increased profits that generate them) are available. Firms may be able to finance their investment internally through the use of retained profits from previous sales to some extent, but if these internal funds are not sufficient to pay for the firms' upfront investment costs, then they need to obtain some form of external finance (loans from banks, or funds obtained by selling bonds or issuing new equity shares). New loans create bank deposits that constitute increased 'money supply', while sales of bonds or issues of new equity do not increase money but can have various other repercussions on financial markets. Similarly, households – who were largely ignored in the early neo-Keynesian literature – may borrow either to maintain a desired standard of living in the face of falling income or to cover luxury consumption (sometimes based on Veblenian emulation or 'keeping up with the Joneses'). We will cover some models that incorporate financial relationships and debt in Chapter 7, but a full treatment of endogenous money and finance is beyond the scope of this book.[43]

Last but not least, the entire neo-Keynesian literature covered in this chapter only allows for a limited influence of aggregate demand on the long-run evolution of the economy. For example, in Robinson's model, since workers always consume 100 per cent of their wage income, the total level of aggregate demand effectively depends on the two-sided aspects of capitalists' spending: on the one hand, their expenditures on luxury consumption (what they do with the part of their profit earnings that they don't save), and on the other hand, their willingness as owners of firms to spend on investment in new capital stock. These spending (or saving and investing) decisions of the capitalists uniquely determine the rates of profits and growth, as well as the real wage (which always varies inversely with the profit rate, for a given technology). There is no room for workers' consumption to affect the equilibrium of the economy, in which respect the neo-Robinsonian model is similar to the Pasinetti and neo-Pasinetti theorems discussed earlier. Moreover, there is no room in the neo-Robinsonian model for output to vary relative to the existing capital stock, or in other words, excess capacity (underutilized capital) cannot emerge, no matter how low demand is or how far the growth rate falls. Output is always tied to the existing capital stock in the fixed proportion of one unit of output for every $a_1$ units of capital, and prices do the adjusting instead of quantities. This aspect of the neo-Robinsonian model, especially as interpreted by Marglin (1984a, 1984b), was criticized by Nell (1985) and Dutt (1987), and arguably is not faithful to the original intent of Robinson (1962). The neo-Kaleckian models developed in the next chapter (and other models covered in subsequent chapters) are direct responses to this surprising lack of attention (or incomplete attention) to demand in the early neo-Keynesian growth models.

## ? STUDY QUESTIONS

1) Explain Harrod's three growth rate concepts and the distinction between the first and second 'Harrod problems'. Which problems relate to which growth rates, and what dilemmas does each problem suggest?
2) How do the two Harrod problems interact? What is the likely dynamic outcome if the warranted rate of growth exceeds the natural rate? What are the likely dynamics in the opposite case?
3) Many development economists have recommended a strategy of increasing national saving rates in order to accelerate economic growth. In Harrod's theory, which growth rate is increased by raising a country's propensity to save, and is that a good policy to pursue? Why or why not?
4) Which of the two Harrod problems are addressed (or not addressed) in the contributions of Kaldor and Robinson?
5) Draw a diagram similar to Figure 3.3 for the early Kaldor model of growth and distribution. Which version of a classical-Marxian model from Chapter 2 does it most resemble, and in what respects is it similar or different?
6) Demonstrate the 'paradox of thrift' and 'widow's cruse' implications of the neo-Robinsonian growth model. What is the *policy* significance of these results?
7) How does Marglin synthesize neo-Marxian and neo-Keynesian principles in his hybrid growth model? In what respects is the hybrid model more classical-Marxian and in what respects is it more Keynesian? What role, if any, does aggregate demand play in Marglin's synthesis?

## NOTES

1 For an account of Keynes's contributions relative to his predecessors as well as later developments in other schools of (mostly mainstream) macroeconomics, see Snowdon and Vane (2005).
2 Harrod (1939) is the classic statement of Harrod's thinking, but strictly speaking his ideas pre-date the publication by Keynes of *The General Theory*. See Harrod (1936) and Asimakpoulos (1991, Chapter 7) for discussion.
3 For later presentations and interpretations of Sraffa's contributions see, among others, Pasinetti (1977), Steedman (1977), Harris (1978), Eatwell and Milgate (1983), Bharadwaj and Schefold (1990), Kurz and Salvadori (2003), Kurz (2008), Roncaglia (2009) and Sinha (2016). Note that Sraffian theory extends beyond value theory and includes analysis of growth and distribution. A modern example of this – the Sraffian supermultiplier theory originally due to Serrano (1995) and subsequently developed by Bortis (1997), Cesaratto (2015), Freitas and Serrano (2015), Pariboni (2016), and Serrano and Freitas (2017) – will be explored in detail in Chapter 7.
4 Ultimately, we will introduce a fourth growth rate – the *expected* rate of growth – to help us understand the interaction of the actual and warranted rates in the course of what is identified below as the second Harrod problem.

**5** See also Sen (1970, Introduction), Fazzari (1985) and Palley (1996a).

**6** Equation (3.11) is originally stated in Sen (1970, Introduction) as:

$$y = 1 - \frac{1 - y^e}{y^e} \cdot y_w$$

**7** Although we have only demonstrated instability in the case where $y^e > y_w$ initially, the second Harrod problem can also be shown to arise if $y^e < y_w$ initially. Demonstration of this second case is left to the reader.

**8** See, for example, Hicks (1950) and Minsky (1959), who also drew upon the earlier work of Samuelson (1939a, 1939b) on multiplier–accelerator interactions.

**9** We will return to investigate this sort of behaviour in detail in Chapter 6, in the course of discussing the controversy over Harrodian instability in contemporary heterodox growth models.

**10** See also Franke (2015), who reaches a similar conclusion by considering the effects on Harrodian instability of modelling in discrete rather than continuous time. The latter, of course, makes no allowance for *lags* (since change is conceived as instantaneous) which, as the analysis above reveals, are essential in the behavioural equations of the Harrod model if the second Harrod problem is to materialize.

**11** As Shaikh (2016, p. 14) notes, 'antagonistic in nature and turbulent in operation . . . [real competition] is as different from so-called perfect competition as war is from ballet.' This vision of competition is compatible with the sometimes-heard quip that 'competition is a game to see who gets to be the monopolist'.

**12** From a Harrodian perspective this must be the case, because the objective of firms when investing in accordance with the accelerator relationship is to expand capacity in line with the expansion of the economy and so keep capacity utilization constant at its normal rate. Were $u \neq u_n$ when $y = y_w$, the warranted rate would not be an equilibrium, because firms would seek to adjust their investment behaviour in response to $u \neq u_n$ in an effort to change the utilization rate (specifically, restore it to its normal value).

**13** We will see other examples of 'first-generation' neo-Keynesian growth models that exhibit long-run, dynamic paradoxes of thrift later in this chapter; the paradox of thrift is also found in the neo-Kaleckian models covered in Chapter 4.

**14** Note that this variability owes to the law of diminishing marginal returns. As remarked in Chapter 1, this is precisely the feature of production that must be overcome (in the aggregate) in order to generate neoclassical endogenous growth results. In this way, neoclassical endogenous growth theory is sometimes said to have 'rediscovered' the Harrodian fixity of the marginal capital–output ratio, which the Solow model deliberately sought to overcome.

**15** For example, the neoclassical production function treats capital as 'putty-like' – an input that can be moulded and remoulded to suit any production process and any capital–labour ratio, and be made consistent with any change in the state of technology. In reality, it might be argued, capital is 'lumpy' and embodies specific engineering standards germane to a specific state of technology, that circumscribe both its capacity to produce output and the extent to which it can be combined in a production process with different quantities of labour. This is the essence of the Leontief technology. See, for example, Harcourt (1972) for further discussion.

**16** Solow himself must be considered the architect of this misleading argument. As early as the first page of his article he claims that 'the fundamental opposition of the warranted and natural rates turns out in the end to flow from the crucial assumption that production takes place under conditions of *fixed proportions*' (Solow, 1956, p. 65; emphasis in original).

**17** Referring again to Solow (1956), we are told in the first paragraph of section II ('A Model of Long-Run Growth') that:

> Part of each instant's output is consumed and the rest is saved and invested. The community's stock of capital $K(t)$ takes the form of an accumulation of the composite commodity. Net investment is then just the rate of increase of this capital stock $dK/dt$ or $\dot{K}$, so we have the basic identity at every instant in time:
>
> (1) $$\dot{K} = sY$$
>
> (Solow, 1956, p. 66)

While this stops short of literally saying 'saving causes investment' it is hard to read it in any other way, especially as the article appeals to no other accumulation function than (1) above, which, read conventionally, states that saving is the independent variable and investment the dependent variable. It is also hard to believe that Solow did not know what he was doing here – namely, ruling out by hypothesis any possibility of Keynesian problems of deficient demand of the kind entertained by Harrod in his formulation of the natural and warranted rates of growth.

**18** See also Kregel (1980), Asimakopulos (1991, pp. 161–4) and Halsmayer and Hoover (2016).

**19** Such an analysis simply assumes away the second Harrod problem of instability of the warranted growth rate, which by that point (the 1950s) was an issue that had been relegated to models of short-run cyclical fluctuations and divorced from long-run growth analysis as noted earlier.

**20** Although Kaldor emphasized the solution of his model for the profit share, it is easy to see that the real wage and profit rate are also endogenous variables in this framework, and the reader is invited to solve for the corresponding equilibrium values $w^*$ and $r^*$ as an exercise. The neo-Robinsonian model, developed in the next section, puts more emphasis on the latter two variables.

**21** Of course, there may be narrower limits to the profit share. For example, it cannot be so high as to imply a wage less than some minimum subsistence wage for workers, or so low as to induce capitalists to cease saving and investing. But we restrict ourselves here to the formal requirements of the early Kaldor model.

**22** Recall that this distinction does not matter in a long-run, steady-state equilibrium in which the capital–output ratio $a_1$ is constant, since constancy of $a_1$ implies that $y = g$.

**23** The one exception is the natural rate of growth closure (which assumes full employment or a constant unemployment rate) in the classical/Marxian models, but as discussed in Chapter 2 that closure was only implicit in the original works of the classicals and our interpretation of it was inspired by Kaldor (1955–56). To see the analogy between that closure and Kaldor's model more clearly, recall that the profit share can be written as $\pi = 1 - wa_0$. Thus, given labour productivity $Q = 1/a_0$, there is a monotonic inverse relationship between the real wage and the profit share. Furthermore, the profit rate can be written as $r = \pi / a_1$. So, given the input–output coefficients, the real wage and profit rate are determined once the profit share is known.

**24** The same intuition would apply if we allow for positive productivity growth $q = \hat{Q} = -\hat{a}_0 > 0$, since a higher $q$ would mean that labour was becoming redundant at a faster rate, thereby putting a downward pressure on wages.

**25** The same approach is taken in contemporary studies that attribute different saving propensities to different social classes. See, for example, Lavoie and Godley (2001) and Saez and Zucman (2016).

**26** In this respect we adopt Hein's (2008, 2014) notation, in which the ratio of saving to capital is represented by a different letter to the ratio of investment to capital. Earlier works, which used the same letter (usually $g$) to represent both ratios (sometimes with $s$ and $i$ superscripts) led to much confusion. Here, we use $g$ only for the ratio of investment to capital.

**27** Equation (3.27) is the same as equation (2.23) in our model of the Ricardian stationary state, except with $\sigma$ replacing $g$.

**28** If the capital–output ratio $a_1$ remains fixed, then $g$ is also the growth rate of output. However, if the capital–output ratio can vary (which it may as a result of technological change, variations in capacity utilization or changes in the sectoral composition of output), then the equivalence between the rates of growth of capital and output breaks down.

**29** One can hear echoes of Kalecki's (1937) 'principle of increasing risk' here, if only implicitly. In Kalecki's view, the increasing risks associated with higher levels of investment would lead to credit rationing or higher interest rates charged by lenders; here, the increasing risk is seen as diminishing the impact of greater expected profits on desired investment, although one could also think of financial constraints as being incorporated into the $f(r^e)$ function.

**30** Note that $g > \sigma \Rightarrow I/K > S/K \Rightarrow I > S$, which, under the hypothesized conditions, implies excess demand in the goods market.

**31** We have omitted the consumption–growth relationship from Chapter 2, because it is not essential to the key points in the neo-Keynesian framework, but it could easily be added into the analysis if desired.

**32** Similarly, substituting equation (3.34) into (2.7) would yield the equilibrium solution for consumption per worker.

**33** We will consider models of inflation and distribution that make different assumptions about wage- and price-setting in Chapter 5.

**34** These derivatives can be obtained directly by partially differentiating the reduced form solutions (3.33) to (3.35), or alternatively by total differentiation of the equilibrium condition (3.32) with the saving and investment functions (3.27) and (3.31) substituted in, under the assumption that $r = r^e$.

**35** Many translations of the Bible, including the canonical King James version, use the word 'vessel' rather than 'cruse', as used by Keynes.

**36** 'Normal' circumstances here exclude the special case of full wage indexing, in which workers' nominal wages are guaranteed to rise in proportion to prices of consumer goods, as has occurred in some high-inflation situations. See Chapter 5 for a model of inflation that incorporates partial or full indexation of nominal wages to price inflation.

**37** Robinson defined her 'bastard golden age' (and various other alternative 'ages') in relation to the benchmark of a pure 'golden age', in which the equilibrium from her core model, as depicted above, implies a growth rate equal to the natural rate of growth $(n + q)$ so that the growth path would be sustainable in terms of maintaining a constant rate of unemployment over time. See the conclusions in section 3.6 below for further discussion.

**38** For simplicity, we omit the consumption–growth inverse relation, equation (2.7), along with the variable $c$. Omitting this one equation along with one variable makes no difference to any of the results in this section.

**39** In his own discussion in the *General Theory*, Keynes (1936, p. 14) recognized that workers care about their relative wages compared to each other; in other passages, he seemed to imply that they would accept reductions in real wages if these were caused by a rise in the general price level and shared by all workers.

**40** Our presentation differs from Marglin's in several respects, although the logic is entirely the same. Among other things, he used changes in discrete time while we use instantaneous rates of change; he used a non-linear investment function (similar to our equation 3.28); and he included 100 per cent depreciation of capital in each period (whereas we assume no depreciation). Our presentation is more similar to the representation of Marglin's model in Dutt (1990).

**41** That is, lowering the target wage would actually stimulate growth in Marglin's model. However, in fairness to Marglin, most of the early neo-Keynesian models covered in this chapter, especially those of Kaldor (1955–56) and Robinson (1962), also imply that growth is normally profit-led, or at least that faster growth requires an accompanying increase in the profit rate and/or profit share. For modern presentations of the wage-led versus profit-led distinction, see Lavoie and Stockhammer (2013) as well as Chapters 4 and 5.

**42** Rochon (2005) shows that Robinson (1956) was aware of this problem and recognized the endogenous character of the money supply.

**43** The interested reader is referred to the post-Keynesian literature on money and finance, including Moore (1988), Wray (1990, 1998), Palley (1996c, 2013b), Rochon (1999) and Lavoie (2014), among many others. A good overview of many of the most important themes in this large literature can be found in Arestis and Sawyer (2007).

## Appendix 3.1 Pasinetti after Samuelson and Modigliani

To begin with, note that if $s_w < s_r$, then it follows that:

$$s_w Y = s_w(wL + \Pi_w + \Pi_c) < s_w(wL + \Pi_w) + s_r\Pi_c \qquad (3A.1)$$

where $\Pi_c$ and $\Pi_w$ are the profit incomes of capitalists and workers, respectively. Meanwhile, from the conditions imposed by Pasinetti (1962) – that $r_c = r_w = r$ and $g_c = g_w = g$ – it follows that:

$$\frac{r}{g} = \frac{r_c}{g_c} = \frac{r_w}{g_w}$$

$$\Rightarrow \frac{\Pi_c/K_c}{I_c/K_c} = \frac{\Pi_w/K_w}{I_w/K_w} \qquad (3A.2)$$

where $I_c = s_r\Pi_c$ and $I_w = s_w(wL + \Pi_w)$ are the flows of total investment spending ($I$) funded by capitalists' and workers' savings, respectively, and $K_w$ is the capital stock owned by workers. It follows that we can rewrite (3A.2) as:

$$\frac{\Pi_c}{s_r\Pi_c} = \frac{\Pi_w}{s_w(wL + \Pi_w)}$$

$$\Rightarrow s_w(wL + \Pi_w) = s_r\Pi_w \qquad (3A.3)$$

Finally, combining the results in (3A.1) and (3A.3) yields:

$$s_w Y < s_r(\Pi_w + \Pi_w) = s_r\Pi$$

$$\Rightarrow s_w < \pi s_r \qquad (3A.4)$$

as stated in the text.

## Appendix 3.2 A note on Robinson's original presentation

Robinson did not write down the equations for the saving and investment functions (3.27) and (3.28) as shown above, but rather represented them as curves on a diagram, and the way she discussed them was conceptually different from what we stated earlier. According to Robinson, what we are calling the saving function (the $\sigma = s_r r$ curve as shown in Figure 3.2) represented 'the expected rate of profit on investment as a function of the accumulation that generates it' (Robinson, 1962, p. 48). This requires inverting the saving function and thinking of it instead as determining profit expectations, given the growth (accumulation) rate and the saving propensity out of profits:

$$r^e = g/s_r \tag{3A.5}$$

This is essentially a multiplier formula showing how a given amount (rate) of investment in the numerator $g$ is blown up into a larger amount (rate) of profits $r$ when multiplied by one over the saving propensity out of profits (since $s_r < 1$, it follows that $1/s_r > 1$ and $r^e > g$). She also referred to what we have called the investment function (equation 3.28) as representing 'the rate of accumulation as a function of the rate of profit that induces it' (Robinson, 1962, p. 48). Assuming that she meant the *realized* rate of profit, this suggests that the investment function should be rewritten as $g = f(r)$ instead of $g = f(r^e)$. She also defined the growth rate at a stable equilibrium (such as point A in Figure 3.2) as 'the desired rate of accumulation', in the sense that this rate of accumulation 'is generating just the expectation of profit that is causing it to be maintained' (Robinson, 1962, p. 49).

This way of presenting the model has the advantage of emphasizing the two-sided nature of profits in the accumulation process: on the one hand, they are the incentive or inducement to invest, and on the other hand, they are the chief source of (*ex post*) finance for investment (what is more conventionally called 'saving'). Nevertheless, Robinson's presentation is confusing, because it puts the profit expectations on the saving side of the ledger, that is, it requires us to think of business executives as having expectations about the profits that will be generated out of their own collective investment spending, rather than about the profit rates that would induce them to carry out a certain level of investment. In fact, the profits that are generated depend on an economy-wide multiplier process, and may not be perceived by firms until they are realized or observed, so it may not make sense to put the expectations of profits into the saving (or profit generation) function. In terms of the previous literature, Harris (1978) covered Robinson's analysis essentially in its original form, but later authors (Marglin, 1984a, 1984b; Dutt, 1990; Foley and Michl, 1999) interpreted the model more as we have here, where saving comes out of realized profits and investment is a function of expected profits.

# 4 Neo-Kaleckian models

## 4.1 Introduction

The classical-Marxian and neo-Keynesian theories covered in the previous two chapters link economic growth to the process of capital accumulation and show how this process in turn is related to the functional (wage–profit) distribution of income. However, most of those theories have the rather disturbing implication that more rapid economic growth generally requires a more unequal distribution of income, in the sense of a higher share of national income going to profits and some degree of repression of wages. Those theories do allow a number of exceptions, especially in the presence of technological progress. As we saw, in many cases technical progress makes labour and/or capital more productive and thereby permits an economy to expand more rapidly without wages (either the real wage or wage share) inevitably having to fall. But the classical-Marxian theories do not incorporate aggregate demand into the analysis, while some of the early neo-Keynesian models do so but only in a limited way. For example, the neo-Robinsonian model privileges investment and saving decisions by capitalists, but gives no role to workers' consumption demand. In order to find theories that take aggregate demand more systematically into account and allow for the possibility that economic systems can be wage-led instead of profit-led, we begin by exploring models in the neo-Kaleckian tradition.

Michał Kalecki was a Polish economist who was self-trained in Marxian economics but did not use an orthodox Marxian analytical apparatus (for example, he did not adopt the labour theory of value). In a series of articles and books he wrote between the 1930s and 1950s (collected in Kalecki, 1990, 1991), Kalecki independently developed a macroeconomic framework that has strong analytical similarities to Keynes's income–expenditure model. Kalecki demonstrated that increases in investment spending, net exports or the government deficit could boost the level of realized profits and the rate of capacity utilization (the ratio of actual output relative to potential output) in the private sector through a multiplier mechanism, in an economy characterized by oligopolistic industries that operate with excess capacity. Unlike Keynes, Kalecki focused on income distribution as well as on total output and employment, and he also grounded his macro model in a micro-level analysis of imperfect competition. Kalecki thus founded a distinctive approach to macro modelling, in which the relative shares of wages and profits in national income depend on the markup pricing behaviour of oligopolistic firms, and those shares in turn have important causal effects on aggregate demand and economic growth.

When Kalecki moved to the Oxford Institute of Statistics in the UK to escape the Nazi occupation of Poland during World War II, he collaborated with Josef Steindl, an Austrian economist (in the sense of Austrian nationality, not 'Austrian' economic theory) who was also working in exile at Oxford at that time. Steindl (1952 [1976]) made fundamental contributions to the development of Kalecki's approach, especially in regard to the process of industrial concentration, the importance of investment finance and the potential for economic stagnation in a highly concentrated and unequal regime. This chapter will present 'neo-Kaleckian' models in the tradition of both Kalecki and Steindl, as well as later extensions of those models; additional critiques, extensions and alternatives are covered in the following three chapters.

The rest of this chapter is organized as follows. Section 4.2 takes a brief but vital detour from macro theory into the microeconomic foundations of the neo-Kaleckian approach. Section 4.3 presents a 'canonical' version of the Kalecki–Steindl macro model based on the neo-Kaleckian literature of the 1970s and 1980s. Section 4.4 discusses several important extensions and critiques of the Kalecki–Steindl approach, which fall broadly within a neo-Kaleckian paradigm; some of these (especially the approach of Bhaduri and Marglin, 1990) revive significant elements from the classical-Marxian and neo-Keynesian traditions, while others (derived from Blecker, 1989a) open up the model to international trade. These alternative models yield a variety of different cases in which aggregate demand and economic growth can be either wage-led or profit-led, depending on various circumstances. The concluding section 4.5 discusses other critiques and responses. Appendix 4.1 presents an important variant of the neo-Kaleckian model that emphasizes the role of overhead labour in explaining short-run, cyclical variations in labour productivity and the profit share.

## 4.2 Kaleckian microfoundations: theory of the oligopolistic firm

To understand Kalecki's macro theories, it is important to start with a brief account of his underlying vision of the pricing mechanism and the operation of business firms at the micro level.[1] His analysis was built upon two key elements that are not found in the neo-Keynesian or classical-Marxian theories covered in the previous two chapters: the importance of 'markup' pricing by oligopolistic firms and the existence of excess capacity in industry. Kalecki assumed that most industries are dominated by oligopolistic firms that have significant market power, which enables them to mark up their prices above their marginal costs. He then rooted his macro-level theory of income distribution (relative shares of wages and profits in national income) in the idea that the profit share in the aggregate economy was fundamentally determined by the average markup rate of the firms.

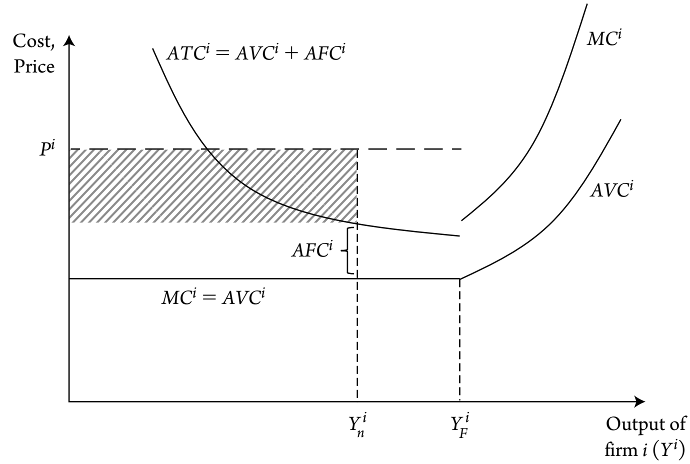
Like Ricardo (1821 [1951]), Kalecki (1954 [1968]) drew a sharp distinction between primary commodities, which were produced in conditions of increasing marginal costs and sold at prices that reflect demand as well as supply, and manufactured goods, which were produced in conditions of constant costs and sold at cost-based prices. Unlike Ricardo, however, Kalecki did not believe that the forces of competition were strong enough to impose a 'normal' rate of profit on all sectors of the economy. Instead, Kalecki saw that by the twentieth century, industrial concentration had proceeded to the point where most industries had an oligopolistic structure, in which the leading firms could administer prices based on a cost-plus-markup principle. Thus, the theory of the firm that we will describe below is meant to be applied only to industrial producers, not to producers of agricultural or mineral products for whom standard considerations of supply and demand (augmented to include possible speculative activity) would apply.

## 4.2.1 Excess capacity and cost functions for industrial firms

Kalecki's theory of the industrial firm is based on the type of cost functions shown in Figure 4.1. Although these differ from the U-shaped cost functions usually taught in neoclassical microeconomics courses (especially at the undergraduate level), they are generally consistent with the cost functions used in many advanced branches of mainstream economics today, such as the theories of international trade with increasing returns to scale originally developed by Krugman (1979). In this diagram, the output of an individual firm $i$ is represented by $Y^i$. Given its capacity, defined by the existing technology, capital stock and social relationships (for example, standard weekly hours and intensity of work effort) at the plant level, each firm has a maximum level of engineer-rated capacity $Y^i_F$, which we will refer to here as output at 'full' utilization.[2] Note that this is not the same as 'normal' utilization, which we will define below.

Note: The shaded area represents net profits at normal utilization.

**Figure 4.1** The Kaleckian model of the firm: cost functions and capacity utilization




At output levels below full utilization $(Y^i < Y^i_F)$, a typical firm $i$ produces with constant marginal cost $(MC^i)$, implying also constant average variable cost $(AVC^i)$.[3] That is, in order to change the level of output at any point below full utilization, the firm simply needs to change the variable inputs (raw materials, intermediate goods and labour hours of production workers) in the same proportion as it wishes to increase or decrease the output. However, each firm also has various fixed costs $(FC^i)$, including overhead labour (managers, administrators, engineers and so on), research and development (R&D) expenses, advertising and other sales effort, debt service and so on, all of which are not proportional to the current level of output $(Y^i)$. Therefore, average fixed costs $(AFC^i = FC^i/Y^i)$ are decreasing in the level of output up to full utilization $(Y^i_F)$ and average total costs $(ATC^i = AVC^i + AFC^i)$ are also decreasing up to that point. Mathematically,

$$ATC^i = AVC^i + AFC^i = MC^i + (FC^i/Y^i) \tag{4.1}$$

where $MC^i$ is constant as long as $Y^i < Y^i_F$.

In contrast, if output rises beyond full utilization, then $MC^i$ and $AVC^i$ start increasing, and of course $MC^i > AVC^i$ as always occurs when costs are rising. At such high levels of output, the firm's productive capabilities would be strained – machines and other systems would suffer more breakdowns and need more frequent repairs, workers might have to be paid overtime or less efficient workers would have to be hired and trained, administrative capacity would also be stretched, normal supplies of inputs (raw materials or intermediate goods) in 'value chains' would become harder to obtain, and so on. It is not impossible to increase output beyond a standard definition of full utilization – many industrial firms have done so in times of war or other boom conditions – but it can only be accomplished at increasing cost, and it is not a situation that most firms would want to allow to persist long term.

Although firms could possibly produce on the upward-sloping, above-full-utilization portions of their cost curves, under normal conditions they try to avoid this. For several reasons that are well known in the field of industrial organization, firms typically produce at output levels significantly below full utilization. The main reasons why firms prefer to operate with some degree of excess capacity include the following:

- *Indivisibilities*: Many types of machinery and equipment can only be obtained in large or discontinuous units, but can be operated at less than 100 per cent of their potential flow-through. Similarly, it may be economical to obtain intellectual property products (software, patent rights and so on) on scales that may exceed a firm's current usage needs. In such cases, firms may be forced to invest in capital (physical or intellectual) in 'lumpy' amounts that generate excess capacity relative to current output levels.
- *Building ahead of demand*: As emphasized by Keynes (1936), firms face fundamental uncertainty about the future demand for their products, in the sense that they do not even know the probability distributions of possible future market situations. Under such circumstances, firms desire to maintain some excess capacity so that they can be positioned to respond to future increases in demand (which are uncertain and unknowable in advance) by ramping up output and selling more, instead of having to concede market share to rival firms. Firms don't want to disappoint potential customers, so if they can plan 'ahead of the curve' by installing some level of excess capacity, they won't have to turn away customers when new ones show up or existing ones increase their orders.
- *Entry deterrence*: Large firms in oligopolistic industries want to prevent the entry of new firms (or the expansion of existing rivals), which would lessen the concentration of the industry and could thereby reduce their market power and cut their profit margins. Having excess capacity on hand constitutes a deterrent to such entry (or expansion) because it enables a given existing firm to increase output and potentially to undersell rivals or new entrants as needed to prevent them from taking away its market share. Of course, this entry deterrence motivation is related to building ahead of demand, but it can also be thought of as a separate reason for maintaining spare capacity.

For all these reasons, we can say that firms' *normal* capacity utilization lies somewhere *below* full utilization, as defined earlier. The normal utilization rate could occur at a particular level of output that a firm seeks to achieve, as illustrated by $Y^i_n$ in Figure 4.1, or it could be a range of levels that a firm is willing to accept, as will be discussed in Chapter 6. But whether it is a fixed level or a range, the normal utilization rate (or range of rates) is definitely less than 100 per cent of capacity; here, we will use a particular level for simplicity. If a firm finds that it needs to operate close to or above full utilization for a sustained period of time, this would be a signal to invest in increased capacity – a point that should be kept in mind when we come to the neo-Kaleckian investment functions discussed later in this chapter. For now, we may simply note that such investment could take the form of either a higher level of potential output $Y^i_F$ in a given production unit, or the construction of new units that would add to the firm's overall capacity. Either way, the upshot is that oligopolistic firms normally operate in a zone where $MC^i = AVC^i$ is constant and $ATC^i$ is decreasing; as a result, they have to be able to set prices above $ATC^i$ (at output levels in the normal range of utilization) in order to make positive net profits when all costs are considered.

### 4.2.2 Prices and profits (markups, margins and shares)

Next, we have to specify how firms set their prices ($P^i$). Since $ATC^i > MC^i$, firms would make losses if they set prices by the neoclassical 'perfect competition' rule, $P^i = MC^i$. Thus, as recognized in all theories of imperfect competition, firms that face decreasing average total costs must set prices by some rule or procedure that ensures $P^i \geq ATC^i$ so that they can make a non-negative net profit (and firms that are unable to do so on a sustained basis will be forced to exit the market). The simplest such rule, often invoked by Kalecki, is that firms set prices by a markup on average variable cost:

$$P^i = (1 + \tau^i)AVC^i \tag{4.2}$$

where $\tau^i > 0$ is the markup rate, which must be not only positive but also high enough so that $P^i > ATC^i$ in order for the firm to obtain a positive net profit. Or, to put it another way, $\tau^i$ is a gross markup, which means that it must cover fixed costs as well as provide net profits to the firm. Indeed, Kalecki included 'overheads' or fixed costs in his list of factors that affect markups, as we will discuss below.

As Figure 4.1 makes clear, whether the price that the firm sets is high enough to generate a net profit depends on the level of output or (equivalently) the utilization rate. The need to take the level of output into account in setting prices has led to the hypothesis of 'normal-cost pricing' (Andrews, 1949), which can be considered an alternative to simple markup pricing.[4] In this view, firms calculate what their average total costs (including fixed costs) would be at their normal rate of capacity utilization, which as noted earlier is less than full utilization. Referring to the output level of the *i*th firm at normal utilization as $Y^i_n$, we can define normal unit costs (that is, average total costs at normal output) as $NUC^i = MC^i + FC^i/Y^i_n$, and then prices are set by charging a 'net costing margin' ($ncm^i$) on $NUC^i$ rather than a gross markup on $AVC^i$. Formally,

$$P^i = (1 + ncm^i)NUC^i = (1 + ncm^i)(MC^i + FC^i/Y^i_n) \qquad (4.3)$$

If firms follow normal-cost pricing, there is the additional question of how they determine the net costing margin $ncm^i$. One important hypothesis in the industrial organization literature is that firms set $ncm^i$ in order to achieve a target rate of return (profit rate) on their capital, calculated at the normal level of output $Y^i_n$, which leads to the idea of 'target-return pricing'. Suppose the firm's capital stock is $K^i$, the price of capital goods is $P_K$ (there is no $i$ superscript assuming this is common to all firms), and a normal rate of return (as dictated, for example, by financial markets in order to satisfy stockholders and bondholders) is $r_n$ (also assumed to be the same for all firms). Then the profits that the firm needs to achieve are $r_n P_K K^i$. If we assume that the ratio of capital to normal output for the *i*th firm is given by $a^i_1 = K^i/Y^i_n$, then the firm's required profits can be written as $r_n a^i_1 P_K Y^i_n$. Therefore, the net costing margin must be set such that $ncm^i \cdot NUC^i \cdot Y^i_n = r_n a^i_1 P_K Y^i_n$, which implies

$$ncm^i = r_n a^i_1 P_K/NUC^i = r_n a^i_1 P_K/(MC^i + FC^i/Y^i_n) \qquad (4.4)$$

Then, the solution (4.4) for $ncm^i$ is substituted into the normal-cost pricing equation (4.3) to obtain the target-return pricing relationship

$$P^i = NUC^i + r_n a^i_1 P_K = MC^i + (FC^i/Y^i_n) + r_n a^i_1 P_K \qquad (4.5)$$

The normal-cost and target-return pricing hypotheses have the advantage that they take fixed costs and expected or normal levels of utilization explicitly into account. However, Kalecki himself often took a simpler view, in which he theorized more informally about the determinants of markups. Kalecki (1954 [1968]) used the concept of the 'degree of monopoly', which he measured by the ratio of gross profits to net sales (or gross profit margin, $GPM^i$):[5]

$$GPM^i = \tau^i AVC^i/P^i = \tau^i AVC^i/[(1 + \tau^i)AVC^i] = \tau^i/(1 + \tau^i) \qquad (4.6)$$

Since $\partial GPM^i/\partial\tau^i > 0$, $GPM^i$ is monotonically increasing in $\tau^i$ and there is no operational difference between the so-called degree of monopoly and the markup rate.[6] Hence, the factors that affect the degree of monopoly can also be thought of as the determinants of markups.

Kalecki advanced several hypotheses about the factors that could cause changes in the degree of monopoly, or, equivalently, changes in average markups. Kalecki (1954 [1968]) spoke of four such factors, and we can add a fifth one that derives from his later work (especially Kalecki, 1971a) and is more explicit in the later neo-Kaleckian literature. The five factors, with the signs showing the direction of their impact on $\tau^i$ (or $GPM^i$), are:

1) *Industrial concentration* (+): The biggest firms in more concentrated industries have a greater ability to set prices and achieve high profit margins because of their lack of effective competition, and the potential for collusion – which may be either explicit (if legally allowed in a particular country or industry) or implicit (for example, via price leadership). As the industrial organization literature has long verified, more concentrated industries tend to have higher average profit margins, after controlling for other factors.
2) *'Overheads' or fixed costs* (+): As noted earlier, these include large-scale machinery and equipment, fees for intellectual property rights (patents, licensing and so on), management and maintenance (overhead labour), R&D expenses and costs of debt service (interest plus principal on outstanding debt). These expenses are explicitly taken into account in normal-cost or target-return pricing, but must also be taken into account in markup pricing to ensure that $P^i > ATC^i$ at the firm's expected output level. Indeed, if firms take the normal rate of capacity utilization into account in deciding how to account for their fixed costs, they are effectively doing normal-cost pricing, and if they incorporate a normal rate of return into their calculations, they are implicitly doing a form of target-return pricing. Otherwise, we can simply treat fixed costs as a factor that has a positive effect on firms' markups.
3) *Sales effort* (+): Advertising and marketing are important, not only for increasing the number of customers a firm has, but also because 'product differentiation' can make consumers more loyal to specific brands and thereby make them less willing to switch to other products if the price is increased. Although the sales effort can be considered part of fixed costs, it is also important in its own right because of its ability to increase the level of demand for a firm's products and also to make that demand less price-elastic. Although it is perhaps more Schumpeterian than Kaleckian, one could also include the part of R&D expenditures that is devoted to

product innovation under this rubric since the development of new and improved products is a key part of sales effort – and new or improved products may be eligible for forms of intellectual property protection (patents, copyrights and so on) that enable firms to charge higher than normal markups. Thus, sales effort – broadly defined – has a positive effect on the markups that firms are able to charge over and above the need for firms to pay the fixed costs involved.

4) *Strength of labour unions* (–): Kalecki recognized that workers could effectively capture part of an oligopolistic firm's potential profits if they have a strong bargaining position, especially through the formation of unions (providing, of course, that the unions are independent of the firms and not under the thumb of the government). How such bargaining occurs may vary across countries; for example, it may be centralized (as in Germany) or decentralized with staggered contracts (as in the US). Assuming that firms cannot pass through 100 per cent of their increases in labour costs (wages or benefits) to consumers without risk of losing too much market share, firms' profit margins (markups) must be squeezed to some extent by such increases.[7]
5) *External competition* (–): This could come from a competitive fringe or non-union firms domestically, or foreign producers in open economies. Such competition can undermine the price-setting power of oligopolistic firms and force them to lower prices and profit margins (markups), for any given degree of concentration, fixed costs and other factors.[8] In a sense, such external competition reduces the effective degree of industrial concentration and so could be considered a subset of the first factor, but empirically (in studies of industry-level or firm-level profit margins) it is often measured by distinct variables such as levels of import penetration in addition to the impact of concentration ratios per se. Most importantly, external competition limits the ability of firms to pass through increases in labour and other costs, and hence amplifies the ability of such cost increases to squeeze markups.

Finally, a very important link in Kalecki's theory is the relationship between the average markup rate of firms and the (gross) profit share of value added. Suppose that the average variable costs of firm $i$ are the sum of its unit (average) labour costs and unit (average) materials costs, $AVC^i = ULC^i + UMC^i$, and for simplicity we assume here that all labour is part of variable costs (thus, there is no overhead labour; a model incorporating overhead labour is presented in Appendix 4.1). Then, the markup pricing equation (4.2) becomes

$$P^i = (1 + \tau^i)(ULC^i + UMC^i) \tag{4.7}$$

Since the firm's value added ($VA^i$) does not include materials costs, it must equal the sum of labour costs plus the gross profits. Thus, measured in per-unit terms for the *i*th firm,

$$VA^i/Y^i = ULC^i + \tau^i(ULC^i + UMC^i) \tag{4.8}$$

Then, the gross profit share of value added for the *i*th firm, $\pi^i$, is defined as the ratio of gross profit per unit $\tau^i(ULC^i + UMC^i)$ to value added per unit (the right-hand side of equation 4.8):

$$\pi^i = \frac{\tau^i(ULC^i + UMC^i)}{ULC^i + \tau^i(ULC^i + UMC^i)} \tag{4.9}$$

Then, if we define $j^i = UMC^i/ULC^i$ as the ratio of materials to labour costs for firm *i*, we can rewrite (4.9) as

$$\pi^i = \frac{\tau^i(1 + j^i)}{1 + \tau^i(1 + j^i)} \tag{4.10}$$

which is a monotonically increasing function of $j^i$. On the one hand, this implies that a rise in $UMC^i$ (owing, for example, to increases in energy or other commodity prices, or a currency depreciation in a country that imports its raw materials) results in a higher gross profit share for any given markup rate $\tau^i$ (also taking $ULC^i$ as given). On the other hand, it implies that a rise in $ULC^i$ (for example, as a result of workers negotiating for higher wages), which reduces $j^i$, will lower the profit share (for any given markup rate $\tau^i$ and unit materials costs $UMC^i$).

This relationship is a crucial feature of the Kaleckian approach, because it links the distribution of income (the relative share of profits in value added) to the markups and cost structures of firms at the micro level. This was a revolutionary new way of thinking about the functional (wage–profit) distribution of income, which Kalecki viewed as deriving from the pricing policies of oligopolistic firms. It represents a different approach to income distribution from what is found in any of the models (classical-Marxian, neo-Keynesian and syntheses thereof) covered in Chapters 2 and 3. This approach is quite distinct from the classical-Marxian view, in which the profit share is a function of the real wage relative to labour productivity (if the real wage is exogenously given), or an exogenous datum (if the wage share itself is fixed), or simply a residual if the economy has to adjust to grow at a natural rate (in which case the equilibrium profit share is determined by the rate of labour supply growth, saving rate out of profits and aggregate capital–output ratio: $\pi^* = a_1 n/s_r$, assuming $q = 0$ for simplicity). It also differs from the neo-Keynesian views of Kaldor and Robinson, in which the profit share has to be flexible in order to adjust the profit rate to the level required for equilibrium between saving and investment. In Kalecki's view, the profit shares of individual firms are determined by their pricing policies, and these firm-level profit shares can then be aggregated up to determine the economy-wide average profit share.[9] We now turn to the types of macroeconomic models that can be built upon these microfoundations and this unique theory of income distribution.

## 4.3 The Kalecki–Steindl macro model

Neo-Kaleckian macro models draw their inspiration from Kalecki's *Theory of Economic Dynamics* (1954 [1968]) and his other writings collected in Kalecki (1990, 1991), as well as Steindl's *Maturity and Stagnation in American Capitalism* (1952 [1976]). However, the presentation in this section is not based directly on these original sources, but rather is an amalgam of later interpretations and formalizations.[10] We will present a simplified neo-Kaleckian model that illustrates the core principles of Kalecki and Steindl in this section; various critiques and extensions of (or alternatives to) this approach will be discussed later in this chapter and in the following three chapters. To clarify terminology, we will use the term 'neo-Kaleckian' for the entire literature that has revived Kalecki's approach and extended his and Steindl's work in various directions (even when sometimes disagreeing with some of their original hypotheses), while designating specific models within this literature by various particular names or labels. This section presents what may be considered a canonical version of a Kalecki–Steindl model, which encapsulates various of their key original ideas as will be explained in more detail below.

### 4.3.1 Markup pricing, income distribution and capacity utilization

The neo-Kaleckian macro models considered in the rest of this chapter will be based on a simplified version of Kalecki's theory of the firm, in which we abstract from raw materials and overhead (fixed) costs.[11] However, a version of the model with overhead labour included is presented in Appendix 4.1, while a neo-Kaleckian model of pricing including imported intermediate goods (which could be raw materials) is covered in section 10.6 in Chapter 10.

Dropping the *i* superscripts for notational simplicity, the price level for a 'representative firm' can be expressed as

$$P = (1 + \tau)Wa_0 \tag{4.11}$$

where $\tau > 0$ is the markup rate, $W$ is the nominal wage rate, $a_0$ is the labour coefficient (labour hours per unit of output) as defined in previous chapters, and average variable costs consist entirely of unit labour costs ($AVC = ULC = Wa_0$). Because there are no raw materials costs, price equals value added per unit ($P = VA/Y$), and therefore the share of profits in value added is the same as the gross profit margin as defined earlier (but with $UMC = 0$):

$$\pi = \frac{P - Wa_0}{P} = \frac{(1 + \tau)Wa_0 - Wa_0}{(1 + \tau)Wa_0} = \frac{\tau}{1 + \tau} \tag{4.12}$$

which is a monotonic increasing function of the markup rate $\tau$. This is the same as equation (4.10) on the simplifying assumption that $j = 0$. On the (admittedly heroic) simplifying assumption of a single representative firm, equation (4.12) represents the aggregate profit share for the whole economy.

As discussed above, industrial firms typically have excess capacity, in the sense that they usually produce a level of output below their maximum feasible capacity. This enables firms to respond to fluctuations in demand by varying their output level and, consequently, the utilization of their capacity. As in Chapter 1, we will define $Y_K$ as potential output determined by the available capital stock, $K$; this definition of potential output corresponds to what we called 'full utilization' in the previous section. Also, as in previous chapters, we define $a_1 = K/Y_K$ as the ratio of capital to full-capacity output. The capacity utilization rate is again defined as $u = Y/Y_K$, where $Y$ is actual or current output. Unlike in Chapters 2 and 3, however, all the models covered in this chapter assume that there can be excess capacity: $0 < Y \leq Y_K$ and $0 < u \leq 1$. Although the capital stock is thus related to potential output, employment of labour (which is treated as an entirely variable input, since we are abstracting from overhead labour here) is proportional to current or actual output, which in turn is linked to the utilization rate as well as the capital stock: $L = a_0Y = a_0uY_K = a_0uK/a_1$.

As in the micro models of the firm covered in the previous section, we will assume that firms generally operate at less than full utilization of capacity ($0 < u < 1$). Unlike in some of those micro models (and also unlike in the classical-Marxian and neo-Keynesian models covered in Chapters 2 and 3, with the exception of Harrod), however, we will not assume here that capacity utilization must converge to a unique 'normal' rate $u_n$ in the long run. In the neo-Kaleckian macro models covered in this chapter, the utilization rate is simply an endogenous variable that has to adjust, as we shall see, to bring saving into equilibrium with investment. Formally, this means that these models solve for equilibrium utilization rates that can lie anywhere at or below unity, and some authors have interpreted these equilibria (regardless of how far below unity they may be) as describing long-run steady states. Alternatively, these models could be interpreted as determining only short-run equilibria, in which case some mechanism would have to be specified that would make the economy adjust towards a normal utilization rate (or normal range of utilization rates) in the long run. The debate about the degree to which the utilization rate can vary in the long run (including whether 'normal' rates are themselves endogenous) and what sorts of adjustment mechanisms could make producers converge towards a long-run normal utilization rate (or range of rates) will be considered in Chapter 6.

As in previous chapters, total national income equals the sum of wage and profit income

$$PY = WL + rPK \tag{4.13}$$

where again we simplify by ignoring depreciation of the capital stock so that $Y$ represents both gross and net output (or income), and there is only one good so $P$ is the price of capital as well as output. Dividing both sides by $PY$ and using the above definitions, we obtain

$$1 = wa_0 + ra_1/u \tag{4.14}$$

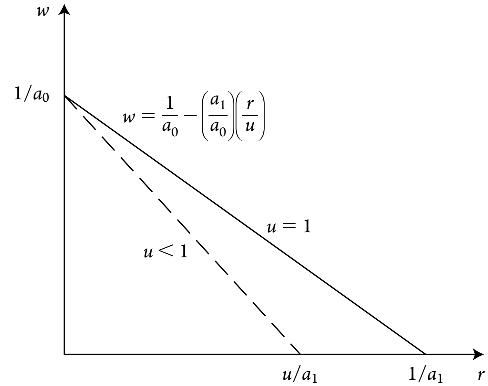
which can be solved for the modified inverse wage–profit relation

$$w = \frac{1}{a_0} - \left(\frac{a_1}{a_0}\right)\left(\frac{r}{u}\right) \tag{4.15}$$

In this equation, the real wage appears to be positively related to the utilization rate (for a given profit rate and technology), but that is quite misleading because the profit and utilization rates are not independent of each other. As we shall see below, it turns out that equilibrium real wage $w^*$ is causally independent of both the equilibrium levels $r^*$ and $u^*$ in this model.[12]

In Figure 4.2, the solid line represents the maximum wage–profit frontier defined by the prevailing technology with full utilization ($u = 1$ or 100 per cent), while the dashed line represents the actual wage–profit trade-off assuming excess capacity or less than full utilization ($u < 1$). Note that the dashed line rotates inward (pivoting on a constant $w$-intercept) as utilization decreases and rotates outward (towards the maximum frontier) as utilization

**Figure 4.2** The inverse wage–profit relation with full utilization ($u = 1$) compared with excess capacity ($u < 1$)




increases. In this model, the maximum technically feasible real wage continues to correspond to the productivity of labour ($1/a_0$), while the equilibrium real wage can be solved for directly by substituting equations (4.11) and (4.12) for the price level and the profit share into the definition of the real wage as follows:

$$w^* = \frac{W}{P} = \frac{W}{(1+\tau)Wa_0} = \left(\frac{1}{1+\tau}\right)\frac{1}{a_0} = \frac{1-\pi}{a_0} \qquad (4.16)$$

where $1 - \pi = 1/(1+\tau)$ is the wage share. Thus, the real wage is determined strictly by the wage share, which is inversely related to the markup, and the productivity of labour ($1/a_0$), and is independent of the utilization rate or profit rate. The intuitive story behind this is that, similar to the neo-Robinsonian story, the nominal wage rate $W$ is set in labour contracts, but unlike in the neo-Robinsonian story, the price level $P$ is determined by firms' markups over their unit labour costs, so that the purchasing power of the workers' nominal wage depends only on the level of those markups and their own productivity.

In Figure 4.2, the maximum profit rate based on actual output ($u/a_1$) depends on the rate of capacity utilization ($u$) as well as the potential productivity of capital ($1/a_1$) and equals the latter only at full utilization ($u = 1$). The actual profit rate, however, also depends on the profit share (and thus on the underlying markup rates of firms). To see this, note that by definition the profit rate is the ratio of profits to capital, and therefore (using the preceding definitions as needed)

$$r = \frac{PY - WL}{PK} = \frac{[(1+\tau)Wa_0 - Wa_0]Y}{(1+\tau)Wa_0K} = \left(\frac{\tau}{1+\tau}\right)\frac{Y}{K} = \frac{\pi u}{a_1} \quad (4.17)$$

Thus, in a *definitional* sense, the actual or realized profit rate is related positively to the profit share $\pi$ (which in turn depends on the average markup) and utilization rate $u$, and inversely to the ratio of capital to full-capacity output $a_1$ (or positively to the productivity of capital at full utilization, $1/a_1$). However, great caution must be exercised because equation (4.17) is *not* a causal statement. Given $\pi$ (as determined by the underlying markup rate) and $a_1$ (which depends on the existing technology), the equilibrium profit rate $r^*$ is determined *simultaneously* with the equilibrium utilization rate $u^*$. As we shall see below, the neo-Kaleckian approach thus allows for the paradoxical result that, in comparisons across equilibrium states, the profit share $\pi$ and profit rate $r$ may (under certain conditions) be *inversely* rather than positively related to each other. That is, a rise in $\pi$ may possibly depress, rather than increase, equilibrium $r^*$, because the negative effect of increased $\pi$ on equilibrium utilization $u^*$ can potentially outweigh the direct positive effect of the rise in $\pi$.

### 4.3.2 Saving–investment equilibrium and the model solution

In constructing a canonical version of the Kalecki–Steindl model, we will continue to assume the same saving function we used for the neo-Robinsonian model, equation (3.27) in Chapter 3, which is reproduced here for convenience:

$$\sigma = s_r r \tag{4.18}$$

This specification assumes that all saving comes out of profits. We will also assume a closed economy (no foreign trade) and a pure private economy (no government) for simplicity, although all of these assumptions can be relaxed (the cases of positive saving out of wages and an open economy will be considered later in this chapter).

In this basic framework, we then face the crucial choice of what kind of investment function to assume. Although many versions have been proposed – and the consequences of assuming a different one will be considered later in this chapter – we will start with what might be called the Rowthorn–Dutt–Taylor investment function that was prevalent in the early 1980s when the modern neo-Kaleckian approach emerged, and which (arguably) reflects much of the original vision of Kalecki and Steindl on the subject.[13] This investment function is

$$g = g_0 + g_1 r + g_2 u \tag{4.19}$$

where desired investment is an increasing function of the actual, realized profit rate $r$ and the capacity utilization rate $u$.[14] We assume a linear functional form here for mathematical convenience, with coefficients $g_0$, $g_1$, $g_2 > 0$, but the same qualitative conclusions would be reached using a more general (implicit) investment function $g = g(r, u)$ as long as we assume $g_r$, $g_u > 0$.[15]

The Kalecki–Steindl approach does not adopt the Robinsonian emphasis on the expected profit rate $r^e$ as the key determinant of firms' desired investment.[16] This is because, in Kalecki and Steindl's view, the profits received by firms influence investment mainly by relieving financial or liquidity constraints, in which case it is actual rather than expected profits that matter most. Kalecki (1937) developed a theory of 'increasing risk' in investment finance, in which firms can be charged higher interest rates on external funds or find those funds rationed as they increase their levels of investment. As developed later by Steindl (1952 [1976]) and Minsky (1975, 1986), this theory implies that a higher level of firms' cash flow (retained earnings or gross corporate savings, equal to net profits after taxes, dividends and interest, but including depreciation allowances) would enable firms either to finance more investment internally (that is, out of their own funds) or to access external finance on better terms (lower interest rates and so on) because they would be seen by lenders as better credit risks. Since these additional complications (interest, dividends, taxes and depreciation) are not included in the present, simplified model, the relevant variable that defines firms' financial positions is their realized rate of profit $r$ (that is, the current flow of profits normalized by the capital stock).

Moreover, many theories of investment imply that it should be sensitive to some measure of demand for firms' products, usually represented by the growth rate of their output or sales ($y = \Delta Y/Y$, which is called the 'accelerator effect'[17]) or their rate of capacity utilization $u$ (which measures the adequacy of the existing capital stock). Steindl (1952 [1976], 1979) included the utilization rate in his investment function, as did many of his followers listed earlier (especially Rowthorn and Dutt). The logic behind using the utilization rate is clear: given the many reasons for firms to desire to hold excess capacity, as discussed earlier, if existing capacity starts to be more heavily utilized that gives firms a signal to invest more in expanding their capacity in order to avoid the risk of running out of excess capacity and hitting upon a full-utilization constraint. Since using the utilization rate is mathematically simpler than using the growth rate of output (because it allows for a static rather than a dynamic solution), it is the specification we will adopt here.[18] The constant term $g_0$ in this function can be thought of as injecting a Keynesian element of 'animal spirits' into the Kalecki–Steindl investment function, since this term can be used as a shift factor to represent changes in the degree of optimism or pessimism of firms about market conditions and future profitability.[19]

In solving the saving–investment part of the model for a closed economy with no government, we adopt the same saving-equals-investment equilibrium condition that we used for the neo-Robinsonian model, which represents goods market clearing:

$$\sigma = g \tag{4.20}$$

Here, in addition to substituting the saving and investment functions (4.18) and (4.19), we also use the correspondence between the rates of profit and utilization given by equation (4.17) to eliminate one of the two endogenous variables ($r$ or $u$) and reduce the resulting equation (the equilibrium condition with the saving and investment functions substituted into it) to one equation in one unknown. It does not matter in the end whether one solves the model in terms of $r$ or $u$, since one can use $r = \pi u/a_1$ to transform the solution for either one of these variables into a solution for the other. In what follows, we will solve the model algebraically in terms of $u$, which is more revealing about the intuitive logic of the model as well as mathematically more convenient, but the same results can be obtained either way.

Thus, by substituting (4.17) into (4.18) and (4.19) and then substituting the resulting expressions into (4.20), we obtain

$$s_r \pi u/a_1 = g_0 + g_1 \pi u/a_1 + g_2 u \tag{4.21}$$

which (given the linear specification of the equations) yields the explicit, reduced form solution for the utilization rate

$$u^* = \frac{g_0}{(s_r - g_1)(\pi/a_1) - g_2} \tag{4.22}$$

The denominator of (4.22) must be positive if the model satisfies the goods market stability condition. To see this, note that the stability condition in this model can be obtained from the excess demand for goods written as a function of the utilization rate:

$$EDG = g - \sigma = g_0 + g_1 \pi u/a_1 + g_2 u - s_r \pi u/a_1 \tag{4.23}$$

In the neo-Kaleckian framework, the utilization of capacity (that is, actual output relative to potential) is the key adjusting variable, and the condition for increases in this variable to eliminate excess demand is

$$\frac{\partial EDG}{\partial u} = (g_1 - s_r)\frac{\pi}{a_1} + g_2 < 0 \tag{4.24}$$

It is easily seen that satisfaction of (4.24) is equivalent to the denominator of (4.22) being positive. Then, using (4.17) and (4.22), we find the solution for the equilibrium profit rate to be

$$r^* = \frac{g_0(\pi/a_1)}{(s_r - g_1)(\pi/a_1) - g_2} \tag{4.25}$$

Finally, we can use the solutions (4.22) and (4.25) in either (4.18) or (4.19) to find the equilibrium solution for $\sigma = g$, and while it is simpler to use the former, either way (after simplification in the latter case) the result should be

$$g^* = \sigma^* = \frac{s_r g_0(\pi/a_1)}{(s_r - g_1)(\pi/a_1) - g_2} \tag{4.26}$$

Note that positive solutions for $u^*$, $r^*$ and $g^*$ require $g_0 > 0$ along with satisfaction of the stability condition (4.24).

This equilibrium solution can be illustrated graphically as shown in Figure 4.3, where we rotate the wage–profit diagram counterclockwise 90 degrees to put it in the left-hand quadrant. For purposes of comparison with the other models (classical-Marxian and neo-Keynesian) covered earlier, we draw the saving–investment part of the graph (right-hand quadrant) in terms of the profit rate $r$ rather than the utilization rate $u$, which requires using (4.17) to transform the investment function (4.19) into:

$$g = g_0 + g_1 r + g_2 r a_1/\pi = g_0 + (g_1 + g_2 a_1/\pi) r \tag{4.19'}$$

The solution shown in Figure 4.3 looks similar to the one in Figure 3.3 for the neo-Robinsonian model, but has several key differences. First, and most obviously, the wage–profit relation rotates inward (downward, as drawn) to the extent that there is excess capacity, and the equilibrium utilization rate is represented as shifting the $r$-intercept (and slope) of this relationship relative to the outer distributional frontier. In effect, there is no longer a strict or fixed trade-off between wages and profits, but rather a relationship that shifts inward or outward depending on the utilization rate. Second, in terms of causality, the illustrated equilibrium is based on a simultaneous solution for

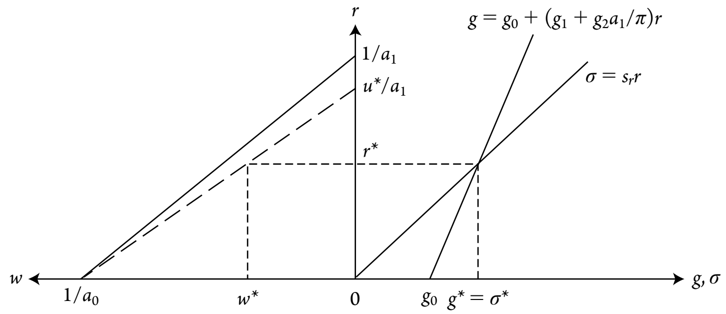



**Figure 4.3** Equilibrium in the Kalecki–Steindl growth model with excess capacity ($u^* < 1$)

three variables: the rates of utilization, profit and growth. Thus, we omit the bold arrows indicating the direction of causality that we have employed in previous diagrams of this type, because all these variables are simultaneously determined. The solutions for $r$ and $g$ are shown along their respective axes, while the solution for $u$ is depicted as determining the $r$-intercept of the dashed wage–profit relation. Third, the real wage is fixed by equation (4.16) independently of the other endogenous variables shown in this diagram. Implicitly, variations in labour demand are felt entirely as changes in the level of employment, and do not impact on the real wage (at least, not unless they affect markups). Fourth, the profit share now appears as a (negative) shift factor in the investment demand function; increases in $\pi$ cause this function to rotate leftward, so that the desired rate of investment is lower for any given profit rate $r$. We will explore the consequences of such a shift in the next subsection.

### 4.3.3 Concentration, stagnation and wage-led growth

The most important and unique contribution of the Kalecki–Steindl growth model lies in the fact that it implies something that appears impossible in most of the classical-Marxian and neo-Keynesian models: a more equitable distribution of income, in the form of a rise in the real wage or wage share of national income, may (even in the absence of technological progress) lead to greater output and more rapid growth. The classical-Marxian models and the neo-Keynesian models of Kaldor and Robinson rest on the foundation of a strict trade-off between the real wage and the profit rate, so that in general faster growth can be achieved only at the expense of greater inequality (a lower real wage and/or wage share) unless there is a technological improvement (which relaxes the wage–profit relation, as shown in Chapter 2). There are a few other exceptions in the classical-Marxian framework, such as the case of an increase in the capitalists' saving propensity with a natural rate of growth and endogenous labour supply (in which case both the real wage and the growth rate rise). But in most of the cases we examined in previous chapters, faster growth requires a redistribution of income towards profits and hence greater inequality.

By making the utilization rate endogenous, the Kalecki–Steindl approach allows for situations in which the real wage, profit rate and growth rate can all rise or fall simultaneously, instead of the first of these and the last two always varying inversely. A change in the utilization rate rotates the effective wage–profit relation (the dashed line in either Figure 4.2 or 4.3) inward or outward, thus enabling the real wage *and* profit rate to change in the *same* direction under certain circumstances, and in these cases the growth rate generally goes in the same direction too because it is an increasing function of both the profit rate and the utilization rate. Or, in a terminology that has become widely accepted in recent years, the Kalecki–Steindl approach implies that growth can be *wage-led*.

In this basic version of the neo-Kaleckian approach, the relative shares of wages and profits depend uniquely on the firms' markup rate per equation (4.12), while the real wage also depends on the markup rate along with labour productivity according to (4.16). Hence, taking labour productivity as given, the real wage or wage share can only rise or fall if the markup rate moves in the opposite direction. In section 4.2.1 above, we reviewed several factors that could raise the markup rate according to Kalecki: increased industrial concentration, higher overhead costs, greater sales effort, reduced power of unions or less external competition.

In addition, Steindl (1952 [1976]) developed a theory of an evolutionary process in which price–cost margins (which, as noted earlier, are positively related to markup rates) would tend to rise secularly over time. He hypothesized that initially competitive industries pass through a process he called 'absolute concentration', in which the higher-cost, less efficient firms are eliminated through cut-throat, price-slashing competition from their lower-cost, more efficient rivals. This process would lead eventually to a phase he called 'monopoly capitalism', in which industries would be dominated by the largest of the remaining firms.[20] These firms would acquire significant oligopoly power – and raise price–cost margins (or markup rates) accordingly. Then, as costs continued to fall due to technological innovations, the oligopolistic or monopolistic firms would capture the benefits through further increases in price–cost margins rather than pass the gains through to consumers in the form of lower prices. The result would be a secular tendency for margins (or, equivalently, markups) to rise, and per equation (4.12) above this would lead to a rising trend in the profit share.[21] Regardless of whether one accepts Steindl's original argument for why markups would tend to rise or not, recent empirical studies have found that average markup rates and profit shares have risen dramatically in the US economy since the 1980s and one contributing factor is a process of concentration in which the firms with higher profit shares (lower labour shares) have increased their market shares (see Autor et al., 2017; De Loecker and Eeckhout, 2017).

Because $\tau$ and $\pi$ rise or fall together per equation (4.12), it is sufficient to analyse the effects of a rise (or fall) in the latter variable on the rest of the model. First, by equation (4.16), one direct effect of a rise in the profit share is to lower the wage share $(1 - \pi)$ and through it the real wage:

$$\frac{\partial w^*}{\partial \pi} = -\frac{1}{a_0} < 0 \tag{4.27}$$

Next, we can use either total differentiation applied to the equilibrium condition written as (4.21) or else partial differentiation of the reduced form solution (4.22) to obtain the effect on capacity utilization:

$$\frac{\partial u^*}{\partial \pi} = \frac{-g_0(s_r - g_1)/a_1}{[(s_r - g_1)(\pi/a_1) - g_2]^2} < 0 \tag{4.28}$$

The negative sign of this effect depends crucially on the stability condition $(s_r - g_1)(\pi / a_1) - g_2 > 0$, which implies that $s_r - g_1 > g_2 a_1/\pi > 0$. Thus, if the equilibrium is stable (as neo-Kaleckian theorists tend to assume), the propensity to save out of profits must exceed the responsiveness of investment to profits, in which case the numerator of (4.28) is negative and a redistribution of income towards profits reduces capacity utilization. The intuitive reason for this outcome is that a higher profit share redistributes income from wages (100 per cent of which are spent on consumption) to profits (of which only the share $1 - s_r < 1$ is spent on consumption), and thus results in a decline in overall consumption demand. This outcome is similar to the idea of 'underconsumptionism' expressed by some nineteenth-century economists and Marx's idea of a 'realization crisis' caused by workers' depressed wages relative to their productivity.

Now, it might seem possible for this fall in consumption to be offset or outweighed by a rise in investment in response to the higher profit share, and in some more general versions of the neo-Kaleckian model (to be discussed later in this chapter) that can indeed happen, but in the canonical specification of the Kalecki–Steindl model presented here it cannot occur. In fact, in this model, not only is it impossible for investment to rise enough to offset the fall in consumption, but investment must also fall because the utilization falls by enough to reduce realized profits. To see this, we first note the paradoxical result that the profit *rate* is *inversely* related to the profit *share* in this model:[22]

$$\frac{\partial r^*}{\partial \pi} = \frac{-g_0 g_2/a_1}{[(s_r - g_1)(\pi/a_1) - g_2]^2} < 0 \tag{4.29}$$

This striking result is what Rowthorn (1981) called the 'paradox of cost', since it implies that a rise in the real wage (which would occur if $\pi$ falls) would increase rather than decrease the profit rate.[23] The reason behind this result was identified by Marglin and Bhaduri (1990) when they pointed out that it assumes that the elasticity of utilization with respect to the profit share in this model is less than $-1$, or greater than 1 in absolute value,[24] so that $u$ falls proportionately more than $\pi$ rises, thereby reducing $r = \pi u/a_1$. Marglin and Bhaduri argued that the necessity of this result in the Kalecki–Steindl model stems from the particular form of the investment function (4.19), which in their view double-counts capacity utilization (which is included in $r$ as well as by itself),[25] in conjunction with the assumptions of no saving out of wages and a closed economy with no government (see also Blecker, 2002a). We will discuss Bhaduri and Marglin's alternative proposal for an investment function in the next section.

Given that a rise in the profit share causes the profit and utilization rates to both decrease, it follows logically (indeed, it is obvious from equation 4.19) that the equilibrium growth rate is also diminished. Using (4.18), (4.20) and (4.29), it is easily seen that[26]

$$\frac{\partial g^*}{\partial \pi} = \frac{-s_r g_0 g_2/a_1}{[(s_r - g_1)(\pi/a_1) - g_2]^2} < 0 \tag{4.30}$$

The graph for this case is shown in Figure 4.4. In this graph, the outer wage–profit frontier at full utilization ($u = 1$) is omitted in order not to clutter the diagram. Thus, the outer dashed line in the left-hand quadrant of Figure 4.4 represents an initial rate of utilization $u^* \leq 1$, while the inner (more lightly dashed) line in that quadrant represents the new, lower utilization rate $u' < u^*$. The right-hand quadrant shows the investment demand function rotating to the left as a result of the increase in $\pi$. Intuitively, the story starts with the reduction in consumption demand caused by the redistribution of income towards profits, which are saved at a higher rate compared to wages, as discussed earlier. In effect, the real wage is reduced (from $w^*$ to $w'$), which lowers workers' purchasing power. The reduction in workers' consumption demand then pulls down capacity utilization, as firms reduce output in response. The lower utilization rate, in turn, also reduces the realized profit rate, and both of these induce firms to cut back on investment spending. Reduced investment then further depresses aggregate demand, utilization, employment and profits, resulting in continued investment cutbacks, until

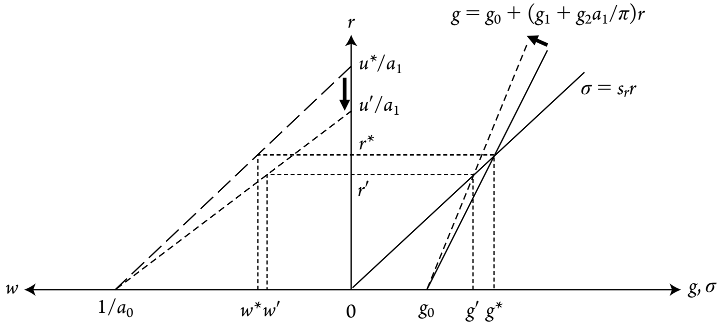



**Figure 4.4** The stagnationist effects of a higher profit share

(assuming that the stability condition holds) the economy converges to a new equilibrium constellation of lower rates of utilization, profits and growth along with a lower real wage.

In Kalecki and Steindl's view, this outcome shows that increasing oligopoly power or industrial concentration, which would lead to a higher profit share of national income, would lead to a tendency towards stagnation (chronically slow growth) in advanced capitalist economies like the US. Based on this conclusion, many authors have referred to this model as 'stagnationist'. Such a tendency towards stagnation could, of course, be offset by other factors, such as a fiscal stimulus or trade surplus, and in modern times we would have to include a debt-financed consumption or housing boom. But the underlying tendency would still be for income to become more unequally distributed and the system to lack internal dynamism on the demand side as a result of the heightened inequality.

By the same token, this model also demonstrates the possibility of what has become known as 'wage-led growth'. That is simply the reverse of the case just examined: a fall in the profit share would lead to simultaneous increases in the real wage, profit rate and growth rate. Since the model thus suggests that raising wages (via lower markups) can create a rising tide that lifts all boats (including realized profits and growth), the 'stagnationist' label can be confusing, so we prefer to refer to it simply as 'wage-led demand and growth'. Indeed, the Kalecki–Steindl model leads to considerable optimism about the economic feasibility of a more equitable form of economic growth. However, Kalecki (1943) observed that the political feasibility of a wage-led growth strategy could be limited if business interests perceive that workers would become more powerful and would be able to gain a greater share of the oligopolistic profits under conditions of rapid growth and full employment.

Finally, note that a rise in the profit share that cuts the real wage reduces employment in the Kalecki–Steindl model, since (as noted earlier) employment is proportional to utilization of capacity and the latter is reduced. This reduction in employment occurs in spite of the fact that firms experience lower labour costs, which under neoclassical logic would induce firms to 'substitute labour for capital' thereby hiring more workers and increasing employment. The conventional idea of a downward-sloping labour demand curve (based on supposed diminishing marginal productivity of labour) and conservative arguments against social policies such as minimum wage (or living wage) legislation assume that lower labour costs should lead to greater employment. In the Kalecki–Steindl model, however, lower labour costs only result in lower employment, as the reduction in consumption demand – amplified by its impact in reducing utilization and investment – ends up lowering output and employment as well as realized profits and growth. Based on this model, higher wages are good for growth and employment because they translate into increased demand and higher utilization of capacity. This model is thus an important intellectual foundation for pro-wage policies, such as the increases in minimum wages and more progressive tax policies now being advocated in many countries around the world in response to growing inequality and stagnation (for example, see Onaran et al., 2017).

### 4.3.4 The paradox of thrift and widow's cruse once again

The Kalecki–Steindl model exhibits a paradox of thrift and a widow's cruse, similar to the neo-Robinsonian model presented in Chapter 3, but with the crucial difference that in the former capacity utilization is also affected but the real wage is not. In the Kalecki–Steindl model, increases in $s_r$ reduce the equilibrium values of $u^*$, $r^*$ and $g^*$, while increases in $g_0$ do the opposite. Proofs of these propositions and depicting them on diagrams of the type shown in Figure 4.3 are left up to the reader as an exercise.

## 4.4 Alternative neo-Kaleckian models: wage-led versus profit-led demand regimes

As we have seen, the basic Kalecki–Steindl model leads to some very strong results. It implies that economies are generally wage-led, in the sense that a reduction in markups (which raises the real wage and lowers the profit share) will also increase the rates of utilization, profits and growth. These strong results in turn are based on many strong assumptions, including not only the basic propositions of excess capacity (variable utilization) and markup pricing, but also certain simplifications (for example, no workers' saving, a closed economy, and no government or fiscal policy) and special cases (for example, the particular investment function 4.19). In this section, we present some of the key alternative neo-Kaleckian models that can be constructed by dropping various of these simplifying assumptions or special functional forms, while still accepting the basic postulates of excess capacity (variable capacity utilization) and markup pricing.[27] In particular, by allowing for positive saving out of wages, a more general investment function and open economy considerations (effects of international trade), we will discover that neo-Kaleckian models can generate a variety of outcomes in terms of whether variables like the utilization and growth rates can be either wage-led or profit-led.

### 4.4.1 Positive saving out of wages

Although this case has received less attention in the literature than some of the others,[28] it is potentially important for two reasons. First, working-class households in some countries (especially in East Asia) have relatively high saving rates, which implies that the assumption of zero saving out of wages would be inappropriate for those cases. One could also be interested in the implications of efforts to raise workers' saving rates, for example through pension reforms. Second, even though we are not covering fiscal policy explicitly here, the effects of taxes levied on different types of income (wages, corporate profits, interest or dividend income and so on) are analogous in many respects to the effects of differential (but positive) saving propensities out of these different types of income.[29]

In order to incorporate saving out of wages, we will use a saving function more similar to the Kaldorian one (equation 3.16 in Chapter 3) in place of the Robinsonian saving function (equation 3.27 or 4.18):

$$\sigma = S/K = [s_r\pi + s_w(1-\pi)]u/a_1 \qquad (4.31)$$

where we assume $0 \leq s_w < s_r \leq 1$. The main difference from the Kaldorian saving function is that when we normalize saving and output by the capital stock here, we have to incorporate the utilization rate $u$ (which was assumed to be fixed at unity in Kaldor's model, and could therefore be ignored). Note that the simpler saving function (4.18) re-emerges in the special case in which $s_w = 0$, since $r = \pi u/a_1$.

In light of Pasinetti's (1962) critique of Kaldor (1955–56) and Kaldor's (1966b) response, discussed in Chapter 3, we assume that the saving propensities $s_w$ and $s_r$ pertain to the type of income received (wages or profits), not to the classes of agents (workers or capitalists) who receive it. In other words, profit income is always saved at a higher rate than wage income, even when some of it is received by working-class households as returns to their accumulated savings, presumably because a portion of profit income is retained by corporate firms as 'corporate saving' before any of it is paid out to households (for example, stockholders and bondholders, who may include workers whose pension funds are invested in corporate assets). Thus, we ignore any possible differences in household saving rates between pure worker households and households that own financial assets; models that include such differences will be considered in Chapter 7.

Using equation (4.31) for the saving function along with the Kalecki–Steindl investment function (4.19) in the equilibrium condition (4.20), we can easily find the solutions for the utilization, profit and growth rates with positive saving out of wages:

$$u^* = \frac{g_0}{(s_r - s_w - g_1)(\pi/a_1) + (s_w/a_1) - g_2} \tag{4.32}$$

$$r^* = \frac{g_0(\pi/a_1)}{(s_r - s_w - g_1)(\pi/a_1) + (s_w/a_1) - g_2} \tag{4.33}$$

$$g^* = \sigma^* = \frac{[(s_r - s_w)\pi + s_w](g_0/a_1)}{(s_r - s_w - g_1)(\pi/a_1) + (s_w/a_1) - g_2} \tag{4.34}$$

Here, the stability condition implies that $(s_r - s_w - g_1)(\pi/a_1) + (s_w/a_1) - g_2 > 0$ (found via the same method as shown earlier), which ensures that the denominators of all these solutions are positive. Once again, economically meaningful (positive) solutions require $g_0 > 0$ as well as fulfilment of the stability condition.

Now, the major difference from the basic Kalecki–Steindl model in which $s_w = 0$ is that, when $s_w > 0$, satisfaction of the stability condition no longer suffices to establish that any of these variables must be wage-led (inversely related to the profit share $\pi$). To see this, note that the effect of a higher profit share on utilization is given by

$$\frac{\partial u^*}{\partial \pi} = \frac{-g_0(s_r - s_w - g_1)/a_1}{[(s_r - s_w - g_1)(\pi/a_1) + (s_w/a_1) - g_2]^2} \tag{4.35}$$

Since the denominator is squared, the sign depends only on the sign of the numerator, and in this case (unlike the Kalecki–Steindl model) the stability condition does not suffice to determine the latter.[30] Thus, the sign of $\partial u^*/\partial \pi$ is ambiguous in general; aggregate demand (as represented by the utilization rate) can be either wage-led or profit-led when there are positive savings out of wages, albeit with a lower propensity than out of profits.

Further inspection of equation (4.35) reveals three important things about the factors that can make an economy more likely to have wage-led or profit-led aggregate demand (utilization). First, in regard to the propensities to save, what matters especially is the *gap* between the saving propensities out of profits and wages, $s_r - s_w$. If this gap is large, there is a big gain in consumer demand when income is redistributed to wages, so the economy is more likely to be wage-led (in the extreme, when $s_w = 0$, the economy must be wage-led, as we saw previously). On the other hand, the closer $s_w$ is to $s_r$, the smaller is that gain, and the more likely the economy is to be profit-led (since then the loss of investment demand from lower profits can outweigh the gain in consumption demand).[31] Second, and this follows from the last point, greater responsiveness of investment to profits ($g_1$) makes the economy more likely to be profit-led, as long as $g_1$ is not so high as to make the goods market unstable. A relatively high $g_1$ enhances the prospects for profit-led demand because it means that the boost to investment from a higher profit share can more than offset the corresponding decrease in consumption, and of course this more likely to be true when $s_w$ is also relatively high. Third, even though $g_2$ does not appear in the numerator, a relatively high $g_2$ still helps to make the system wage-led, because if $g_2 > s_w$ then the model can only be stable if $s_r - s_w - g_1 > 0$ (in which case $\partial u^*/\partial \pi < 0$ must hold).

Using equation (4.34), we find that the effect of a higher profit share on the growth rate[32] is given by

$$\frac{\partial g^*}{\partial \pi} = \frac{\partial \sigma^*}{\partial \pi} = \frac{(g_0/a_1)\left[(s_w - g_2)(s_r - s_w) - (s_w/a_1)(s_r - s_w - g_1)\right]}{\left[(s_r - s_w - g_1)(\pi/a_1) + (s_w/a_1) - g_2\right]^2} \tag{4.36}$$

This derivative is also ambiguous in sign, since both $s_r - s_w - g_1$ and $s_w - g_2$ are ambiguous, so growth can either be profit-led or wage-led (but under conditions that are different from the conditions for demand to be profit-led or wage-led). The intuition for the sign of this derivative is more difficult, given the greater complexity of the solution, but a few points are clear. First, a greater sensitivity of investment to profits ($g_1$) makes profit-led growth more likely to occur, provided again that $g_1$ is not so high as to make the system unstable. Second, a relatively high accelerator (utilization) effect on investment $g_2$ tends to make growth more likely to be wage-led, since in this case there can be a positive feedback from higher wages to increased consumption to greater investment, as firms respond strongly to the higher utilization rates brought about by increased consumer demand. On the other hand, the impact of the gap in the saving propensities $(s_r - s_w)$ is less clear in regard to growth, given its interaction with other terms, one of which $(s_w - g_2)$ is itself ambiguous in sign, in equation (4.36). In any event, this analysis also reveals that, once we allow for additional complications like positive saving out of wages, it is possible that an economy could be wage-led by one criterion (demand or utilization) and yet profit-led by another (growth or capital accumulation).

### 4.4.2 A more general investment function (the Bhaduri–Marglin model)[33]

As the preceding discussion reveals, the relative strength of the two coefficients ($g_1$ and $g_2$) in the Kalecki–Steindl investment function (4.19) is an important determinant of whether an economy has wage- or profit-led demand and growth, once we allow for positive saving out of wages. However, even if there are zero savings out of wages, the same point can be seen by considering alternatives to that investment function, which Bhaduri and Marglin (1990) and Marglin and Bhaduri (1990) argued was a special case that imposed strong implicit assumptions. To see their point, first consider that the Kalecki–Steindl investment function (4.19) can be written as

$$g = g_0 + [(g_1\pi / a_1) + g_2]u \tag{4.19''}$$

by substituting $r = u\pi / a_1$. This shows that utilization is effectively double-counted in the function, in the sense that it appears once by itself and also as part of the realized profit rate, both times with positive coefficients ($g_1$ and $g_2$).

Second, and more deeply, consider the meaning of the apparently innocuous assumption that $g_2 > 0$. This coefficient represents the effect of a rise in the utilization rate on investment, holding the profit *rate* constant, that is, $g_2 = (\partial g/\partial u)|_{r=\bar{r}}$. But since $r = u\pi / a_1$, then for any given $a_1$, the only way that $r$ can stay constant when $u$ rises is for $\pi$ to fall by the same proportion as $u$ rises. Thus, to assume that $g_2 > 0$ is, in effect, to assume that if firms experience a simultaneous increase in utilization *and* fall in their profit share (which is positively related to their price–cost margin and markup rate) in the same proportions, the firms will *necessarily* desire to invest more. Marglin and Bhaduri (1990) argued that this was not a reasonable assumption, and hence in general $g_2$ could have either sign (or be zero, which of course would take us back to the neo-Robinsonian investment function 3.31 in Chapter 3).

A quick inspection of equations (4.28) to (4.30) in conjunction with the stability condition (4.24) reveals that most of the main results of the Kalecki–Steindl model are radically changed if it is possible for $g_2 < 0$: satisfaction of the stability condition no longer rules out the possibility that $\partial u^*/\partial\pi > 0$; in addition, $\partial r^*/\partial\pi$ and $\partial g^*/\partial\pi$ also become ambiguous in sign. Thus, according to Marglin and Bhaduri, to assume $g_2 > 0$ is to implicitly assume a 'strong accelerator condition' and to rule out the possible alternative of a 'strong profitability condition' (although, as we have seen, the latter can still emerge if we allow for positive saving out of wages, even with the investment function 4.19).

To construct their alternative analysis, Bhaduri and Marglin (1990) and Marglin and Bhaduri (1990) postulated what they claimed was a more general investment function that effectively combines Robinsonian and Kaleckian elements. They start from the premise, borrowed from the former, that desired investment depends fundamentally on expected profits: $g = f(r^e)$. Since the profit rate is increasing in both the profit share and the utilization rate, Bhaduri and Marglin then argued that (taking the capital–capacity ratio $a_1$ as given) these two variables should also determine firms' expected profit rate: $r^e = r^e(\pi, u)$. Evidently, then, the investment function can be written as

$$g = f[r^e(\pi, u)] = h(\pi, u) \tag{4.37}$$

where $h_\pi > 0$ and $h_u > 0$. For this function, Bhaduri and Marglin were willing to assume that $h_u > 0$, because $h_u = (\partial g/\partial u)|_{\pi=\bar{\pi}}$, that, is, it makes sense to assume that the partial derivative of investment with respect to utilization is positive when holding the profit *share* (rather than the profit rate) constant. In other words, if firms experience a rise in utilization *without* any sacrifice of their profit share (margin or markup), *then* they would definitely be willing to invest more.

Using this investment function, both wage-led and profit-led results are possible even in a model with no saving out of wages (and still no international trade). Substituting equations (4.37) for investment and (4.18) for saving into the equilibrium condition (4.20), and again using (4.17) to substitute for the profit rate, goods market equilibrium can be characterized by

$$s_r u^* \pi / a_1 = h(\pi, u^*) \tag{4.38}$$

Since $h(\cdot)$ is an implicit function, we cannot derive an explicit solution for $u^*$ as we did before, but by total differentiation of (4.38) we find that the effect of an increase in the profit share on the equilibrium utilization rate is

$$\frac{\partial u^*}{\partial \pi} = \frac{h_\pi - s_r(u^*/a_1)}{s_r(\pi/a_1) - h_u} \tag{4.39}$$

evaluated at the equilibrium utilization rate $u^*$. As usual, the stability condition (analysed in the same way as we did for the Kalecki–Steindl model earlier) implies that the denominator must be positive, but in this case that has no bearing on the sign of the numerator, which can be either positive or negative. If the profitability effect on investment is relatively strong compared to the propensity to save out of profits $(h_\pi > s_r u^*/a_1)$, then $\partial u^*/\partial \pi > 0$ and demand is profit-led; in the opposite case $\partial u^*/\partial \pi < 0$ and demand is wage-led.

Marglin and Bhaduri (1990) coined the phrase 'exhilarationist' to refer to their profit-led case, since one could think of a strong profitability effect as indicating that capitalists' expectations are highly excited by a rise in the profit share. They referred to the Kalecki–Steindl model (or their own wage-led case) as 'stagnationist' because, as noted earlier, in that model a higher profit share leads to depressed aggregate demand as a reduction in the real wage leads to a large drop in workers' consumption (and no compensating increase in the desire to invest). Although this terminology was popular in the 1990s, it was ultimately seen as confusing because a stagnationist economy could be booming (if the profit share is low and the real wage is high) while an exhilarationist economy could be depressed (under the same conditions). Thus, in recent years the stagnationist and exhilarationist cases have more commonly been referred to by the more descriptive labels of 'wage-led' and 'profit-led' demand, and we will use the latter terminology here.

In the case where demand is wage-led $(\partial u^*/\partial \pi < 0)$, the results for the profit rate and growth rate then depend solely on the *elasticity* of utilization with respect to the profit share.[34] By differentiating $r^* = u^*\pi\,/a_1$ and $g^* = \sigma^* = s_r u^*\pi\,/a_1$ with respect to $\pi$ and using equation (4.39), it is easily seen that the Kalecki–Steindl results $(\partial r^*/\partial \pi < 0$ and $\partial g^*/\partial \pi < 0)$ occur if and only if $\frac{\partial u^*}{\partial \pi}\frac{\pi}{u^*} < -1$, or in absolute value $|\frac{\partial u^*}{\partial \pi}\frac{\pi}{u^*}| > 1$. In other words, for the equilibrium rates of profit and growth to be negatively affected by increases in the profit share, it is not sufficient for the profit share to have a negative effect on utilization; rather, the latter effect must be *more than unit elastic* (in absolute value). This, of course, is just a fancy mathematical way of saying that $u$ must fall proportionately more than $\pi$ rises in order for $r$ to decrease, but it is also a neat way of demonstrating that the Kalecki–Steindl results for the profit and growth rates depend on implicit assumptions (strong accelerator effect, weak profitability effect) that effectively ensure a strongly negative impact of a higher profit share on utilization and growth.

However, the possibility of profit-led demand in the Bhaduri–Marglin model (with no saving out of wages) also rests on what may be considered another strong elasticity condition: the elasticity of investment with respect to the profit share must exceed unity. To see this, note that the condition for the numerator of equation (4.39) to be positive is equivalent to $h_\pi a_1/u^* s_r > 1$, and using the fact that in equilibrium $g^* = \sigma^* = s_r \pi u^*/a_1$, that inequality can also be written as $(\pi / g^*)h_\pi > 1$ where $h_\pi = \partial g^*/\partial\pi$. Thus, a very strong profitability effect on investment is required for demand (utilization) to be profit-led. Mathematically, this is only possible with some functional forms for the investment relationship (4.37) and not with others, as long as we continue to assume a closed economy with no government and no saving out of wages.[35]

Bhaduri and Marglin saw their model as restoring the emphasis on profits as a driver of investment that was found in the classical-Marxian and neo-Keynesian approaches, but which they believed had been underplayed in the Kalecki–Steindl version of a neo-Kaleckian model. Even though the Bhaduri–Marglin model is one that assumes demand-driven output and growth, the emphasis on profitability brings the cost side of production back into the analysis. Much like Marglin (1984a) tried to synthesize neo-Marxian and neo-Keynesian approaches, Marglin and Bhaduri (1990) tried to blend neo-Marxian and neo-Kaleckian elements. However, defenders of the Kalecki–Steindl approach would argue that the utilization or accelerator effect should be strong in principle because of the importance of maintaining excess capacity for oligopolistic firms and because firms will not desire to invest more (which would increase their capacity) in the face of weak demand no matter how high is their profit share. In this view, the double-counting of utilization in the investment function (4.19) is not problematic, because utilization affects realized profits (and hence the availability of cash flow to relieve financial constraints) as well as having an independent effect on investment. Empirically, although the evidence does vary, most econometric studies that have estimated investment functions in recent years have tended to find relatively stronger accelerator effects than profitability (or financial constraint) effects in most countries,[36] thus calling into question whether the investment function is likely to be a key driver of profit-led demand.

Nevertheless, whatever one thinks of their investment function, Bhaduri and Marglin made a fundamental contribution in opening up the neo-Kaleckian approach to a wider range of outcomes in terms of the relationship between income distribution and different measures of economic performance in a model with excess capacity (variable utilization) and markup pricing. As summarized in Table 4.1, we may distinguish three alternative cases of the

**Table 4.1** Alternative taxonomies for neo-Kaleckian model results

| | Signs of partial derivatives | | |
|---|---|---|---|
| $\partial u^*/\partial \pi$ | − | − | + |
| $\partial g^*/\partial \pi$ | − | + | + |
| | **Alternative terminologies** | | |
| This book | Wage-led demand and growth | Wage-led demand, profit-led growth | Profit-led demand and growth |
| Marglin and Bhaduri (1990) | Cooperative stagnationist | Conflictual stagnationist | Exhilarationist |
| Palley (2013a) | Wage-led | Conflictive | Profit-led |

impact of distribution on demand and growth, which have been given different labels by different authors.[37] The first column describes the results implied by the canonical Kalecki–Steindl model, which Marglin and Bhaduri (1990) call 'cooperative stagnationist', Palley (2013a) simply calls 'wage-led', and we prefer (following Blecker, 2011) to call a situation of wage-led demand and growth. Of course, this outcome can also result in the other models (for example, with positive saving out of wages or the Bhaduri–Marglin investment function) under certain parameter values and elasticity conditions, as discussed earlier. The second column shows the case in which demand is wage-led but growth is profit-led, which Marglin and Bhaduri call 'conflictual stagnationist' and Palley (2013a) calls simply 'conflictive'; this is the case where $\partial u^*/\partial \pi < 0$ but the elasticity is less than unity in absolute value. Finally, the third column shows the case in which both demand and growth are profit-led (it is not possible for growth to be wage-led if demand is profit-led). This last case corresponds to what Marglin and Bhaduri call 'exhilarationist', and Palley (2013a) simply calls 'profit-led'.

Bhaduri and Marglin's terminology was designed to call attention to the fact that some of these regimes are more likely to facilitate a 'class compromise' than others. Especially, the cooperative stagnationist (all wage-led) case is one in which capitalists get increased rates of profit and growth in exchange for accepting a lower profit share, so they may have incentives to compromise with workers, while the conflictual stagnationist case is one in which capitalists get no gains whatsoever from a fall in their profit share and hence have no incentive to compromise in terms of accepting a lower profit share.[38]

### 4.4.3 The open economy neo-Kaleckian model

One important simplification in both the Kalecki–Steindl and Bhaduri–Marglin models (at least the basic versions covered above) is that they deal only with closed economies and domestic sources of aggregate demand. In this section, we will present the open economy analysis in the context of a fairly general neo-Kaleckian model that allows for positive saving out of wages, but the same general points can be demonstrated using various alternative specifications of the domestic economy.[39] The basic idea is a simple one: higher unit labour costs ($Wa_0$, or wages adjusted for productivity) in a 'home' country (relative to foreign countries) can, if passed through (at least partially) into prices of traded goods (exports and import-competing products), lead to a loss of competitiveness of home products and, under certain conditions, a reduction in the country's net exports (trade balance). Such a reduction in external demand in turn can potentially outweigh the gains in consumption from higher wages of workers, thereby leading to a contraction instead of an expansion of output. In such a case, aggregate demand (utilization) can be profit-led in an open economy, even if the domestic economy (the sum of consumption plus investment) is wage-led.[40]

The problem that has to be confronted in constructing such a model is that the simple neo-Kaleckian model of markup pricing discussed in section 4.2 does not allow for changes in nominal unit labour costs to influence the relative shares of wages and profits in national income, which depend only on the markup rate.[41] Blecker (1989a) solved this problem by introducing the idea that oligopolistic firms in open economies adjust their markup rates so as to maintain (to some degree) their international competitiveness in response to fluctuations in real exchange rates, that is, the relative prices of foreign compared with home goods. This core idea is found in a wide range of studies of pricing behaviour from both heterodox and mainstream perspectives (for example, Dornbusch, 1987; Feenstra, 1989; Arestis and Milberg 1993–94).

For mathematical convenience, the flexible markup rule is written in terms of the price–cost margin or one plus the markup, $1 + \tau = P/Wa_0 > 1$, as follows:

$$1 + \tau = \mu\left(\frac{EP_f}{P}\right)^{\eta} \tag{4.40}$$

where $\mu > 1$ is a target or desired markup rate of firms, $E$ is the nominal exchange rate (home currency per unit of foreign currency), $P_f$ is the foreign price level and $\eta > 0$ is the elasticity of the price–cost margin with respect to the real exchange rate.[42] Foreign prices $P_f$ and the nominal exchange rate $E$ are taken as exogenously given for simplicity. Equation (4.40) says that when the real exchange rate $(EP_f/P)$ rises (foreign goods become relatively more expensive, or the home currency depreciates in real terms), firms respond by raising their markups to take advantage of their improved competitiveness. Conversely, when the real exchange rate falls (foreign goods become relatively cheaper, or the home currency appreciates in real terms), domestic firms squeeze their profit margins in order to 'price to market' and avoid at least some of the loss of competitiveness that would otherwise result. It may be noted that this flexible markup rule embodies one of the causes of changes in markups noted earlier (section 4.2.1), namely the role of external competition.

Substituting the price equation (4.11) for $P$ in equation (4.40) and rearranging, we obtain the following solution for the price–cost margin:

$$1 + \tau = \mu^{\frac{1}{1+\eta}} z^{\frac{\eta}{1+\eta}} \tag{4.41}$$

where $z = EP_f/Wa_0$ is the ratio of the price of foreign goods to domestic unit labour costs, which is a measure of the country's international competitiveness in terms of labour costs.[43] Thus, in this model a country's unit labour costs do affect its markup (inversely), while other factors that influence markups – such as industrial concentration, overhead costs and unionization – are captured in the 'target' markup rate $\mu$. In this case, the profit share also becomes an increasing function of the competitiveness ratio $z$ and the target markup $\mu$:

$$\pi = \frac{\mu^{\frac{1}{1+\eta}} z^{\frac{\eta}{1+\eta}} - 1}{\mu^{\frac{1}{1+\eta}} z^{\frac{\eta}{1+\eta}}} = \pi(\mu, z) \tag{4.42}$$

where $\pi_\mu > 0$ and $\pi_z > 0$, and similarly the wage share $1 - \pi$ is inversely related to these two variables.

Next, we turn to the modelling of net exports, also known as the trade balance, in relation to relative prices and aggregate demand. The trade balance (measured as a ratio to the capital stock) is assumed to be positively related to the real exchange rate, which reflects the country's international price competitiveness; this requires that the price elasticities of export and import demand are high enough to satisfy the relevant Marshall–Lerner condition.[44] The trade balance (net export) ratio is also assumed to be inversely related to the output–capital ratio $u/a_1$ (increases which are associated with rising demand for imports relative to capital). Thus, the trade balance function is written as:

$$b = b(EP_f/P, u/a_1) \tag{4.43}$$

with partial derivatives $\partial b/\partial(EP_f/P) > 0$ and $\partial b/\partial(u/a_1) < 0$. This formulation makes perfect sense in a static model, that is, holding the capital stock and potential output constant, in which case a rise in domestic output (income) would be expected to increase imports without raising exports, thereby lowering net exports. However, this formulation is more problematic in a dynamic context in which the capital stock and potential output are also increasing, because if $Y$, $Y_K$ and $K$ all rise in the same proportion so that $u/a_1$ stays constant, equation (4.43) says that net exports would be unaffected, yet we would expect imports to rise and (assuming exports would remain unchanged) net exports to fall. However, (4.43) can then be justified by the implicit assumption that exports rise in proportion to the capital stock and potential output when these increase, and that the increase in exports would be just sufficient to keep net exports unchanged.[45]

For the investment function, we choose a special case of a linear Bhaduri–Marglin investment function that is especially relevant for an open economy:[46]

$$g = h_0 + h_1(\pi - \pi_f) + h_2 u/a_1, \quad h_1, h_2 > 0 \tag{4.44}$$

where $\pi_f$ is the foreign profit share (taken as exogenously given). Thus, $h_1$ is the sensitivity of domestic investment to the *difference* in profitability between the home country and abroad. This specification is motivated by the idea that both domestic and foreign firms will be looking at the profit margins in domestic and foreign locations in deciding where to produce goods, and hence where to invest in increasing their productive capacity. In addition, the utilization of domestic capacity $u$ still serves as a signal to firms of when they need to augment that capacity via additional investment. Although it is not essential, we choose to normalize the utilization rate here by the capital–capacity ratio $a_1$. This means that the accelerator effect is measured by the actual output–capital ratio $Y/K = u/a_1$, which seems more appropriate given that investment has also been normalized by the capital stock ($g = I/K$). Also, note that we do *not* have to assume a positive intercept ($h_0 > 0$) here, as it is not necessary for a positive solution $u^* > 0$, in contrast with the simple Kalecki–Steindl model presented earlier in which a positive intercept ($g_0 > 0$) was required for this purpose.

In an open economy with no government, the goods market equilibrium condition is that saving has to equal the sum of domestic investment plus the trade balance (net exports), where the latter can be thought of as the net outflow of domestic saving or increase in net foreign assets:

$$\sigma = g + b \tag{4.45}$$

Substituting the saving function with positive saving out of wages (equation 4.31) along with the investment function (4.44) and the trade balance function (4.43) into this equilibrium condition, and using (4.42) and the various definitions as needed, we can write the equilibrium solution for the utilization rate in implicit form as

$$[(s_r - s_w)\pi(\mu, z) + s_w)]u^*/a_1 = h_0 + h_1[\pi(\mu, z) - \pi_f] + h_2 u^*/a_1$$
$$+ b(\mu^{-1/(1+\eta)} z^{1/(1+\eta)}, u^*/a_1) \qquad (4.46)$$

where we use the fact that the real exchange rate can be written as $EP_f/P = EP_f/(1 + \tau)Wa_0 = z/(1 + \tau) = \mu^{-1/(1+\eta)} z^{1/(1+\eta)}$. We cannot obtain an explicit solution for $u^*$ here because the trade balance is expressed as an implicit function, but we can use total differentiation of equation (4.46) to obtain the comparative static results.

This brings us to the chief complexity of the open economy model. Because the markup rate and profit share are endogenous variables, we cannot simply vary them exogenously to determine the impact of distributional shifts on utilization (or other endogenous variables, such as the profit rate and growth rate). Rather, we have to consider changes in the two exogenous factors that affect income distribution: the labour cost competitiveness ratio $z$ and the firms' target markup rate $\mu$. Before we show the mathematics for these changes, there is an important point of intuition that must be stressed. That is, increases in $z$ and $\mu$ both lead to a higher profit share, but they have opposite effects on a country's external competitiveness as measured by the real exchange rate: a rise in $z$ improves external competitiveness (by reducing domestic labour costs relative to foreign prices), while a rise in $\mu$ worsens it (by increasing profit markups on domestic products). Thus, increases in $z$ and $\mu$ will have qualitatively similar effects on domestic demand (consumption and investment), but opposite effects on external demand (the trade balance or net exports). As a result, we no longer can have a unique characterization of a given economy as having either wage-led or profit-led demand; rather, in the open economy model, whether an economy behaves in a wage-led or profit-led fashion depends on the *source* of the distributional shift (that is, changes in relative labour cost competitiveness versus changes in industrial structure and market power of firms).

To see these different possibilities, we totally differentiate the equilibrium solution (4.46) to obtain the following partial derivatives:

$$\frac{\partial u^*}{\partial \mu} = \frac{h_1\pi_\mu - (s_r - s_w)(\pi_\mu u^*/a_1) - [z/\mu(1+\eta)(1+\tau)]b_1}{\Psi} \quad (4.47)$$

$$\frac{\partial u^*}{\partial z} = \frac{h_1\pi_z - (s_r - s_w)(\pi_z u^*/a_1) + [1/(1+\eta)(1+\tau)]b_1}{\Psi} \quad (4.48)$$

where $\Psi = [(s_r - s_w)\pi + s_w - h_2 - b_2]/a_1 > 0$ by the stability condition for this model,[47] $b_1 = \partial b/\partial(EP_f/P) > 0$ and $b_2 = \partial b/\partial(u/a_1) < 0$ are the partial derivatives of the trade balance function (4.43), and we use (4.41) and the fact that $\mu^{-1/(1+\eta)}z^{1/(1+\eta)} = EP_f/P = z/(1+\tau)$ in simplifying the solutions. Since a rise in either $\mu$ or $z$ leads to a higher profit share by equation (4.42), wage-led demand results if either $\partial u^*/\partial\mu$ or $\partial u^*/\partial z$ is negative, and profit-led demand occurs if either is positive.

To understand equations (4.47) and (4.48), note that the first two terms in each numerator are parallel to each other and correspond to similar terms representing the effects of the profit share on investment and saving in the solutions for the various closed economy models covered earlier (compare especially with equations 4.35 and 4.39). If the sum of the first two terms is positive, domestic demand (that is, the sum of consumption plus investment, ignoring net exports) is profit-led, and if the sum is negative then domestic demand is wage-led.

But in order to find the overall effect of a distributional shift on demand, we also have to take the last term in each numerator (which represents the impact on net exports) into account. The last terms in each of the numerators have different signs: negative in (4.37) and positive in (4.38). These signs correspond to the opposite effects of increases in $\mu$ and $z$ on international competitiveness noted above. Since a rise in the target markup $\mu$ makes home goods more expensive (less competitive) and thereby reduces net exports at the same time as it increases the profit share $\pi$, it has a negative effect on $u^*$ and therefore makes wage-led demand more likely to result. In contrast, a rise in labour cost competitiveness $z$ (for example, as a result of a cut in the nominal wage) makes home goods relatively cheaper and thereby increases net exports while also increasing the profit share $\pi$, thereby having a positive impact on $u^*$ and making profit-led demand more likely to obtain. Thus, *in the open economy case we cannot uniquely classify a given country as having either wage-led or profit-led demand*. Rather, we need to know the source of a 'shock' to income distribution: shocks to the industrial structure that affect target markups are more likely to have wage-led effects, while shocks to labour costs relative to other countries are more likely to generate profit-led responses (all else being equal).[48]

Of course, the impact of foreign trade can still potentially be offset by the domestic impact of either kind of distributional shift. For example, if the domestic economy (consumption plus investment) is strongly wage-led and the international competitive effects are relatively weak, then the net (overall) impact of higher relative labour costs (a lower $z$) could still be expansionary. This scenario is most likely to occur in a large country or one that is relatively closed to foreign trade. Similarly, if the domestic economy is strongly profit-led (for example, because of a high profitability effect on investment $h_1$ and a small gap in the saving propensities $(s_r - s_w)$), then a rise in the target markup $\mu$ could end up having an expansionary overall effect if the domestic gains outweigh the losses on the trade side. On the other hand, if the international trade effects are large, then those effects could outweigh the domestic impact of the distributional shift. For example, if relative labour costs rise (so $z$ falls) in a very open economy where trade is a large share of gross domestic product (GDP) and exports are highly price-elastic, the reduction in the trade balance could outweigh any possible gains in domestic demand and cause equilibrium utilization to fall, so the economy would behave in a profit-led manner even if domestic demand is wage-led.[49]

However, we do need to sound one note of caution about the implication that overall demand is more likely to be profit-led in an open economy in response to changes in relative unit labour costs $z$. In reaching this conclusion, we have considered only a single, small country that experiences a change in its unit labour costs while foreign unit labour costs (and prices) remain constant. However, if there is a simultaneous change in unit labour costs in the same direction globally, the competitive effects would largely cancel out, and many countries would exhibit wage-led demand even if they would have profit-led demand for a change in their own unit labour costs alone (see Onaran and Galanis, 2013; von Arnim et al., 2014). After all, the entire global economy is a closed system, so the closed economy models may be better guides to what would happen in response to a global redistribution of income, and in those models demand is more likely to be wage-led unless there are very strong profitability effects on investment combined with small gaps between savings out of profits and wages. We do stress 'more likely', however – if the underlying national economies have profit-led domestic demand (consumption plus investment), the global system could still be profit-led even in response to a simultaneous worldwide redistribution of income.[50]

Returning to the model of a single country, the open economy neo-Kaleckian model has some additional implications for the classic debates in international economics about whether currency depreciations are expansionary or contractionary and how effective they are for improving a country's balance of payments. Indeed, the sign of $\partial u^*/\partial z$ (equation 4.48) tells us very directly whether a real depreciation (an increase in $z = EP_f/Wa_0$) is expansionary or contractionary in a country that fits the assumptions of the neo-Kaleckian model (excess capacity, constant marginal costs, flexible markup pricing and so on). In the open economy neo-Kaleckian model, an expansionary depreciation corresponds to the case where a country exhibits a profit-led response to a change in relative labour cost competitiveness, while a contractionary depreciation corresponds to the case where the country exhibits a wage-led response to such a change.[51] In other words, there is a one-to-one correspondence between having a wage-led versus a profit-led response to a rise in $z$ and whether a real currency depreciation is contractionary or expansionary (respectively). This explains why some neo-Kaleckian economists are sceptical about the use of currency depreciation as a tool for correcting balance of payments (trade or current account) deficits: in addition to making the distribution of income more unequal (by reducing the wage share), a depreciation is likely to be contractionary if the country has wage-led demand.

By the same token, however, the sign of $\partial u^*/\partial z$ also has implications for whether a depreciation is relatively effective or ineffective for improving a country's net exports. To see this, note that the total effect of a rise in $z$ on the equilibrium trade balance $b^*$ is given by

$$\frac{db^*}{dz} = \frac{1}{(1 + \eta)(1 + \tau)} b_1 + \left(\frac{\partial u^*}{\partial z}\right)\left(\frac{1}{a_1}\right) b_2 \qquad (4.49)$$

The first term on the right-hand side is the standard Marshall–Lerner effect, modified by the parameters representing the adjustment of markups when the currency depreciates ($z$ rises), which we have assumed to be positive ($b_1 > 0$). This effect is stronger or weaker depending on how high the price elasticities of demand for exports and imports are. The second term is the unique contribution of the neo-Kaleckian approach: it is the income effect on the trade balance (net exports) created by the response of utilization to the redistribution of income towards profits that occurs as a result of a depreciation (rise in $z$). Assuming $b_2 < 0$ (because a rise in income increases imports), this term will have the opposite sign from $\partial u^*/\partial z$. In other words, if demand is wage-led ($\partial u^*/\partial z < 0$), the income effect is positive and the depreciation will be more effective for improving the trade balance; if demand is profit-led ($\partial u^*/\partial z > 0$), the income effect is negative and the depreciation will be less effective for improving the trade balance. This, of course, makes perfect intuitive sense: the more the depreciation reduces demand (income), the more the trade balance will improve, while the more the depreciation increases demand (income), the less the trade balance will improve. This trade-off between the domestic impact of a depreciation (contractionary versus expansionary) and the external impact (improving or failing to improve the trade balance) is well known; what is unique in the neo-Kaleckian approach is the recognition of how this trade-off is linked to the distributional impact of the depreciation (which always raises the profit share, and hence increases inequality).

## 4.5 Conclusions, critiques and extensions

In the core Kalecki–Steindl macro model, a redistribution of income towards wages (brought about by a fall in the profit markups of firms) always increases aggregate demand (measured by capacity utilization) as well as the realized profit rate and the growth rate of the capital stock. This occurs not only because of the stimulus to workers' consumption (since the real wage rises), but also because of a strong accelerator effect of the resulting rise in capacity utilization that boosts realized investment in spite of the decline in the profit share. Indeed, the utilization rate rises so much that the realized profit *rate* increases even though the profit *share* decreases. Thus, the model demonstrates the possibility of what have come to be known as 'wage-led' demand and growth, which could not occur in any of the classical-Marxian or neo-Keynesian models covered in the previous chapters. In this respect, the neo-Kaleckian approach has generated much excitement among progressive economists who would like to be able to advocate redistributive policies that would favour workers without fearing that such policies would undermine aggregate economic performance.

Nevertheless, these wage-led results were originally demonstrated under certain strong and restrictive assumptions, including no saving out of wages, no government or foreign trade, and a strong accelerator effect in the investment function. In more general neo-Kaleckian models in which these assumptions are relaxed, a redistribution of income towards labour can have more varied effects. For example, if there is positive saving out of wages (albeit with a lower propensity than saving out of profits), demand (utilization) can be either wage-led or profit-led, and growth can possibly be profit-led if demand is only weakly wage-led (and is definitely profit-led if demand is profit-led). If an alternative (allegedly more general) investment function is used, the system may generate either wage- or profit-led demand and growth, even in a closed economy with no savings out of wages. If a country is open to foreign trade and markups are flexible in response to international competitive pressures, the profit share becomes endogenous and there is no unique relationship between distribution and demand; the results depend on the cause of a distributional shift (monopoly power versus labour costs). One key result, however, is that a redistribution towards wages brought about by higher unit labour costs at home relative to the rest of the world is likely to reduce the country's net exports, and if this effect is large enough it can possibly cause the overall economy to behave in a profit-led fashion even if domestic demand is wage-led. Thus, in a broader set of neo-Kaleckian models, whether variables such as utilization and growth rates are wage-led or profit-led becomes an empirical question, which has become the subject of a vast econometric literature that will be critically surveyed in Chapter 5.

In addition, several aspects of the neo-Kaleckian framework have been subject to more fundamental criticisms. The models as specified in this chapter implicitly take full-capacity or potential output as given, so that variations in the utilization rate only reflect demand-driven changes in actual output. But this restricts the application of such models to short-run periods. For medium-run or long-run analysis, potential output would have to be allowed to adjust, but the macro models covered in this chapter are silent on how such adjustments would be likely to occur. This lacuna is problematic not only in theory, but also in regard to the empirical studies that estimate the determinants of utilization rates. As we will see in the next chapter, such studies use time-series data that incorporate changes in potential output as well as actual output, but often interpret their findings as if they applied only to actual output or demand. Some alternative theoretical approaches that may help to address this shortcoming are covered in Chapters 6 and 7.

Furthermore, some critics of the neo-Kaleckian approach (for example, Skott, 1989; Duménil and Lévy, 1999; Shaikh, 2009) have argued that it is not valid to allow the utilization rate to vary in comparisons across long-run equilibria. According to this critique, utilization should adjust to a desired or normal rate in the long run through some kind of adjustment mechanism, which could involve endogenous responses of firms' investment spending, corporate retention ratios (which affect the propensity to save out of profits) or central bank monetary policy. In response, defenders of the neo-Kaleckian approach have countered that the utilization rate may be variable in the long run within some range, the 'normal' rate of utilization may vary endogenously in response to actually experienced utilization (this is called 'hysteresis' in the utilization rate), and a unique long-run equilibrium utilization rate may not exist (see Lavoie, 1996; Dutt, 1997). Other economists have expressed scepticism about the whole exercise of long-run analysis in this context (for example, Chick and Caserta, 1997), a viewpoint that clearly coincides with Kalecki's own view that 'the long-run trend is but a slowly changing component of a chain of short-period situations; it has no independent entity'

(Kalecki, 1971b, p. 165). These debates are surveyed in Chapter 6 (see also Lavoie, 2014, pp. 387–405).

Also, all the neo-Kaleckian macro models covered in this chapter take unit labour costs and (except in the open economy model) firms' markup rates as exogenously given. However, there are likely to be feedbacks from utilization and growth – both of which affect the unemployment rate inversely – to the bargaining strength of workers and their ability to win wage increases in labour contracts. And when workers win wage increases, the ability of firms to pass these wage increases on to consumers in the form of higher prices may depend on market structure (the degree of competition or concentration), the state of demand (booming or depressed), openness to international competition, the level of the real exchange rate and other economic conditions. Although we have taken labour productivity as exogenously given in this chapter, it is likely to vary endogenously in the short run as a result of changes in capacity utilization in the presence of overhead labour, as shown in Appendix 4.1. In the medium and long run, productivity can also be affected by the growth rates of output (which can affect the rate of adoption of new technologies) and real wages (because of inducements to labour-saving technical change).[52] Depending on how wages, prices and productivity all respond to changes in utilization, growth and each other, realized markups and profit shares will change endogenously. These issues are addressed in the models of 'conflicting claims' inflation that are covered in Chapter 5 as well as neo-Kaldorian models of cumulative causation in Chapter 8.

Furthermore, the two-class structure (workers and capitalists) assumed in the basic neo-Kaleckian models – similar to their classical-Marxian and neo-Keynesian predecessors – is too simple to describe a modern society. At a minimum, it is essential to divide labour into production workers and professional-managerial employees (scientists and engineers, corporate managers and so on) and to distinguish firms as institutions from the financial investors ('rentiers') who own their equity (as stockholders) or lend them funds (via bond purchases or bank loans). These sorts of distinctions have been made in various extensions of neo-Kaleckian models, a small sample of which will be covered in Chapter 7.[53] In addition to recognizing heterogeneity of workers, neo-Kaleckian macro models need to incorporate the heterogeneity of firms (in terms of size, cost structure, oligopoly power and so on) that was originally emphasized by Kalecki (1954 [1968]), Steindl (1952 [1976]) and Eichner (1976). Heterogeneous firms have now been incorporated in many branches of mainstream economics, including international trade theory (since Melitz, 2003), but to date there have been only a few efforts to meld them into heterodox macro models. A few notable exceptions include

Setterfield and Budd (2011) and Gouri Suresh and Setterfield (2015), both of which apply agent-based or multi-agent methods.

Last, but certainly not least, Kalecki and Steindl both emphasized financial constraints on firms' investment spending, and Steindl included financial variables such as firms' 'gearing ratios' (debt burdens) and 'own capital' (the part not encumbered by debt) in his original investment function, in addition to the utilization rate. To incorporate these financial factors, it is important not only to distinguish the parts of profits that are kept by firms from the parts that are paid out to rentiers as interest or dividends, but also to model the dynamics of debt accumulation and asset pricing over time. The dynamics of corporate finance were key to Minsky's (1986) theory of financial fragility. Moreover, the financial crisis of 2007–09 suggests the importance of taking into account household debt (that is, consumer and mortgage lending) as well as corporate debt. Although complete coverage of heterodox monetary and financial models would be beyond the scope of this book, some aspects of such an analysis will be covered in Chapter 7.[54]

## ? STUDY QUESTIONS

1) How do the neo-Kaleckian and neo-Keynesian models differ in their treatment of prices, capacity utilization and income distribution?
2) Do the neo-Kaleckian models imply a paradox of thrift, similar to what we saw in the neo-Robinsonian model in Chapter 3? Using the model with positive saving out of both wages and profits, evaluate the effects of an increase in either saving propensity $s_r$ or $s_w$ on the equilibrium levels of $u^*$, $r^*$ and $g^*$, and state any assumptions or conditions you need to sign your results (including, but not limited to, the stability condition). How do the results for the real wage and profit share differ from the corresponding results in the neo-Robinsonian model?
3) Same as question 2, but for the 'widow's cruse': consider the effects of an increase in the animal spirits of business firms that increases the intercept $g_0$ in the investment function (4.19). What is the key difference in the adjustment mechanism for the widow's cruse (and also for the paradox of thrift) in the neo-Kaleckian model as compared with the neo-Robinsonian one?
4) Demonstrate how the Kalecki–Steindl model of a closed economy predicts that a rise in the average markup rate, as a result of increased monopoly power of firms, will lead to sustained economic stagnation. Does the same result necessarily obtain in more general neo-Kaleckian models? Why or why not?
5) Are the results of the open economy model *qualitatively* changed if we use the Kalecki–Steindl investment function (4.19) in place of equation (4.44)? Analyse and discuss your results.
6) Suppose you wanted to estimate empirically whether a given country has wage-led or profit-led demand (output or utilization). Which parameters in the saving, investment and net export functions would you need to estimate and how would you determine whether the economy in question had wage-led or profit-led demand? What econometric issues could arise in attempting to conduct such an estimation?
7) Should the neo-Kaleckian models be considered useful for short-run analysis only, or can they be considered to represent long-run, steady-state equilibria as defined in Chapter 1? Discuss.

## NOTES

1 For much more detailed accounts of heterodox or post-Keynesian views on the theory of the firm and oligopolistic pricing, see Eichner (1976), Lee (1998) and Lavoie (2014).

2 In reality, most modern firms operate multiple plants and also operate within various 'value chains' or 'supply chains' that supply goods at different stages of production (raw materials, intermediate goods and final goods). We abstract from these complications here but recognize that a theory of the firm that omits them is a heroic simplification.

3 Kalecki himself and many of his followers have used an older terminology, in which variable costs (labour, raw materials) are referred to as 'direct' or 'prime' costs. Here, we use the terminology more commonly used in modern micro theory, which is likely to be more familiar to most readers.

4 This presentation largely follows Lavoie (2014, pp. 157–63), but with considerable simplifications and omitting some of the additional hypotheses about pricing that he also discusses.

5 Another related concept, often used in the industrial organization literature, is the price–cost margin ($PCM$), where cost is usually measured by $MC$ or $AVC$ (suppressing the $i$ superscripts for the firm here for simplicity). Since $MC = AVC$ as long as the firm has some excess capacity, it does not matter which of these we use, and we can see that $PCM = P/AVC = (1 + \tau)AVC/AVC = 1 + \tau$. Because $\tau$, $GPM$ and $PCM$ are all positively related to each other, for many purposes these concepts can be used interchangeably.

6 For critical discussions of Kalecki's views on pricing and monopoly power see, among others, Riach (1971) and Kriesler (1988).

7 This same idea is recognized in the literature on 'labour rents', according to which workers in certain industries are able to capture a portion of the 'oligopolistic rents'. See, for example, Krueger and Summers (1988), Katz and Summers (1989) and Blanchflower et al. (1990).

8 A classic example is the US steel industry between the 1960s and early 2000s, in which the oligopolistic power of the largest firms was first battered by import competition and later undercut by small-scale, often non-union domestic producers called 'minimills'. See Blecker (1989b, 1991, 2008).

9 Kalecki (1943) supported Marx's belief that the distribution of income could be influenced by the 'class struggle' between workers and owners/managers of firms, but in Kalecki's approach the class struggle could affect relative shares only if it impacted the firms' markup rates.

10 These include Harris (1974), Asimakopulos (1975), Del Monte (1975), Steindl (1979), Rowthorn (1981), Taylor (1983, 1985, 1991, 2004), Dutt (1984, 1987, 1990), Amadeo (1986) and Kurz (1990, 1994), among others. Blecker (2002a), Hein (2014) and Lavoie (2014) provide comprehensive surveys and additional references.

11 We follow most of the neo-Kaleckian literature cited above in abstracting from raw materials costs. Some exceptions include Taylor (1983) and Ribeiro et al. (2017a, 2017b), who focused on imported raw materials or intermediate goods in models for developing countries or the global 'South' (the latter are covered in sections 7.2.3 in Chapter 7 and 10.6 in Chapter 10). The literature has varied more in regard to overhead labour, which was included in the models of Harris (1974), Asimakopulos (1975), Rowthorn (1981) and Lavoie (1995b), among others. Nevertheless, most other neo-Kaleckian models since the 1980s have ignored overhead labour. The theoretical implications of including overhead labour are considered in Appendix 4.1, while Chapter 5 will discuss how the omission of overhead labour has affected the interpretation of empirical estimates of these models.

12 Although we will not focus on it in this chapter, one can also find the analogous inverse consumption–growth relation for the neo-Kaleckian model:

$$c = \frac{1}{a_0} - \left(\frac{a_1}{a_0}\right)\left(\frac{g}{u}\right)$$

13 Lavoie (1995b) notes that this investment function resembles one used earlier by Kaldor (1957).

14 The general approach to the investment function assumed here, in which investment depends fundamentally on a demand variable and profitability variable, has been validated in a large empirical literature, albeit with some (potentially non-trivial) differences in the specifications. These empirical studies include Fazzari et al. (1988), Fazzari (1993) and Chirinko et al. (1999, 2011), among others. These studies confirm that accelerator effects (usually measured by lagged output or sales growth rather than

the utilization rate) are always a significant positive factor in explaining investment, both statistically and economically. They also find that the measure of profits that affects investment is the corporate cash flow, or (approximately) after-tax gross retained earnings, as hypothesized by Minsky. Blecker (2007, 2016c) observes that the cash flow effects in these studies are really short-run in nature, since they operate by relieving financial (liquidity) constraints but do not affect desired capital stocks in the long run. Finally, these studies tend to find a statistically significant, but economically small, negative effect of the 'user cost of capital', which incorporates tax-adjusted costs of borrowed funds as well as relative prices of capital goods.

**15** The assumption that $g_0 > 0$ is only required here because of some of our simplifying assumptions, which imply that (as will be seen below) positive solutions for the rates of utilization, profit and growth require us to assume this. In more general models, for example including a government sector or foreign trade, $g_0$ could have any sign.

**16** See the discussion of Marglin and Bhaduri (1990) in section 4.4.2 below for an effort to reconcile the Robinsonian and Kaleckian approaches to the investment function.

**17** See the discussion of the accelerator principle in Chapter 3, section 3.2.

**18** Del Monte (1975) constructed a dynamic neo-Kaleckian model using a true accelerator effect. That is, he specified the investment function as $g = f(r, \Delta Y/Y)$ and obtained results that are equivalent to those we will obtain in this section in comparisons across steady states. This article, which was published in Italian, unfortunately had little impact on the subsequent literature, which was written mainly in English.

**19** The $g_0$ coefficient can also be thought of as incorporating the role of fiscal policy, for example public investment expenditures, albeit in a very simple way (and with no attention to how such expenditures are financed). One could also think of $g_0$ as including the level of the fiscal deficit relative to the capital stock.

**20** Thus, Steindl's analysis influenced neo-Marxian theories of monopoly capitalism, for example Baran and Sweezy (1966) and Foster (2014).

**21** A counter-argument is that any such oligopolistic position could eventually be eroded by the intrusion of new sources of external competition, such as imports, or through Schumpeterian innovation that radically changes either the production process or the nature of the products in ways that undermine existing oligopolistic advantages. (Think about what happened to Kodak film after the invention of digital photography.) One can also see that industries may pass through different phases of increased and reduced concentration followed by re-concentration (where the latter now often occurs at a global level).

**22** This derivative can be found either by partial differentiation of (4.25), or else by applying the chain rule to (4.17) using (4.22) and (4.28).

**23** See also Lavoie (1995b, 2014) and Foley and Michl (1999) for further discussion of this paradox.

**24** See section 4.4.2 below for further discussion of this point and a more formal mathematical statement.

**25** Marglin and Bhaduri (1990) referred to this investment function as embodying a 'strong accelerator condition' in the assumption that $g_2 > 0$, which they argued might not hold. This is an important point because the signs of all three derivatives (4.28) to (4.30) depend crucially on this assumption. However, Marglin and Bhaduri do not deny that the paradox of costs is possible; they rather argue that the opposite case is also possible. See section 4.4.2 below.

**26** Alternatively, this derivative can be obtained by partially differentiating (4.29), or by applying the chain rule to (4.24) and using (4.31) and (4.32).

**27** Other critiques, alternatives and extensions – some of which reject some of these fundamental assumptions – will be covered in Chapters 5 to 7.

**28** One of the present authors developed a model with positive saving out of wages in an appendix to his doctoral dissertation (Blecker, 1986), but he did not publish models with this feature until much later (Blecker, 1999, 2002a). Taylor (1990) included positive saving out of wages, although without much emphasis, and the implications were developed further by Mott and Slattery (1994b).

**29** For neo-Kaleckian and related models that incorporate fiscal policy, see Mott and Slattery (1994a), Blecker (2002a), Isaac (2009), Palley (2013a) and Tavani and Zamparelli (2017).

**30** Here, the stability condition only tells us that $s_r - s_w - g_1 > (g_2 - s_w)(a_1/\pi)$, and since $g_2$ could be either greater or less than $s_w$, it is possible that $s_r - s_w - g_1 < 0$ (in which case $\partial u^*/\partial \pi > 0$) and yet the model could still be stable.

**31** Also, the higher is $s_w$, the more likely it is that $s_w > g_2$, in which case the stability condition in this model can be satisfied while $s_r - s_w - g_2 < 0$, as is required for profit-led demand.

**32** The results for the profit rate are not shown for reasons of space, but the sign of $\partial r^*/\partial\pi$ is also ambiguous.

**33** Hein (2014, 2017) calls this model 'post-Kaleckian', but we see it merely as a different species of a neo-Kaleckian model with a different investment function compared to the Kalecki–Steindl version, and hence we name it rather for its progenitors.

**34** In the case where demand is profit-led ($\partial u^*/\partial\pi > 0$), the profit rate and growth rate are always increasing in the profit share, so growth is definitely profit-led as well in the Bhaduri–Marglin model.

**35** For example, Blecker (2002a) shows that profit-led demand cannot occur in a Bhaduri–Marglin model if the investment function is linear with all positive coefficients ($g = h_0 + h_1 u + h_2\pi$, where $h_0, h_1, h_2 > 0$), but it can occur with a 'Cobb–Douglas' functional form ($g = A\pi^{\alpha}u^{\beta}$) where $A > 0$ is a constant and $\alpha$ and $\beta$ are positive exponents, provided that $0 < \beta < 1$ (for stability) and $\alpha > 1$ (high profit-share elasticity of investment). The reader should prove these results as an exercise.

**36** See, for example, Stockhammer et al. (2011), Ballinger (2013), Schoder (2013) and Onaran and Galanis (2012).

**37** To keep the discussion focused on the variables of main policy interest, we confine our classification scheme here to the results for utilization and growth ($\partial r^*/\partial\pi$ and $\partial g^*/\partial\pi$). In regard to the profit rate, $\partial r^*/\partial\pi$ always has the same sign as $\partial g^*/\partial\pi$ in some of the simpler models (for example, Kalecki–Steindl and Bhaduri–Marglin with no saving out of wages), but this need not be true in more complex models, such as the model with positive saving out of wages discussed above or the open economy model considered in the next subsection.

**38** Marglin and Bhaduri (1990) also postulated cooperative and conflictual cases of exhilarationist regimes, based on whether a profit-led increase in the growth rate of the capital stock also increases employment or not. That distinction has not received much attention in the literature, although perhaps it deserves to receive more.

**39** The Kalecki–Steindl approach was extended to open economy issues by Dutt (1984, 1988, 1990), but in models that (as a result of various simplifying assumptions) could only generate wage-led outcomes. Blecker (1989a) was the first to show the possibility of profit-led demand and growth in an open economy in which output is demand-determined by introducing a flexible markup rate and relative price effects on net exports into a Kalecki–Steindl modelling framework. Bhaduri and Marglin (1990) also discussed an open economy extension of their model, while Blecker (1999, 2002a, 2011) later extended his work to more general neo-Kaleckian models. The presentation in this section is most similar to Blecker (2002a). Von Arnim et al. (2014) present a two-country version of an open economy neo-Kaleckian model.

**40** Razmi (2016b) and Ros (2016) present alternative open economy models based on a small country framework (intended for developing countries), in which the country is a price-taker for tradable goods. In these models, a wage increase in the tradable goods sector is generally contractionary because it squeezes the profit margins of firms that cannot pass through the increases in unit labour costs into prices (since their prices are exogenously fixed), and the resulting fall in profits in turn diminishes investment. The Razmi and Ros models have more classical-Marxian or neo-Robinsonian rather than neo-Kaleckian roots, as those models assume that output is proportional to the capital stock in the modern or traded goods sector (utilization in this sector is implicitly fixed at a normal rate). They are also in the tradition of two-sector analysis in dual economy or 'structuralist' models (Taylor, 1979, 1983), as they contain a separate sector producing non-traded goods and services. See Chapter 10 for Razmi's and Ros's work on models of balance-of-payments-constrained growth.

**41** See equation (4.12) above, and note that the wage share is $1 - \pi = 1/(1 + \tau)$.

**42** Note that the higher is $\eta$, that is, the more the markup (or price–cost margin) adjusts, the lower is the degree of 'pass-through' of a change in the exchange rate into prices of domestic products.

**43** Alternatively, if we specify foreign prices as also incorporating a markup on unit labour costs, $P_f = (1 + \tau_f)W_f a_{f0}$, we could measure relative labour costs more directly as $W_f a_{f0}/Wa_0$, and the real exchange rate could be decomposed into relative markup factors and labour costs (home/foreign): $EP_f/P = [(1 + \tau_f)/(1 + \tau)](EW_f a_{f0}/Wa_0)$. However, since we are modelling a small country that takes foreign prices as given and we don't model the determination of foreign markups, we don't use this specification here.

**44** The Marshall–Lerner condition is the condition for a real depreciation of the currency (a rise in $EP_f/P$) to improve the trade balance. If trade is initially balanced and supply curves for both exports (home products) and imports (foreign products sold domestically) are infinitely elastic (as assumed in the neo-Kaleckian model), the Marshall–Lerner condition requires that the sum of the relative price elasticities of export and import demand must exceed unity in absolute value. Using the notation to be introduced in Chapter 9, this means $|\varepsilon_X + \varepsilon_M| > 1$. If trade is not initially balanced, this condition becomes more complicated (see Appendix 9.1), but it remains true that these price elasticities must be sufficiently high in absolute value for a depreciation to improve the trade balance.

**45** We are indebted to Arslan Razmi for pointing out this anomaly in the formulation, but we use this specification because of its analytical simplicity and because it is commonly used in the neo-Kaleckian literature. For a more rigorous model of the trade balance in which the export–capital ratio is explicitly held constant, see Blecker (1989a).

**46** Similar specifications were used by Blecker (1996, 2002a).

**47** The stability condition is found by setting $EDG = g + b - \sigma$ and, after making the appropriate substitutions, finding the condition for $\partial EDG/\partial u < 0$.

**48** We do not derive the impact of changes in $\mu$ and $z$ on $r^*$ and $g^*$ here for reasons of space, but the signs of the corresponding partial derivatives are also ambiguous, and the intuition is a combination of what has been discussed earlier for the closed economy models and what is discussed here for the impact of the induced shifts in international competitiveness.

**49** The empirical literature, which will be reviewed in detail in Chapter 5, often finds precisely that trade effects are larger in more open economies, as a result of which those countries are more prone to be profit-led than less open economies. Most of the empirical literature that has studied this issue has taken the profit share as exogenously given and has failed to identify the effects of different causes of shifts in that share, but we can infer from these sorts of results and from the way that the wage share is often measured as real unit labour costs that the studies are largely finding effects of variations in $z$ rather than $\mu$.

**50** For a sceptical view of the argument that the entire global economy must be wage-led, see Razmi (2018).

**51** Krugman and Taylor (1978) showed that a devaluation is contractionary in a model of a country that, by construction, has a strictly wage-led economy. The contrary case of a devaluation being expansionary in a profit-led economy was demonstrated by Blecker (1999).

**52** For neo-Kaleckian and related models with endogenous labour productivity, see (among others) Dutt (2006a), Rada (2007), von Arnim (2011) and Storm and Naastepad (2012).

**53** Models that make these sorts of distinctions include Onaran et al. (2011), Hein (2012a), Palley (2017) and Vasudevan (2017), among many others.

**54** Without attempting to be comprehensive, a few key references include Godley and Lavoie (2007), Charles (2008), Hein (2012a, 2012b, 2014), Isaac and Kim (2013), Setterfield and Kim (2017) and Nikolaidi and Stockhammer (2017) – all of whom provide additional citations.

# Appendix 4.1 The profit share with overhead labour in the short run

For long-run analysis, it is sensible to assume that employment is proportional to output, although of course this proportion is likely to decrease gradually over time as labour productivity increases as a result of technological improvements (as well as increased organizational and managerial efficiencies). However, in the short run it is more realistic to assume that firms employ some 'overhead' labour, which does not vary in proportion to the current level of output. Overhead labour is a broad category that can include supervisors, managers, executives, professionals, maintenance workers and office staff. As we will see, the presence of overhead labour dramatically changes our solution for the profit share and makes it a positive function of the level of output or rate of capacity utilization in the short run.[a]

For simplicity, we will assume that employment (measured in worker-hours) of production workers $L_0$ is proportional to current output, while employment of overhead employees is fixed at $L_1$ in the short run. Thus, total employment is

$$L = L_0 + L_1 \tag{4A.1}$$

where $L_0 = a_0Y$, $a_0$ is the fixed ratio of production workers to output, $Y$ is real output and $L_1$ is exogenously given. National income (which must equal the nominal value of output) equals the sum of total nominal profits $\Pi$ plus wages $WL$ (where 'wages' includes all labour income, assumed to be paid at the uniform nominal rate $W$ per worker-hour for all workers for simplicity):

$$PY = \Pi + WL \tag{4A.2}$$

The price level $P$ is determined by a markup over average variable costs ($AVC$), which equal unit labour costs for production workers only ($AVC = Wa_0$):

$$P = (1 + \tau)Wa_0 \tag{4A.3}$$

Thus, $\tau > 0$ is a gross markup rate that must cover overhead labour costs as well as provide net profits for the firm.

Combining these equations, we can solve for the gross profit share as

$$\pi = \frac{\Pi}{PY} = \frac{\tau - (L_1/a_0Y)}{1 + \tau} \tag{4A.4}$$

which reduces to equation (4.12), $\pi = \tau/(1 + \tau)$, if there is no overhead labour ($L_1 = 0$). For any positive level of overhead labour ($L_1 > 0$), and also holding the markup rate $\tau$ and production worker labour coefficient $a_0$ constant, the profit share is an increasing function of output ($\partial\pi/\partial Y > 0$). Given that the capacity utilization rate is defined as $u = Y/Y_K$, output can be written as $Y = uK/a_1$, where the capital stock $K$ and ratio of capital to full-capacity output $a_1 = K/Y_K$ are both taken as exogenously given in the short run. Substituting this expression for $Y$ into equation (4A.4), we can also see that $\partial\pi/\partial u > 0$, or in other words, the profit share varies procyclically.

In addition, overall labour productivity (the average product of labour for all employees, both overhead and production workers, $Q = Y/L$) is an increasing function of output or utilization in this model:

$$Q = \frac{Y}{L} = \frac{Y}{a_0Y + L_1} \tag{4A.5}$$

where $\partial Q/\partial Y > 0$ as long as $L_1 > 0$. Multiplying the right-hand side of (4A.5) by $1 = (1/Y_K)/(1/Y_K)$ and using the definition $u = Y/Y_K$, we can also see that $Q = u/(a_0u + L_1/Y_K)$ and, holding $Y_K$ constant and assuming $L_1 > 0$, $\partial Q/\partial u > 0$.

Of course, output (or utilization) is an endogenous variable, which we can assume is determined by saving-equals-investment equilibrium. Using an extremely simple specification, assume that profits are saved at the rate $s_r$ ($0 < s_r < 1$), there are no savings out of wages, and real investment is exogenously given at $\bar{I}$ in the short run.[b] Goods market equilibrium therefore requires

$$s_r\Pi = P\bar{I} \tag{4A.6}$$

Using the previously given equations for the price level, national income and employment in this equilibrium condition, we can solve for equilibrium output

$$Y^* = \frac{(1 + \tau)\bar{I} + (s_rL_1/a_0)}{s_r\tau} \tag{4A.7}$$

and the reduced form solution for the gross profit share is

$$\pi^* = \frac{\tau\bar{I}}{(1 + \tau)\bar{I} + (s_rL_1/a_0)} \tag{4A.8}$$

Thus, as long as there is positive overhead labour ($L_1 > 0$), a positive shock to aggregate demand in the form of an increase in investment (a rise in $\bar{I}$) will increase the profit share: $\partial\pi^*/\partial\bar{I} > 0$.

All these results will be of special importance when we study empirical tests of the relationship between income distribution (the profit or wage share of national income) and measures of output or utilization in Chapter 5. To preview that discussion, empirical studies that do not take into account the procyclical behaviour of the profit share (in other words, the impact of aggregate demand shocks in raising the profit share) may be biased towards finding a positive short-run effect of the profit share on output or utilization (that is, profit-led demand), when in fact it is the latter variables ($Y$ and $u$) that have positive effects on the former ($\pi$) in the short run.

Notes:

a The model presented here is based on Harris (1974) and Asimakopulos (1975). Case 1 in Harris is essentially the same (aside from notational differences) as the Asimakopulos model, except that the former includes saving out of wages (which the latter treats in an appendix) whereas the latter includes an autonomous component of capitalists' consumption. We abstract from both of these complications here since they are not crucial for our purposes; we have also altered both of their notations to conform with the rest of this chapter (this includes reversing their definitions of $L_0$ and $L_1$).

b For a model with overhead labour and endogenous investment, see Rowthorn (1981). Harris (1974) included saving out of wages along with overhead labour.

# Part II

# Extended models of distributional conflict and cyclical dynamics

# 5

# Distributional conflict, aggregate demand and neo-Goodwin cycles

## 5.1 Introduction

The neo-Keynesian and neo-Kaleckian models covered in the previous two chapters take a number of key distributional variables as exogenously given. For example, the neo-Keynesian models generally assume a given nominal wage, while the neo-Kaleckian models (except the open economy version) usually treat firms' markup rates (or price–cost margins) as exogenous. Neither type of model gives an explicit portrayal of inflation (continuous increases) in wages and prices (except in Marglin's synthesis of neo-Keynesian and neo-Marxian models, discussed in Chapter 3). Yet, all of the models covered so far recognize (some more implicitly, some more explicitly) that wage- and price-setting in a capitalist economy are closely tied to the distribution of income between labour and capital and that this distribution is rooted in class conflict over the division of the social product.

This chapter will first address the connections between wage-setting, price inflation and income distribution using what we will call a 'conflicting claims' approach. Since a comprehensive discussion of alternative theories of inflation would be beyond the scope of this book, this chapter will concern itself primarily with models that treat inflation in conjunction with income distribution, which can then be used to analyse the dynamic interactions between distribution and aggregate demand (or capacity utilization). Based on this combined framework, the chapter will then review and evaluate the empirical literature that has attempted to estimate demand–distribution linkages, especially in regard to whether aggregate demand is wage-led or profit-led (as discussed in Chapter 4). One branch of theory and empirics on this topic has adapted the neo-Marxian framework of 'Goodwin cycles', covered in Chapter 2, to generate models of business cycles that contain more neo-Kaleckian features (for example, by using the utilization rate instead of the employment rate as the measure of economic activity). Therefore, this chapter will cover what may be called the 'neo-Goodwinian' approach; alternative explanations of cyclical fluctuations based on neo-Harrodian approaches – in which the distribution of income is sometimes (but not always) treated as endogenous – will be discussed in Chapter 6.

The chapter begins by presenting a series of models of inflation and distribution based on conflicting claims in section 5.2, starting with a very simple core model and proceeding to more complex versions. Ultimately, this section develops a distributive relationship or curve, which reflects how the wage share of national income responds to changes in aggregate demand (output or utilization). Section 5.3 then shows how the distributive relationship can be combined with a neo-Kaleckian model of aggregate demand to produce a dynamic analysis of medium-run shifts and adjustment dynamics in both distribution and output (capacity utilization), including (but not limited to) the neo-Goodwin cycle model (mathematical details for which are relegated to Appendix 5.2). Section 5.4 then reviews the empirical literature that has estimated econometric versions of these models and discusses their findings, implications and limitations. Section 5.5 concludes. Appendix 5.1 discusses how the conflicting claims approach to inflation can be used to derive a type of Phillips curve; a related model will be covered in Chapter 6.

## 5.2 Conflicting claims models of inflation and distribution

All of the heterodox theoretical approaches recognize the fundamental element of conflict that is inherent in the determination of distributional outcomes in a capitalist economy. In the classical, neo-Marxian and neo-Keynesian approaches, this is clear from the existence of an inverse wage–profit relationship for any given technology. In the neo-Kaleckian approach, it is clear from the inverse relationship between firms' markups and the real wages of workers (or the wage share). Building upon these foundations, heterodox macroeconomists have modelled wage- and price-setting behaviour – and the associated distributional outcomes – as reflecting *conflicting claims* of workers and firms (or the owners of labour and capital) over the total income generated in a society.[1] When the *ex ante* claims of workers and firms – reflected in the wages that the former bargain for and the prices that the latter set – would result in claims exceeding 100 per cent of national income, the pressure is relieved via inflation: continuous increases in nominal wages and prices that limit the real income actually received *ex post* by one or both groups until the sum of their realized claims equals the total available social product. As this inflationary process unfolds, the markups of firms, real wages of workers and shares of profits and wages in total income all adjust endogenously.

This conflicting claims approach to inflation and distribution has been adopted in a wide range of heterodox models (cited below) and has implicitly crept into some mainstream models of labour markets and the Phillips curve (for example, Ball and Moffitt, 2001; Blanchard, 2017). In a sense, the conflicting claims approach constitutes the heterodox alternative to mainstream concepts of 'aggregate supply' and the Phillips curve (the relationship between the price level or inflation rate and either the unemployment rate or output gap). Importantly, the conflicting claims approach explicitly links the analysis of inflation to the forces that determine the distribution of income, since (as we shall see) the equilibrium inflation rate is solved for simultaneously with an equilibrium for the wage (or profit) share of national income.

Before proceeding further, however, it is important to acknowledge what may seem to be the most glaring omission in this section – the role of money in the inflationary process. Our purpose here is not to give a complete account of all theories of inflation, but rather to present a theory of inflation that is compatible with the neo-Kaleckian models of aggregate demand covered in the previous chapter, and which can be used to provide an analysis of how inflation and income distribution are simultaneously determined in an economy characterized by chronic excess capacity and demand-driven output.

In any modern economy, the supply of money is clearly an endogenous variable, which accommodates rather than causes inflation. Monetary policy is still important, of course, but it operates by central banks setting or regulating interest rates – not by them setting or determining the quantity of money. This basic proposition is now accepted not only by heterodox monetary theorists, but also by mainstream macroeconomists who have replaced the exogenously given money supply with a monetary policy reaction function of the central bank in standard macro models (see Carlin and Soskice, 2005, 2009). Of course, the famous 'equation of exchange' $MV = PY$ must always hold, where $M$ is the money supply, $V$ is the velocity of money, $P$ is the price level and $Y$ is real output. But this equation is an identity or tautology, since velocity is *defined* as $V \equiv PY/M$. The real issue is the direction of causality, which does not start with the money supply $M$ but rather originates in the processes determining output $Y$ and prices $P$, with $M$ and $V$ adjusting.

The theory of endogenous money has been well developed elsewhere,[2] and to present it here (and debate the nuances of the alternative views) would be beyond the scope of this book. Our purpose is to present a view of inflation that does *not* depend on 'too much money chasing too few goods', and which does *not* accept that 'inflation is always and everywhere a monetary phenomenon' in a causal sense,[3] but instead shows how inflation can arise fundamentally out of *distributional conflict* between workers and firms – the endogenous determination of income distribution being our ultimate focus.

## 5.2.1 A simple model

Just to fix ideas, we start with a highly simplified model that displays very starkly the core principles of the conflicting claims approach in regard to both inflation and distribution.[4] For this simple model, we assume that labour productivity $(Q = 1/a_0)$ is exogenously fixed, so we can concentrate on the real wage defined as $w = W/P$, where $W$ is the nominal wage rate and $P$ is the price level. Wages are determined in a bargaining process between workers and firms, in which workers' objectives are formulated in terms of a real wage target but labour contracts are signed in terms of the nominal wage. Workers bargain for nominal wage increases in proportion to the gap between their target real wage $w_w$ and the actual real wage $w$ as specified in the following reaction function

$$\hat{W} = \varphi(w_w - w) \tag{5.1}$$

where $\varphi > 0$ is a speed of adjustment parameter and a circumflex $(\hat{\ })$ over a variable indicates a growth rate in continuous time (or difference in natural logarithms) unless otherwise stated.[5] The workers' target $w_w$ can be influenced by many factors, including the strength of labour unions, the type of labour market regulations, conventional norms of fairness, race and gender relations, and the unemployment rate; for the moment, we simply take this target level as exogenously given. Similarly, the speed of adjustment also depends on many institutional aspects of the labour market and labour bargaining, such as the frequency of contract renegotiations or wage increases and the ability of workers (or their unions) to win the increases they seek (which could be limited, for example, by threats of outsourcing or offshoring of jobs – see Rodrik, 1997; Setterfield, 2006a, 2007).

Firms in turn are assumed to have a target $\tau_f$ for their markup rate, which depends on underlying factors such as industrial concentration, product differentiation and competition policies. For present purposes, we will not model how this target is determined explicitly,[6] but simply take it as given. The firms' target for the markup then translates into a corresponding (implicit) target for the real wage 'desired' by firms: $w_f = 1/[a_0(1 + \tau_f)]$. If the actual real wage rises above (falls below) this implicit target of firms, the latter respond by raising prices more rapidly (more slowly) according to the reaction function

$$\hat{P} = \theta(w - w_f) \tag{5.2}$$

where $\theta > 0$ is the speed of adjustment of prices. Greater monopoly power of firms and weaker antitrust regulation would enable firms to target a higher markup (implying a lower $w_f$) as well as to raise prices more rapidly (higher $\theta$).

Conflict over the distribution of income arises as long as $w_w > w_f$, that is, as long as the real wage objective of workers is greater than the real wage that firms implicitly desire to pay in order to achieve their target markup. In this situation, it is not feasible for both social classes to achieve their objectives, which would require that total wage and profit income would exceed 100 per cent of the social product.[7] A medium-run equilibrium inflation rate will be reached when wages and prices increase at the same rate ($\hat{W} = \hat{P}$), which is equivalent to the real wage being constant ($\hat{w} = 0$). Setting (5.1) equal to (5.2), we can solve for the equilibrium real wage[8]

$$w^* = \frac{\varphi w_w + \theta w_f}{\varphi + \theta} \tag{5.3}$$

Then using (5.3) in either (5.1) or (5.2), we obtain the equilibrium inflation rate

$$\hat{P}^* = \hat{W}^* = \frac{\varphi\theta(w_w - w_f)}{\varphi + \theta} \tag{5.4}$$

In this solution, inflation is positive if and only if there is distributional conflict in the sense that $w_w > w_f$, and the bigger is the gap $(w_w - w_f)$ the higher is the equilibrium inflation rate. Also note that the markup rate becomes endogenous in this model; the equilibrium markup rate is monotonically inversely related to the equilibrium real wage (5.3) as follows

$$\tau^* = \frac{1}{a_0 w^*} - 1 \tag{5.5}$$

In general – assuming that both speeds of adjustment $\varphi$ and $\theta$ are finite – both social classes will be partially frustrated in the medium-run equilibrium as neither will be able to fully reach its distributional target. As shown in Figure 5.1(a), in this case the equilibrium real wage ends up in-between the

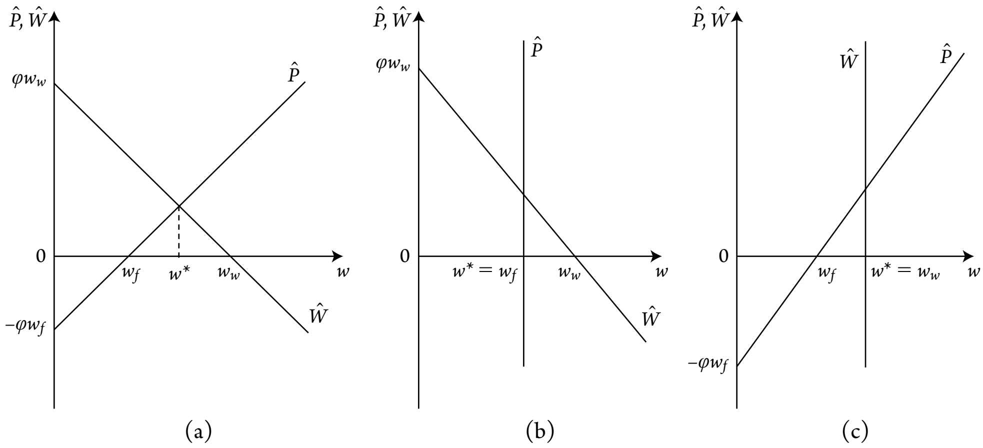



**Figure 5.1** Simple model of conflicting claims, inflation and the real wage (alternative cases)

two classes' target levels. Inflation is essentially the adjustment mechanism that reduces the real income claims of both workers and owners of firms so that they do not exceed 100 per cent of total output. The stronger is either social class, the closer the equilibrium real wage will be to that class's target and the more the other social class will be frustrated.

However, there are two special cases in which one class achieves its target fully and the other class has no influence on the distributional outcome. If firms are so powerful (say, because of strong monopoly power, high tariffs and no antitrust regulation) that they can instantly raise their prices to achieve their target markup, then they can essentially force workers to accept their implicit target for the real wage. Treating this as the case in which the speed of adjustment of prices $\theta$ becomes infinite, and applying L'Hôpital's rule to equation (5.3), we see that

$$\lim_{\theta \to \infty} w^* = w_f$$

In this case, the price reaction function becomes a vertical line as shown in Figure 5.1(b), and increased bargaining strength of workers (a higher or steeper $\hat{W}$ line) only causes higher equilibrium inflation but has no impact on the equilibrium real wage.[9] In parallel fashion, if workers are all-powerful (say, because of strong unions and full indexation to inflation) so that the speed of adjustment of wages $\varphi$ becomes infinite,[10] then

$$\lim_{\varphi \to \infty} w^* = w_w$$

This is the case depicted in Figure 5.1(c), where firms' price-setting only affects the inflation rate but workers always get their target real wage in equilibrium.

While this starkly simple model illustrates the core principles of the conflicting claims approach to inflation and distribution, it is obviously missing many factors that may affect both variables. Among other things, the simple model ignores the importance of demand factors, which may influence the objectives of both workers and firms and their respective abilities to raise wages or prices in pursuit of those objectives. Furthermore, the simple model ignores the role of inflationary expectations or wage indexation in influencing inflationary outcomes. We will start, however, by filling in another key omission in the basic model: the role of productivity growth.

### 5.2.2 Incorporating productivity growth

Increases in labour productivity ($Q = 1/a_0$) may affect the income claims (both desired and realized) of workers and firms in several ways. For firms, increasing labour productivity (equivalent to a falling labour coefficient, $a_0$) implies that nominal unit labour costs ($Wa_0$) rise more slowly for any given rate of increase in nominal wages $W$, thereby relieving pressure on realized profit markups and allowing firms to moderate price increases. For workers, increasing labour productivity means that they may not be satisfied with a constant target for the real wage; rather, workers are likely to pursue real wage objectives that rise in proportion to their productivity, in which case they will effectively target the wage share ($a_0 w$) rather than the real wage ($w$). In addition, some countries may have incomes policies or union movements that effectively seek to ensure that workers are partially or fully remunerated for their productivity improvements, or labour contracts may be written to create incentives for worker effort; in such cases, nominal wage increases could be tied directly to productivity gains. Furthermore, to the extent that competition compels firms to reduce prices (or raise them more slowly) in response to productivity increases, more rapid productivity growth can raise real wages further by lowering prices for workers' consumption goods.

To incorporate these possibilities, we will recast the distributional side of our model in terms of the wage share rather than the real wage.[11] Defining the wage share as $\psi = 1 - \pi = a_0 w = a_0 W/P$, distributional equilibrium is achieved when the wage share is constant, which requires

$$\hat{\psi} = \hat{W} - q - \hat{P} = 0 \tag{5.6}$$

where $q = \hat{Q} = -\hat{a}_0 > 0$ is the rate of labour productivity growth. The rates of nominal wage and price inflation are determined by the following reaction functions

$$\hat{W} = \varphi(\psi_w - \psi) + \beta q \tag{5.7}$$

$$\hat{P} = \theta(\psi - \psi_f) - \gamma q \tag{5.8}$$

Here, $\psi_w$ is the workers' target for the wage share, $\psi_f$ is the firms' implicit target for the wage share (equal to $1/(1 + \tau_f)$, where $\tau_f$ is the firms' target markup), $\beta$ measures the degree to which productivity gains are reflected in nominal wage increases directly through the bargaining process, and $\gamma$ is the degree to which firms are forced by competition to pass through productivity gains in the form of slower price increases. Note that this specification assumes that productivity growth affects the rates of wage and price adjustments, but not the target wage shares of either workers or firms.

In the next subsection, we will consider a model in which the productivity growth rate $q$ is endogenous, but here we make a first pass at productivity issues by taking this rate as an exogenous variable. Substituting equations (5.7) and (5.8) into (5.6), we can solve for the medium-run equilibrium wage share

$$\psi^* = \frac{\varphi\psi_w + \theta\psi_f - (1 - \beta - \gamma)q}{\varphi + \theta} \tag{5.9}$$

In this solution, the effect of an increase in productivity growth $q$ on the wage share depends inversely on the sign of $(1 - \beta - \gamma)$, and is ambiguous in general. If there is sufficient pass-through of productivity gains into slower price increases and sufficient productivity bargaining in wage-setting such that $\beta + \gamma > 1$, then $\partial\psi^*/\partial q > 0$ and faster productivity growth redistributes income to workers (although the negative employment effects of faster productivity growth are not considered here, so the overall benefits to workers remain uncertain[12]). But if those forces are weak enough so that $\beta + \gamma < 1$, then $\partial\psi^*/\partial q < 0$ and faster productivity growth redistributes income to owners of capital. Note that the absence of any such pass-through or productivity bargaining ($\beta = \gamma = 0$) is sufficient (but not necessary) for productivity growth to have a negative effect on the wage share, while in the borderline case where $\beta + \gamma = 1$ it has no effect.

The model of conflicting claims with exogenous productivity growth is graphed in Figure 5.2. With positive productivity growth, there is now a wedge between the equilibrium rates of wage and price inflation, since

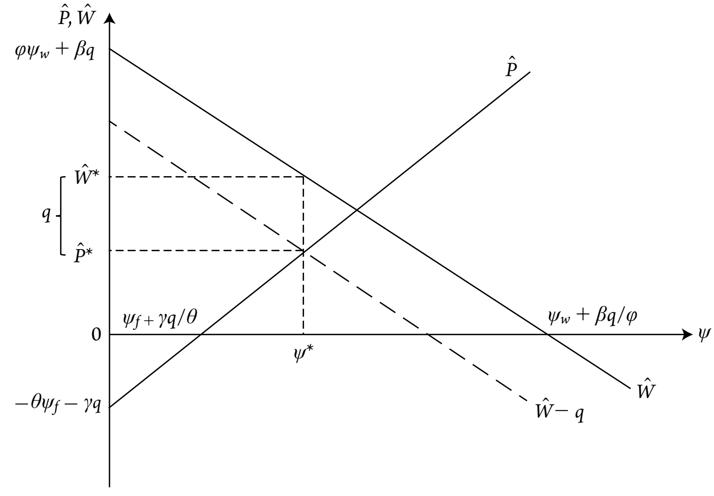



**Figure 5.2** Wage and price inflation and the wage share with exogenous productivity growth

the medium-run equilibrium condition (5.6) requires $\hat{P} = \hat{W} - q$. Thus, medium-run equilibrium is reached where the $\hat{P}$ line intersects $\hat{W} - q$, which lies below the $\hat{W}$ line. Substituting equation (5.9) into (5.8) and rearranging, we find the equilibrium price inflation rate

$$\hat{P}^* = \frac{1}{\varphi + \theta}[\varphi\theta(\psi_w - \psi_f) - (1 - \beta + \gamma\varphi)q] \qquad (5.10)$$

Once again, greater distributional conflict (a larger gap $\psi_w - \psi_f$) implies that higher equilibrium inflation is required to force workers and firms to accept income shares that sum to 100 per cent. Also, whether an increase in $q$ causes lower or higher equilibrium inflation depends on the sign of $-(1 - \beta + \gamma\varphi)$. Because it seems intuitively implausible that faster productivity growth would worsen inflation, we can assume that this effect is normally negative, which requires $\beta < 1 + \gamma\varphi$.

An important implication of this model is that the real wage rises at the same rate as productivity grows in a medium-run equilibrium with a constant wage share. This can be seen from the fact that the equilibrium condition (5.6) can be written as $\hat{W} - \hat{P} = q$, and it is also visible in Figure 5.2. Historically, there have been periods in which this was roughly true; indeed, a constant trend of relative share of labour (apart from cyclical fluctuations) was considered a 'stylized fact' of capitalist development by Kaldor (1961 [1989]).

However, in recent decades wage shares have exhibited a declining trend in many countries (see Karabarbounis and Neiman, 2014; Kiefer and Rada, 2015; Stockhammer and Wildauer, 2016).[13] In the context of the present model, the main causes of such a decline in the wage share must be either reductions in the bargaining power of labour (which would reduce the workers' target $\psi_w$) or increases in the monopoly power of corporations (which would raise their target markups $\tau_f$, and thereby lower $\psi_f$), or a combination of the two.[14] Since the decline in labour shares has coincided with a period of low or falling inflation rates, it seems likely that $\psi_w$ has fallen by more than $\psi_f$, so that the gap $(\psi_w - \psi_f)$ has been reduced. In addition, a decrease in the prevalence of productivity bargaining (a fall in $\beta$) or less competitive pressures for firms to pass through productivity gains into lower prices for consumers (a fall in $\gamma$) could also have contributed to a declining wage share, although only the former would contribute to lower inflation while the latter would tend to increase it.[15]

## 5.2.3 Linking distribution to demand: the distributive curve

Next, we incorporate demand effects into the analysis of wage- and price-setting in order to derive what we will call the 'distributive curve'. One of the most widely held tenets in macroeconomics, cutting across various heterodox and mainstream schools of thought, is the idea that workers' ability to bargain for higher wages is adversely affected by the rate of unemployment in the labour market. In Chapter 2, we examined classical-Marxian models in which changes in the real wage are driven by either the gap between the growth rates of labour demand and supply or (in the Goodwin cycle version) the ratio of labour demand to supply (that is, the employment rate). In a Keynesian framework, the Phillips curve embodies the idea that workers are able to win higher increases in nominal wages when the unemployment rate is lower.[16] In many formulations, the unemployment rate specifically affects the workers' target for the real wage or wage share (see Ball and Moffitt, 2001; Setterfield and Lovejoy, 2006; Stockhammer, 2011), and we will follow this approach here.

In order to maintain a simple and tractable model (and consistent with most of the literature on this type of model), we will invoke Okun's law to postulate that the unemployment rate is inversely related to the capacity utilization rate, $u = Y/Y_K$. For simplicity – and although it is not strictly realistic – we can further assume that $u$ is directly proportional to the employment rate $e = L/N$, as defined in earlier chapters, and of course $e$ equals one minus the unemployment rate.[17] On this basis, we can specify the workers' target wage share as an increasing (and linear) function of the utilization rate:

$$\psi_w = \lambda_0 + \lambda_1 u \tag{5.11}$$

The intercept term $\lambda_0 > 0$ reflects the institutional factors that affect workers' bargaining objectives, such as social norms, the strength of labour unions, the manner of bargaining (centralized or decentralized) and various aspects of labour regulation (including, but not limited to, minimum wage laws, the ease of firing workers and the generosity of unemployment insurance). The more these institutional factors favour workers, the higher is $\lambda_0$. In contrast, the slope parameter $\lambda_1 > 0$ reflects the degree to which higher utilization of capacity (corresponding to a lower unemployment rate) boosts workers' bargaining strength and induces them to set a higher target for the wage share.

In addition, we assume that in a high-inflation environment, workers might be able to achieve some degree of indexation of their nominal wages to price inflation, denoted by $\alpha$ $(0 \leq \alpha \leq 1)$. If inflation is typically low and/or labour is very weak, we would expect no indexation ($\alpha = 0$); the other extreme, which might occur if inflation is very high and labour is very strong, would be full indexation ($\alpha = 1$). Of course, intermediate cases are also possible. Including a term for indexation in the nominal wage reaction function, but omitting direct productivity bargaining (thus assuming $\beta = 0$) for simplicity,[18] we replace equation (5.7) with

$$\hat{W} = \varphi(\psi_w - \psi) + \alpha\hat{P} \tag{5.12}$$

Then, substituting equation (5.11) for the target wage share into (5.12), the nominal wage adjusts according to

$$\hat{W} = \varphi(\lambda_0 + \lambda_1 u - \psi) + \alpha\hat{P} \tag{5.13}$$

On the pricing side, the utilization rate could affect firms' targets for the markup rate and profit share in various ways. For example, if firms follow target-return pricing (as discussed in subsection 4.2.2 in Chapter 4), the relationship would be inverse, as firms would try to raise profits per unit when sales are slack. On the other hand, more robust demand conditions could enhance firms' ability to raise prices without losing customers, leading to a positive relationship. Also, we have not considered non-labour costs explicitly in this model, but assuming that these other costs (for example, prices of energy and other raw materials) would tend to rise as economic activity expands, this could further justify assuming that firms raise prices faster (relative to unit labour costs) when utilization is higher.[19]

Again, using a linear function for mathematical simplicity, we can write the firms' target *profit* share (written as one minus their implicit target for the wage share, as defined earlier) as

$$1 - \psi_f = \eta_0 + \eta_1 u \tag{5.14}$$

where $\eta_0 > 0$ reflects firms' monopoly power and $\eta_1$ reflects the response of their profit markups to demand conditions (which can be either positive or negative, for the reasons just mentioned). Substituting $\psi_f$ as specified in (5.14) into the price reaction function (5.8), while ignoring direct productivity effects in price-setting (that is, assuming $\gamma = 0$) for simplicity, we obtain

$$\hat{P} = \theta[\psi - (1 - \eta_0) + \eta_1 u] \tag{5.15}$$

We will focus mainly on the case of $\eta_1 > 0$, but the opposite case could also be considered.

The last element we will introduce in this more complete model of conflicting claims is endogeneity of labour productivity growth, $q$. Theoretically, $q$ could depend on several variables in our macro framework, including the following:

- Capacity utilization $u$: As discussed in Appendix 4.1 in Chapter 4, productivity is an increasing function of output and utilization in the short run in the presence of overhead labour. Intuitively, this occurs because layoffs are not proportional to output declines when $u$ falls in a recession, and hiring is less than proportional to output increases when $u$ rises in a recovery. Strictly speaking, this implies a positive static relationship between $u$ and labour productivity $Q = Y/L$, but for present purposes it could be considered to imply a positive relationship between $u$ and the growth rate of labour productivity, $q = \hat{Q}$. In the longer term, such a positive relationship could also reflect the likelihood that firms invest more in new capital equipment when utilization is higher, thereby resulting in faster growth of labour productivity.
- The wage share $\psi$: A higher wage share could induce firms to invest more in labour-saving equipment, thereby making labour productivity grow faster when labour costs are higher.[20] Although incentives to invest in productivity-enhancing equipment could be attenuated if firms have the option of 'offshoring' production to lower-wage locations in response to high domestic labour costs, average productivity can still rise assuming that the most labour-intensive (lowest-productivity) activities are the ones most likely to be relocated. In addition, a higher wage share could boost workers' effort and thereby raise labour productivity, as recognized in the concept of 'efficiency wages'.[21] Efficiency wage effects are likely to be stronger in countries where industries are highly capital-intensive and use advanced technology, and weaker where industries are more labour-intensive and use standardized technologies.

- Output growth $y$: Since the pioneering work of Verdoorn (1949) and Kaldor (1961 [1989], 1966a [1989]), it has long been thought that productivity growth is an increasing function of output growth, at least in the manufacturing sector (see also Cornwall, 1977). Kaldor and later Kaldorian economists have attributed this to dynamic increasing returns to scale, which include not only traditional (static) economies of scale from indivisibilities in capital equipment and high fixed costs, but also induced innovation and adoption of new technologies. However, Kaldor–Verdoorn effects are likely to operate over a much longer time frame than the adjustments of wages, prices and distribution considered in this chapter, and therefore they will not be formally incorporated in the present model. Kaldor–Verdoorn effects will be covered extensively in relation to Kaldorian models of long-run growth in Chapters 8–10.

Therefore, combining the first two elements, we can specify the productivity growth equation as follows[22]

$$q = q_0 + q_1 u + q_2 \psi \tag{5.16}$$

where $q_0 > 0$ (assuming there is some underlying trend of productivity growth not attributable to $u$ or $\psi$, for example because of ongoing R&D efforts in advanced economies and technological catch-up in less developed economies) and $q_1, q_2 \geq 0$ (where the possibility of either of these being zero allows for cases where the corresponding effects are negligibly weak).

Substituting equations (5.13), (5.15) and (5.16) into the equilibrium condition (5.6) and solving for the wage share $\psi$, we obtain the equation that defines the distributive curve (DC). This equation depicts equilibrium wage share as a function of the capacity utilization rate as follows

$$\psi = \frac{\varphi\lambda_0 + (1-\alpha)\theta(1-\eta_0) - q_0 + [\varphi\lambda_1 - (1-\alpha)\theta\eta_1 - q_1]u}{\varphi + (1-\alpha)\theta + q_2} \tag{5.17}$$

This equation for DC does not give us a single-valued solution for the equilibrium wage share; rather, it describes an equilibrium relationship between the wage share $\psi$ and utilization rate $u$.[23] The slope of this DC relationship is ambiguous in general. Given that the denominator is likely positive,[24] the sign of the slope ($\partial\psi/\partial u$) depends on the sign of the term in brackets $[\cdot]$ in the numerator.

Several possible cases are depicted in Figure 5.3. Panel (a) depicts the case known as a 'profit squeeze', in which rising utilization raises the wage share

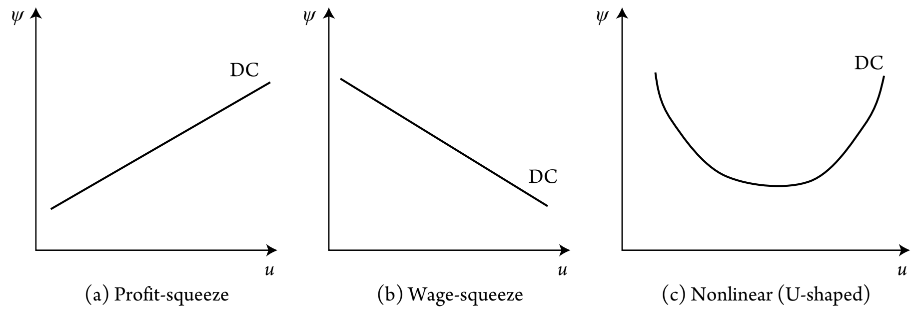



**Figure 5.3** Three alternative cases for the distributive curve (DC)

and reduces the profit share, which occurs if $\varphi\lambda_1 > (1 - \alpha)\theta\eta_1 + q_1$, or in other words, if the impact of higher utilization and employment on wage increases exceeds the impact on price increases and productivity growth. The profit-squeeze notion has a long and venerable history in heterodox macroeconomics, dating back to the neo-Marxian literature of the 1970s and 1980s (for example, Boddy and Crotty, 1975; Weisskopf, 1979; Bowles and Boyer, 1988), as discussed by Lavoie (2017). Panel (b) depicts the opposite case, which may (following Kiefer and Rada, 2015) be dubbed a 'wage-squeeze'. In this case, $\varphi\lambda_1 < (1 - \alpha)\theta\eta_1 + q_1$ so that rising utilization causes prices and productivity to rise faster than nominal wages, thereby lowering the wage share.[25] Finally, if the relevant parameters are not constant but vary with the level of $u$, various nonlinear shapes of DC become possible. Panel (c) shows one such possibility, the U-shaped DC curve found in an econometric study of the US economy by Nikiforos and Foley (2012). This U-shaped DC curve embodies a wage-squeeze at low levels of $u$ and a profit-squeeze at higher levels.[26]

The same fundamental determinants of income distribution that we saw in the simpler models of the previous sections (especially workers' bargaining power, here reflected in $\varphi\lambda_0$, and firms' monopoly power, reflected in $\theta\eta_0$) now operate by shifting the DC curve upward or downward in $u \times \psi$ space (that is, raising or lowering $\psi$ for any given $u$). Many variations on the solution (5.17) can also be considered, including exogenous productivity growth ($q_1 = q_2 = 0$), no indexation of wages ($\alpha = 0$), and full indexation ($\alpha = 1$). In general, the greater is the degree of wage indexation, the lower is the weight $(1 - \alpha)$ on firms' objectives and responses, and the greater is the influence of workers' behaviour. In the extreme case of full indexation, firms' behaviour disappears completely from equation (5.17), and workers have more power to achieve their target wage share subject only to productivity and demand conditions. Also, productivity bargaining for labour and productivity gains to consumers could be added back into the model, in which case each $q_i$ ($i =$ 0, 1, 2) term in (5.17) would be multiplied by $(1 - \beta - \gamma)$ and the effects of all these parameters would be reversed in sign if $\beta + \gamma > 1$. Thus, this model can be adapted to a wide variety of circumstances depending on the economic structure and policy regime of a given country in a particular historical period, while still reflecting the underlying logic of conflicting claims.

The solution for the inflation rate as a function of the utilization rate can be found by substituting equation (5.17) into the price reaction function (5.15) and simplifying, the details of which are left to the reader as an exercise. However, it should be noted that the resulting solution will incorporate all three of the fundamental determinants of inflation: distributional conflict, cost-push pressures and demand-pull factors, all in turn moderated by productivity growth but possibly exacerbated by indexation. This very general approach to inflation stands in contrast to some other approaches, such as the Marglin (1984a) model covered in Chapter 3, in which only one possible type of inflation (in that case, demand-pull) is included. Appendix 5.1 discusses how the conflicting claims approach can be extended to derive a Phillips curve model of inflation, and explains how such a Phillips curve analysis differs from the model of distribution and inflation considered in the rest of this chapter.

## 5.3 Dynamics of demand and distribution

We now turn to showing how the distributive relationship (DC curve) derived above can be combined with a neo-Kaleckian analysis of effective demand, as discussed in Chapter 4, to generate a model of the simultaneous determination of aggregate demand (the utilization rate) and income distribution (the wage share). To facilitate this analysis, it is helpful to recast the results of the neo-Kaleckian models in terms of the wage share instead of the profit share. In a model where profits and wages are the only two categories of income and there are no indirect taxes, the shares of profits and wages must sum to unity, so $\pi = 1 - \psi$ and $1 - \psi$ can be substituted for $\pi$ in all the equations of the neo-Kaleckian models from Chapter 4. The cases of wage-led and profit-led demand can then be represented by aggregate demand (AD) curves that slope upward or downward in $u \times \psi$ space, respectively, as shown in panels (a) and (b) of Figure 5.4. Nonlinear or nonmonotonic cases are also possible. For example (and it is only one possibility), panel (c) of Figure 5.4 shows an inverted U-shaped AD curve, which assumes that demand is wage-led at low rates of utilization and profit-led at high rates. This could occur, for example, if firms put a greater weight in their investment decisions on accelerator (demand) effects in a slump and profitability

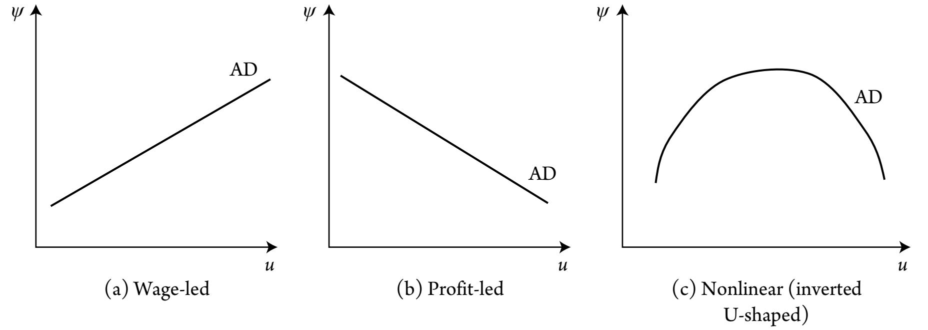



**Figure 5.4** Three alternative cases for the aggregate demand (AD) curve

considerations in a boom. Alternatively, although it is not shown in Figure 5.4, increments in utilization might provoke only small investment responses in a slump (when utilization is low) but larger responses in a boom, and conversely for profitability effects, in which case AD could have a U-shape (rather than an inverted U-shape). Clearly, the slope and shape of the AD curve are empirical questions.

An AD curve can then be combined with a DC curve of the type derived in the previous section to find the equilibrium solution(s) for utilization and the wage share (noting that multiple equilibria are possible, if at least one of the curves is nonlinear). However, modelling the equilibrium solution of the AD and DC relationships is not sufficient; we also need to specify the dynamics of adjustment and study their stability (or instability) properties and the possible emergence of cyclical behaviour. As will be seen, these dynamics depend crucially on the relative speeds of adjustment of output and prices: does the goods market clear (is saving–investment equilibrium reached) much faster than wages, prices and distribution adjust, or not? If nominal wages and prices are 'rigid' or 'sticky' in the short run, as assumed in most Keynesian and Kaleckian macro models, then output could adjust at a rate that is an order of magnitude faster than the speed of adjustment of distribution; in this case, we can use a static AD equation. Alternatively, if the adjustments in both output (utilization) and distribution (wages, prices and wage share) occur over relatively similar periods of time, then we need to use a dynamic specification of AD. These two alternative specifications will be covered in the next two subsections, respectively.

Before proceeding further, however, it is important to clarify how the terminology of wage-led and profit-led cases will be used in the rest of this chapter. There are at least three possible ways of applying this terminology in models that include a distributive relationship along with a demand relationship: (1) outcomes are wage-led (profit-led) if there is a positive (negative) association between $u$ and $\psi$ in comparisons across equilibrium states (intersections of AD and DC); (2) outcomes are wage-led (profit-led) if an upward shift in DC leads to an increase (decrease) in equilibrium $u$ (regardless of which way equilibrium $\psi$ changes); or (3) the terms wage-led and profit-led are applied only to what may be called the demand regime, that is, they are associated solely with a positive or negative slope of AD. Because the change in the equilibrium wage share is endogenous in the medium run in both definitions (1) and (2), it is not clear in what sense the change in distribution can be said to 'lead' the change in utilization.[27] Therefore, in the remainder of this chapter we will restrict ourselves to definition (3), and apply the terms wage-led and profit-led only to the cases in which AD slopes upward or downward (respectively) regardless of how the comparative dynamics play out in response to shifts in the underlying parameters. Moreover, we will see that even this last definition can be problematic when we allow for adjustments in potential output ($Y_K$) as well as actual output ($Y$) in driving changes in the utilization rate $u = Y/Y_K$.

Finally, one other point of clarification is that – similar to the neo-Kaleckian models in Chapter 4 – we will assume in the rest of this section that the equilibrium utilization rate is bounded only by zero and one ($0 < u^* < 1$) and there is no 'normal' utilization rate to which firms must adjust in the long run. In the models presented below, the equilibrium level of $u^*$ will emerge as the outcome of a dynamic interaction between utilization and distribution – that is, between demand-side and so-called supply-side (cost and price) factors – not as a result of aggregate demand factors alone. This dynamic equilibrium for $u^*$ is an endogenous result of the model; it is not an exogenous datum as the normal utilization rate $u_n$ is in some macro models (see Chapter 6 for further discussion of such models).

### 5.3.1 Dynamics with slow adjustment of income distribution

We begin with the commonly assumed case in which demand and output adjust relatively rapidly, while adjustments of nominal wages and prices as well as income distribution occur more slowly over much longer periods of time. In the real world, this could mean that output would adjust within just a few quarters or a single year, while distribution could take several years or longer to adjust, in response to any exogenous shock. Since we are using continuous time for mathematical convenience, we will represent this as instantaneous adjustment of utilization combined with gradual adjustment of the wage share. Assuming instantaneous adjustment of output implies that, in effect, the economy is always on the AD curve; it will be pulled either towards or away from the equilibrium point depending solely on the dynamics of adjustment of the wage share.

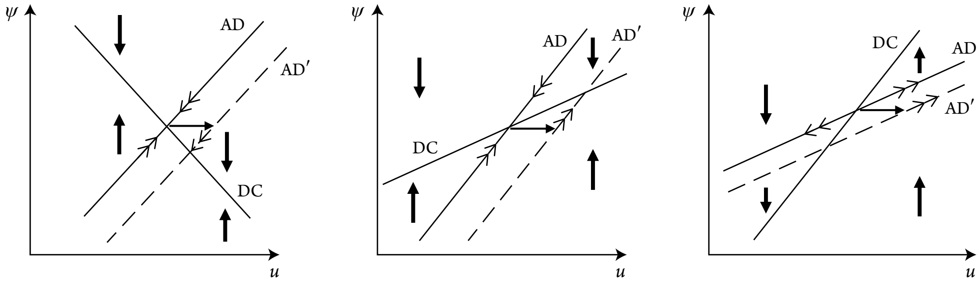
To summarize the many different versions of the neo-Kaleckian approach in a compact form, we will simply specify that aggregate demand (normalized by the capital stock) is an implicit function of capacity utilization, income distribution and other factors (such as fiscal policies, interest rates, animal spirits, financial positions). Then the goods market equilibrium condition that output equals aggregate demand can be written as

$$u = AD(u, \psi, Z_D) \tag{5.18}$$

where $AD(\cdot)$ is the aggregate demand function (where the level of aggregate demand is normalized by potential output, $Y_K$) and $Z_D$ is a vector of other determinants of aggregate demand. The partial derivative $\partial AD/\partial u = AD_u$ represents the marginal propensity to spend (on consumption and investment together, as well as any procyclical government spending) out of income, net of all 'leakages' due to saving, taxes and (in an open economy) imports. If the Keynesian stability condition holds, we can assume that $0 < AD_u < 1$ (the alternative case of $AD_u > 1$ is conceptually similar to Harrodian instability, which was introduced in Chapter 3 and will reappear in Chapter 6).

Equation (5.18) defines the AD curve. To find its slope in $u \times \psi$ space, we can totally differentiate this equation with respect to $u$ and $\psi$ and solve to obtain $\mathrm{d}\psi/\mathrm{d}u = (1 - AD_u)/AD_\psi$. Assuming Keynesian stability, the numerator of this expression is positive, so the slope has the same sign as $AD_\psi$. Thus, the cases of wage-led and profit-led demand can be represented by the signs of the partial derivatives $\partial AD/\partial\psi = AD_\psi > 0$ and $\partial AD/\partial\psi = AD_\psi < 0$, respectively.[28] Exogenous changes in the elements of $Z_D$ will shift AD to the right or left, depending on whether they are expansionary or contractionary.

Many different configurations of the AD curve (5.18) and DC curve (5.17) are possible, depending on their respective slopes. In all cases the intersection of AD and DC represents a simultaneous solution for the wage share and utilization rate. Implicitly, the inflation rate will also stabilize and the markup rate will adjust to an equilibrium level in such an equilibrium. Although multiple equilibria are possible if either (or both) of the curves is (are) nonlinear or nonmonotonic, we confine our discussion here to cases in which both curves are linear and monotonically increasing or decreasing, so that the equilibrium (assuming it exists) is unique. Since we are assuming that output (utilization) adjusts instantaneously, the economy is always on the AD curve, but it is not necessarily at the medium-run equilibrium point where AD




(a) Wage-led demand with a wage-squeeze (stable, overshooting)

(b) Wage-led demand with a weak profit-squeeze (stable, undershooting)

(c) Wage-led demand with a strong profit-squeeze (unstable)

Note: Bold vertical arrows show the direction of motion of $\psi$; lighter arrows show the dynamics following a positive shock to demand.

**Figure 5.5** Three alternative configurations of aggregate demand and the distributive relationship with slow adjustment of the wage share, all assuming wage-led demand

intersects DC. Hence, we need to focus on the adjustments of the wage share and how those will pull the economy towards (or push it away from) the medium-run equilibrium.

To see how the wage share adjusts, recall that the rate of change in the wage share is given by $\hat{\psi} = \hat{W} - q - \hat{P}$ and DC can be seen as the demarcation curve (nullcline) along which $\hat{\psi} = 0$. Furthermore, the direction of adjustment in $\psi$ is given by the sign of $\partial\hat{\psi}/\partial\psi$. Using equations (5.13), (5.15) and (5.16) for wage-setting, price-setting and productivity growth, respectively, under the assumptions made previously, this derivative is $\partial\hat{\psi}/\partial\psi = -[\varphi + (1 - \alpha)\theta + q_2] < 0$. In this case, $\psi$ is always falling at points above the DC curve and rising at points below it, as shown by the bold arrows in Figure 5.5. Thus, the 'own' adjustments in $\psi$ are self-stabilizing, but the overall stability of the economy also depends on the relative slopes of the two curves as we shall see.

Figure 5.5 illustrates three possible cases assuming that demand is wage-led; the corresponding cases for profit-led demand are left to the reader as an exercise. In panel (a), there is a wage-squeeze along with wage-led demand, so DC slopes down while AD slopes up. Since the economy must always be on AD in the short run, the movements along the AD curve are towards the equilibrium, which is therefore stable. The same is true in panel (b), where both curves slope upward (so there is a profit-squeeze along with wage-led demand) but AD is steeper. Although it is not drawn here, the equilibrium would also be stable in the borderline case of a horizontal DC curve (in which distribution would be independent of demand).

However, instability will result if both curves slope upward and DC is steeper than AD, as illustrated in panel (c) of Figure 5.5. In this case, the reactions of distribution to changes in demand are so extreme and self-reinforcing that they actually pull the economy away from the equilibrium. For example, suppose that the economy is initially at a point along AD which is above and to the right of the equilibrium point. Since AD lies below DC to the right of the equilibrium, $\psi$ must be increasing. However, because demand is wage-led, the rise in the wage share is expansionary, causing $u$ to rise further, which then causes $\psi$ to increase more, and so on in an explosive fashion. Of course, such an economy would eventually hit upon some constraint not considered in this model, such as a capacity constraint, a financial constraint (debt crisis) or a policy response (monetary contraction or fiscal austerity) to high and rising inflation. But the dynamics of such an economy would be chronically unstable so we would expect it to bounce between booms and busts with recurrent crises rather than growing smoothly.[29]

Slow adjustment of nominal and distributional variables also has implications for the dynamics of adjustment to a 'shock' that shifts one of the curves. If output and utilization adjust instantaneously while the wage share adjusts more slowly, then the wage share is a state variable that is given at any point in time and adjusts gradually according to whether $\psi$ is above or below the DC curve, as discussed above. Hence, if the AD curve shifts and the economy has to remain on the AD curve even in the short run, the economy will 'jump' from the original equilibrium to the point on the shifted AD curve corresponding to the initial level of $\psi$; then, the economy will move gradually along the new AD curve towards the new medium-run equilibrium. This behaviour can result in either undershooting or overshooting of the utilization rate in the short run, compared to its new medium-run level.

For example, in Figure 5.5, panels (a) and (b), we show the dynamics following an expansionary demand shock (such as a fiscal stimulus). In panel (a), AD shifts rightward to AD′, and since DC is downward sloping (wage-squeeze case) the economy overshoots in the short run following the shock. Because the new short-run equilibrium point is above the DC curve, $\psi$ begins falling, and with wage-led demand $u$ must then decrease (moving down and to the left along the new AD curve until the new equilibrium is reached). In contrast, in panel (b), where DC is upward sloping (profit-squeeze), the economy undershoots the new equilibrium in the short run. Then the new short-run equilibrium point on AD′ is below DC, so $\psi$ begins rising, which then causes $u$ to increase further given that demand is wage-led. Finally, in the unstable case depicted in panel (c), a demand shock moves the economy horizontally onto the new AD′ curve in the short run, and since the new short-run equilibrium point lies below DC, $\psi$ increases leading to further increases in $u$ and the result is explosive growth (until, presumably, some constraint is reached or a crisis results).

In contrast, if DC shifts (for example, because of a change in the monopoly power of firms that alters $\psi_f$ or a change in labour market institutions that alters $\psi_0$), no 'jumping' will occur. In the short run following a shock to DC, the economy remains at the initial equilibrium point (since $\psi$ does not change instantaneously), and then it moves gradually up or down along the AD curve towards the new medium-run equilibrium (or away from it, if the equilibrium is unstable) as dictated by the sign of $\hat{\psi}$. Illustrations of shifts in DC are left to the reader as an exercise, as are the dynamics if demand is profit-led.

## 5.3.2 Dynamics with similar speeds of adjustment: neo-Goodwin cycles and other cases

A different kind of dynamic analysis emerges if we assume that the adjustments of output and utilization towards their equilibrium values takes place at a qualitatively similar speed to the adjustments of wages, prices and distribution. This does not require exactly equal speeds of adjustment, only that the speeds of adjustment are not so different that one variable can be assumed to adjust instantaneously while the other adjusts slowly as in the previous subsection. With relatively similar speeds of adjustment, the AD–DC model implies dynamics that can, under certain assumptions discussed below, generate arguably realistic portrayals of business cycles. As will be seen, these cycles are a variant of the neo-Marxian Goodwin cycles we examined in Chapter 2 (section 2.8), but with some important differences. This variant was originally called a 'structuralist Goodwin model' by its progenitors (Barbosa-Filho and Taylor, 2006), but we will follow Stockhammer (2017) and call it the 'neo-Goodwinian' model or a model of 'neo-Goodwin cycles'.

To see the possibility of neo-Goodwin cycles and other types of dynamics, we need to convert the AD–DC model into a system of two simultaneous differential equations in $u$ and $\psi$.[30] By substituting equations (5.13), (5.15) and (5.16) into $\hat{\psi} = \hat{W} - q - \hat{P}$ and using the definition that $\hat{\psi} = \dot{\psi}/\psi$, we obtain

$$\dot{\psi} = (\omega_0 + \omega_1 u + \omega_2 \psi)\psi \qquad (5.19)$$

where $\omega_0 = \varphi\lambda_0 + (1-\alpha)\theta(1-\eta_0) - q_0$, $\omega_1 = \varphi\lambda_1 - (1-\alpha)\eta_1 - q_1$ and $\omega_2 = -[\varphi + (1-\alpha)\theta + q_2]$. The assumptions made earlier imply that $\omega_2 < 0$, which means that increases in $\psi$ reduce its own rate of increase and hence are self-stabilizing. The sign of $\omega_1$ depends on whether the response of wages to demand pressures dominates the responses of prices and productivity (in which case, $\omega_1 > 0$, which is the profit-squeeze case), or if the responses of prices and productivity dominate the response of wages ($\omega_1 < 0$, wage-squeeze). The sign of $\omega_0$ is ambiguous in general.

To derive a dynamic version of the AD curve, we will follow Barbosa-Filho and Taylor (2006) and replace the static aggregate demand function (5.18) with a specification that assumes gradual adjustment of the utilization rate.[31] By definition, the rate of change in the utilization rate is $\hat{u} = \hat{Y} - \hat{Y}_K$. Barbosa-Filho and Taylor specify linear differential equations for the growth rates of actual and potential output as follows

$$\hat{Y} = d_0 + d_1 u + d_2 \psi \tag{5.20}$$

$$\hat{Y}_K = b_0 + b_1 u + b_2 \psi \tag{5.21}$$

Assuming that output growth in (5.20) is determined by aggregate demand, $d_0$ represents the impact of all sources of 'autonomous demand' (such as government expenditures and exports) and other exogenous determinants of aggregate demand (interest rates, wealth effects, debt levels, business and consumer confidence and so on). We can assume that $d_1 < 0$ for Keynesian stability,[32] while the sign of $d_2$ is ambiguous and depends on whether demand is wage-led ($d_2 > 0$) or profit-led ($d_2 < 0$).

Since $Y_K = K/a_1$, changes in potential output $\hat{Y}_K$ are assumed to be driven by capital accumulation ($g = \hat{K}$) on the simplifying assumption of a constant ratio of capital to full-capacity output $a_1$. Thinking about equation (5.21) as largely reflecting investment behaviour, we see that it is analogous to a linearized version of the Bhaduri–Marglin investment function (4.37) in Chapter 4. We can assume $b_1 > 0$ due to a positive accelerator effect, as assumed in all versions of the neo-Kaleckian model (either Kalecki–Steindl or Bhaduri–Marglin). Because the effect of the wage share on investment is assumed to be negative in the Bhaduri–Marglin model, Barbosa-Filho and Taylor assume $b_2 < 0$.

However, there are three reasons why this last sign could be reversed. First, as we saw in Chapter 4, the effect of a rise in $\psi$ on the *equilibrium level* of $g = \hat{K} = \hat{Y}_K$ (taking into account indirect effects through endogenous changes in $u$) can potentially be positive in the presence of a sufficiently strong accelerator effect (as in the Kalecki–Steindl model, or even in the

Bhaduri–Marglin model under certain parameter values).[33] Second, the standard view that the wage share has a negative effect on investment assumes that all investment is corporate (business) investment, but in most countries a large portion of total investment is residential (housing). A higher wage share could have a positive effect on workers' demand for housing and hence on residential investment, in which case the net effect of the wage share on total investment is ambiguous even if the effect on corporate investment is negative.[34] And third, if we were to use the definition of potential output in terms of full employment of labour ($Y_N$) rather than full utilization of the capital stock ($Y_K$), then a higher wage share could induce more labour-saving technological innovations that would make labour productivity grow faster, hence increasing potential output in the former sense.[35] This is important because the two concepts of potential output are often not distinguished in empirical studies, and hence this third effect could be empirically important. Therefore, we will assume that $b_2$ could be either positive or negative.

Substituting equations (5.20) and (5.21) into $\hat{u} = \hat{Y} - \hat{Y}_K$ and using the fact that $\hat{u} = \dot{u}/u = (du/dt)/u$, we obtain

$$\dot{u} = (v_0 + v_1 u + v_2 \psi)u \tag{5.22}$$

where $v_i = d_i - b_i$ ($i = 0, 1, 2$) and AD is defined by $\dot{u} = 0$. It is easily seen that $v_1 = d_1 - b_1 < 0$ under our assumptions, while the sign of $v_2 = d_2 - b_2$ is ambiguous since both $d_2$ and $b_2$ are ambiguous in sign. Equations (5.19) and (5.22) constitute a system of two simultaneous differential equations in the variables $u$ and $\psi$. This system is more complicated than it may appear, however, because it is nonlinear and has multiple equilibria. Setting $\dot{\psi} = 0$ in equation (5.19), there are two nullclines (demarcation curves) given by $\psi = 0$ (which is the horizontal axis) and $\psi = -(\omega_0 + \omega_1 u)/\omega_2$. Similarly, setting $\dot{u} = 0$ in (5.22) implies two nullclines, $u = 0$ (the vertical axis) and $\psi = -(v_0 + v_1 u)/v_2$. Hence, there are equilibria at all points where these four nullclines intersect.[36]

One solution to this problem would be to set aside the zero-valued solutions for $u$ and $\psi$ as trivial, and conduct local stability analysis in the neighbourhood of an all-positive equilibrium point. As an alternative, we will express the variables in natural logarithms (as is done in many empirical studies anyway for econometric reasons), in which case we can write the system as a log-linear one that may have a unique solution (and which can still generate cyclical behaviour):

$$\hat{u} = d\ln u/dt = \upsilon_0 + \upsilon_1 \ln u + \upsilon_2 \ln \psi \tag{5.23}$$

$$\hat{\psi} = d\ln \psi/dt = \omega_0 + \omega_1 \ln u + \omega_2 \ln \psi \tag{5.24}$$

Now we can define the AD curve (in $\ln u \times \ln \psi$ space) as the nullcline (demarcation curve) on which $\hat{u} = d\ln u/dt = 0$, with slope$(\partial \ln \psi / \partial \ln u)_{\hat{u}=0} = -(\upsilon_1/\upsilon_2)$, while DC is the one on which $\hat{\psi} = d\ln \psi/dt = 0$, with slope $(\partial \ln \psi/\partial \ln u)_{\hat{\psi}=0} = -(\omega_1/\omega_2)$. Also note that the logarithm of the utilization rate can be seen as the logarithmic output gap, since $\ln u = \ln Y - \ln Y_K$. The mathematical conditions for this system of equations to be stable (or unstable) and for it to generate cyclical oscillations are analysed in Appendix 5.2. What follows here is a more intuitive and graphical presentation.

Assuming $\upsilon_1 < 0$ for Keynesian stability, AD slopes upward when $\upsilon_2 > 0$ and downward when $\upsilon_2 < 0$. Following Barbosa-Filho and Taylor, these signs will be referred to as denoting the wage-led and profit-led cases, respectively. However, it should be noted that this way of defining those cases includes responses of capacity (potential output) as well as of demand (actual output) to distributional shifts, since $\upsilon_2 = d_2 - b_2$. Barbosa-Filho and Taylor's use of the wage-led/profit-led nomenclature can thus lead to some difficulties in interpreting empirical estimates of the slope of the AD relationship. To make the point briefly here, suppose that aggregate demand is wage-led so that $d_2 > 0$ in equation (5.20). If persistent increases or decreases in output lead to endogenous adjustments in potential output in the same direction, then $b_2 > 0$ in equation (5.21), in which case $\upsilon_2 = d_2 - b_2$ in (5.22) is ambiguous in sign. Empirical estimates of $\upsilon_2$ may therefore fail to identify the wage-led nature of demand. In fact, they could reach the opposite conclusion if $b_2 > d_2 > 0$, which is possible if the positive impact of a higher wage share on potential output is greater than the impact on current output. Despite these difficulties, we will continue to refer to the cases of upward and downward sloping AD curves based on equation (5.23) as the wage-led and profit-led cases, for consistency with the literature in this genre.

Many different configurations of the AD and DC curves, with correspondingly distinct dynamics, are possible (especially if we admit cases other than those assumed here for the signs of the underlying parameters). In Figure 5.6, we show four possible types of dynamics that can arise in the profit-led case, in which AD slopes downward; the dynamics for cases of wage-led demand (an upward-sloping AD curve) are left to the reader as an exercise. Damped neo-Goodwin cycles are found in the case of a stable focus depicted in panel (a) of Figure 5.6, which requires profit-led demand (downward-sloping AD,

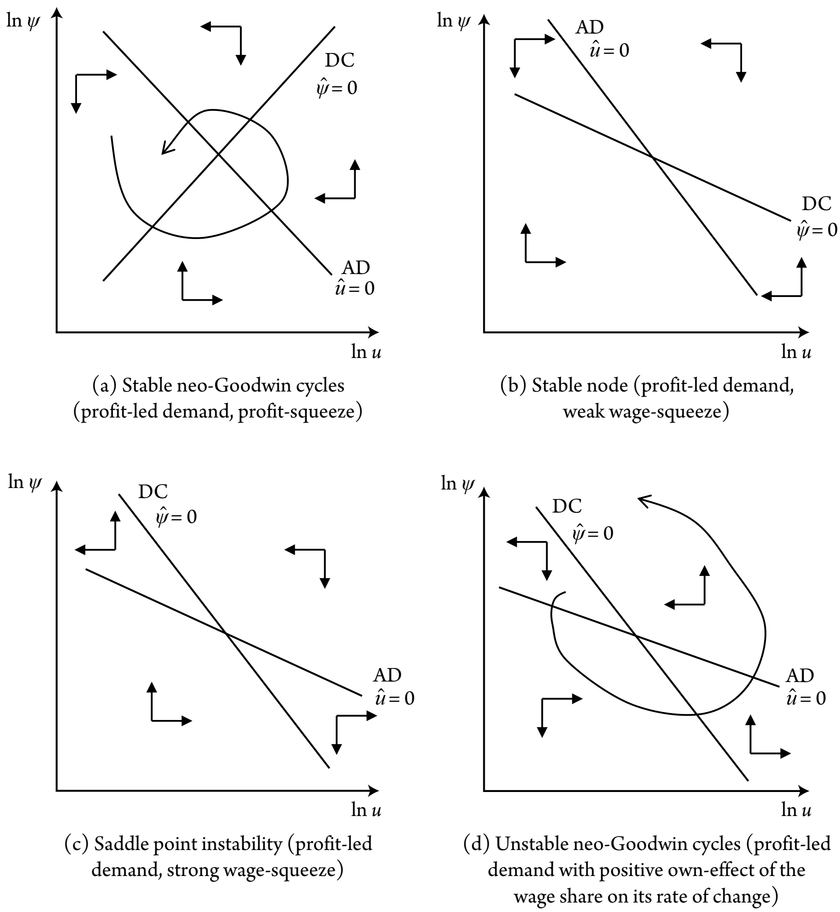



**Figure 5.6** Four alternative dynamics of distribution and utilization (neo-Goodwin cycles and other cases), all assuming profit-led demand

or $\upsilon_2 < 0$) and a profit-squeeze in distribution (upward-sloping DC, or $\omega_1 > 0$). This generates a counterclockwise cyclical rotation, which accords with empirical evidence about US business cycles found in Barbosa-Filho and Taylor (2006) and Barrales and von Arnim (2017). However, as discussed in Appendix 5.2, the existence of cyclical behaviour in this case depends on the particular values of the parameters in equations (5.23) and (5.24), especially relatively strong 'cross effects' of each variable on the changes in the other variable ($\upsilon_2$ and $\omega_1$).

Panels (b) and (c) in Figure 5.6 illustrate two other possibilities, in which DC is downward sloping (indicating a wage-squeeze in distribution) and there are no cycles. In panel (b) AD is steeper than DC, which makes the equilibrium a stable node, while in panel (c) DC is steeper than AD, resulting in a saddle point equilibrium.[37] However, these two cases do not exhibit cyclical behaviour, and are therefore inconsistent with the stylized facts associated with actual business cycles, as described above.

Barbosa-Filho and Taylor also considered a case that can generate unstable or divergent neo-Goodwin cycles, illustrated in panel (d) of Figure 5.6. To generate this result, they have to assume that wage share adjustment is self-destabilizing ($\omega_2 > 0$), which would be impossible under the assumptions made earlier in this chapter. This would require that workers have such strong bargaining power that, when the wage share rises, they are able to obtain even greater wage increases than previously (rather than moderating their wage demands when the wage share rises, as we have assumed previously).[38] If this assumption is combined with $\omega_1 > 0$ (profit-squeeze), then DC will slope downward, but in this case it must be steeper than AD to avoid saddle point instability. Then, under certain parameter values (see Appendix 5.2 for details), there can be a counterclockwise rotation of the neo-Goodwin cycle variety, but the cycles will be explosive: they become more amplified over time and the economy moves further and further away from the equilibrium (with ever-wider cyclical fluctuations), as shown in panel (d).

The emphasis on accounting for business cycles in the neo-Goodwinian approach represents a significant shift in the focus of the models of growth and distribution that we have studied up to now. Indeed, the main theoretical contribution of the neo-Goodwinian approach is to provide a rationale for the observed counterclockwise rotations of the data on wage shares and utilization rates for the US and other economies in $u \times \psi$ space at business cycle frequencies (although other explanations are also possible, as discussed later in this chapter and in Chapters 6–7). But at best this model *only* tells us about the direction of cyclical movements in utilization and distribution; it does *not* tell us about the causes of longer-term shifts in these variables or in growth rates, that is, what causes the equilibrium values of $\psi$ and $u$ (or other measures of output, like the growth rate) to change across decades or longer periods. For example, a model of short-run cycles cannot tell us why average growth has stagnated while labour shares have declined in many countries since the 1990s.

Even if demand is profit-led in the short run, as must be assumed to generate neo-Goodwin cycles,[39] this merely implies that falling profits can instigate a recession while increased profits help to spark a recovery. It does not guarantee that the economic system as a whole is profit-led in any meaningful sense in comparisons across medium-run equilibria. For example, financial liberalization could drive shifts in both AD (depressing demand) and DC (weakening labour), causing changes in both demand and distribution that increase profits and reduce growth (or utilization) but without profits 'leading' the results (Blecker, 2020). Thus, the neo-Goodwinian model should not be interpreted as suggesting that efforts to boost profitability are the best strategy for achieving high average rates of capacity utilization and employment or more robust long-term growth. Rather, the neo-Goodwinian model tells a specific story about what drives business cycles that is generally consistent with the original (simpler) Goodwin model, namely, rising wages in the expansion phase squeeze profits, and as profitability falls, capital accumulation falters and the economy eventually enters a recession, which then restores profitability by lowering wages, and rising profits then spark a recovery. In the next section (and in later chapters), we will discuss reasons to doubt that this story is an accurate portrayal of typical business cycle dynamics.

## 5.4 Empirical methodologies, findings and critiques[40]

The neo-Kaleckian models developed in the previous chapter and their extensions to incorporate conflicting income claims in this chapter reveal many possible cases: demand can be either wage-led or profit-led, distribution can exhibit either a profit-squeeze or a wage-squeeze, and various different kinds of dynamics (stable or unstable, cyclical or acyclical) can result. As a result, an enormous empirical literature has sprung up, attempting to estimate the relationships involved and to determine which cases are most realistic or prevalent in real-world economies. This section will review alternative estimates of the slopes of the demand and distributional relationships, with a primary focus on the former, as well as critiques of the methodologies employed.

### 5.4.1 Analytical framework and alternative methodologies

There are three empirical approaches for estimating the effects of income distribution on aggregate demand, which we will call the 'structural', 'aggregative' and 'reduced form' approaches. If the aggregate demand relationship is estimated (using either a structural or aggregative approach) simultaneously with the distributional relationship, the estimates can also be regarded as a 'systems' approach. In a reduced form approach, since the distributional relationship is substituted into the aggregate demand relationship, no systems or simultaneous equations methods are needed. The first two of these methods yield estimated slopes of the AD relationship, while the third method effectively estimates how shocks to exogenous variables affect the AD–DC equilibrium.

To see the distinctions between these methods, consider a standard version of aggregate demand in a neo-Kaleckian macro model, taken (with some modifications) from Stockhammer et al. (2011)[41]

$$Y = AD = C(Y, \psi, Z_C) + I(Y, \psi, Z_I) + G + NX(Y, P, Z_X, Z_M) \quad (5.25)$$

$$P = P(\psi, Z_P), P_\psi > 0 \quad (5.26)$$

where $Y$ is output, $AD$ is aggregate demand, $C$ is consumption, $I$ is investment, $G$ is government purchases, $NX = X - M$ represents net exports ($X$ and $M$ are exports and imports, respectively), $P$ is the domestic price level and $Z_J$ is a vector of exogenous (control) variables affecting each endogenous variable $J$ ($J = C, I, X, M, P$). Net exports can be decomposed into separate functions for export and imports:

$$X = X(P, Z_X), X_P < 0 \quad (5.27)$$

$$M = M(P, Y, Z_M), \quad M_P > 0, M_Y < 0 \quad (5.28)$$

where foreign income ($Y_f$) would usually be included in $Z_X$ with a positive effect on $X$. Since the nominal exchange rate and foreign price level are typically included in $Z_X$, $Z_M$ and/or $Z_P$, the domestic price level $P$ is used to capture the relative price of home goods compared to foreign goods.[42] It is generally hypothesized that $C_Y > 0$, $C_\psi > 0$, $I_Y > 0$, $I_\psi < 0$, $NX_Y < 0$ and $NX_P < 0$.[43] The hypothesis that $I_\psi < 0$ assumes that investment is predominantly corporate or business investment and that a lower profit share depresses investment demand; as discussed earlier, this sign could turn positive if residential investment is included and the wage share has a strongly positive effect on workers' housing demand, or if firms respond to a higher wage share (real unit labour costs) by investing heavily in equipment embodying labour-saving technology.

The slope of the AD relationship – which is the effect of a change in the wage share on output, holding all exogenous terms $Z_J$ constant – is given by totally differentiating equation (5.25) and solving for

$$\frac{\partial Y}{\partial \psi} = \frac{\dfrac{\partial AD}{\partial \psi}}{1 - \dfrac{\partial AD}{\partial Y}} \quad (5.29)$$

where $\frac{\partial AD}{\partial Y} = \frac{\partial C}{\partial Y} + \frac{\partial I}{\partial Y} - \frac{\partial M}{\partial Y}$. Assuming $\partial AD/\partial Y < 1$ for Keynesian (goods market) stability, the sign of $\partial Y/\partial \psi$ (and hence the slope of AD, taking potential output $Y_K$ as given at any point in time) depends only on the sign of the numerator, $\partial AD/\partial \psi$. Then the total effect on output shown by derivative (5.29) can be calculated as the product of the direct impact of distribution on aggregate demand $\partial AD/\partial \psi$ and the multiplier $1/[1-(\partial AD/\partial Y)] > 1$. Where the first two methods differ is on how to estimate the slope of AD econometrically.

The **structural approach** estimates the individual components of aggregate demand using *separate econometric equations* for *C*, *I*, *X*, *M* and *P*.[44] In this approach, $\partial AD/\partial \psi$ is then calculated by summing the estimated partial derivatives for consumption, investment and net exports with respect to the wage share. Assuming that net exports are decomposed into exports minus imports as explained above, this approach relies on the following calculation[45]

$$\frac{\partial AD}{\partial \psi} = \frac{\partial C}{\partial \psi} + \frac{\partial I}{\partial \psi} + \left(\frac{\partial X}{\partial P} - \frac{\partial M}{\partial P}\right)\frac{\partial P}{\partial \psi} \tag{5.30}$$

Whether the AD curve is upward or downward sloping (wage-led or profit-led) is inferred from the sign of this derivative. Then, equation (5.29) is applied to determine the total impact of a distributional shift in favour of labour on output, including multiplier effects.

In contrast, the **aggregative approach** relies on direct estimation of the effects of the wage share and any control variables on aggregate demand, that is, the *sum* of all the terms on the right-hand side of equation (5.25), which can be written in implicit form as

$$Y = AD(\psi, G, Z_C, Z_I, Z_X, Z_M, Z_P) \tag{5.31}$$

The derivative $\partial Y/\partial \psi = \partial AD/\partial \psi$ is then found from the coefficient on the wage share (including any lagged effects) in an econometric estimate of equation (5.31). In practice, output in this equation is usually normalized by potential output $Y_K$, so that in effect what is estimated is a capacity utilization function

$$u = u(\psi, G, Z_C, Z_I, Z_X, Z_M, Z_P) \tag{5.31$'$}$$

Also in practice, since the models are estimated in discrete time (usually using quarterly data, sometimes annual), lags of the dependent variable ($u$) are often included on the right-hand side of (5.31′) in econometric implementations,[46] and the number of exogenous control variables included varies as we shall discuss below.

Lastly, at least one study (López et al., 2011) has adopted what could be called a **reduced form approach**. Suppose we specify a simple, implicit function for the distributive relationship or DC curve,

$$\psi = \psi(u, Z_\psi) \qquad (5.32)$$

where $Z_\psi$ is a vector of exogenous factors affecting income distribution. Substituting equation (5.32) into (5.31′), we obtain the reduced form solution for the utilization rate as a function only of exogenous variables

$$u = u(G, Z_C, Z_I, Z_X, Z_M, Z_P, Z_\psi) \qquad (5.33)$$

Economists using this approach estimate the effects of the underlying determinants of the wage share in the vector $Z_\psi$ on equilibrium utilization, but do not attempt to estimate the effects of the wage share (which is treated as endogenous) per se. Thus, this method does not attempt to determine the slope of the AD curve; instead, it focuses on identifying the changes in the equilibrium utilization rate caused by shocks to the underlying determinants of both demand and distribution.[47] This approach recognizes that distribution can change for different reasons (such as changes in labour's bargaining power, monopoly power of firms or the real exchange rate), which could have different effects on equilibrium utilization. For exogenous variables in $Z_\psi$ that may also be included in other $Z$ vectors, this method will find the total effects of those variables not just the effects that operate through the wage share.

Turning to the distributive relationship or DC curve, fewer studies have sought to estimate this and the methodologies are less well defined. However, the most common approach is to estimate some version of equation (5.32) for the wage share along with (5.31′) for the utilization rate. Such efforts to control for endogeneity by estimating the AD and DC relationships simultaneously can be called **systems estimates**. Although in principle both structural and aggregative models could be estimated as systems with an endogenous wage share, in practice this has been done mainly using the aggregative approach. For example, Fernandez (2005) used two-stage least squares (2SLS) with instrumental variables (IV) to control for the endogeneity of the profit share in his AD equation, using annual US data.[48]

Barbosa-Filho and Taylor (2006) and Carvalho and Rezai (2016) instead used vector autoregression (VAR) methods, which omit the contemporaneous values of the variables and hence have only lags (which are assumed to be predetermined) on the right-hand side of the regression equations, applied to quarterly US data.[49] Barbosa-Filho and Taylor estimated a discrete-time version of the wage share equation in log differences (5.24), using quarterly US data with two lags of ln $u$ and ln $\psi$. Neither they nor Carvalho and Rezai included any control variables or covariates, other than lags of ln $u$ and ln $\psi$. Nikiforos and Foley (2012) estimated a similar model but included a quadratic term for utilization ($u^2$) in their DC equation to test for nonlinearity (recall that they found a U-shaped DC curve). Nikiforos and Foley also included a government-spending variable to control for fiscal policy in their AD equation and a dummy variable to capture 'changes in the political-economic environment' in their DC equation, and used 2SLS to control for endogeneity.

Kiefer and Rada (2015) estimated a more complex systems model that allows for shifts in the equilibrium values $u^*$ and $\psi^*$ over time

$$u_{it} - u_{it-1} = \beta_0(\psi_{it-1} - (\psi_t^* - \beta_1 u_t^*) - \beta_1 u_{it-1}) + \delta_{it} \qquad (5.34)$$

$$\psi_{it} - \psi_{it-1} = \alpha_0(\psi_{it-1} - (\psi_t^* - \alpha_1 u_t^*) - \alpha_1 u_{it-1}) + \varepsilon_{it} \qquad (5.35)$$

where $i$ indexes the country, $t$ is time, $\alpha_j$ and $\beta_j$ ($j = 0,1$) are coefficients, and $\delta$ and $\varepsilon$ are random errors. Kiefer and Rada found that demand is profit-led and distribution exhibits a profit-squeeze in the short run, thus supporting a neo-Goodwin cycle interpretation. In addition, they found robust evidence (using a variety of econometric techniques as sensitivity tests) that equilibrium $u_t^*$ and $\psi_t^*$ declined in the long run over the period 1971–2012. However, aside from various lags, the only control variable they included was the average wage share for all other countries, intended to represent a 'race to the bottom' in distribution (individual countries lowering their wage shares to compete with other nations that are also lowering their wage shares).

Since the vast majority of studies have used either a structural or aggregative/systems approach, we will focus our discussion mainly on these two. A priori, each of these two methodologies has its strengths and weaknesses. One major advantage of the structural method is that it can identify the sign and magnitude of the effect of distribution on each component of demand, and thus allows for a distinction between domestic demand effects (measured by $\partial C/\partial\psi + \partial I/\partial\psi$) and the total effect including net exports per equation (5.30). The structural method also allows for comparing the impact of distributional shifts within an individual country with the impact of simultaneous global shifts in income distribution (Onaran and Galanis, 2012; Onaran and

Obst, 2016). Simultaneous shifts are more likely to result in wage-led demand outcomes, since the negative effects of higher labour costs on net exports largely cancel out when all countries raise their wage shares at the same time. However, most estimates using this approach (whether for individual countries or global impacts) have ignored the systems aspect of their models – the potential endogeneity of income distribution and other included variables as well as common shocks affecting the different components of aggregate demand – by estimating the individual equations using OLS.[50]

In contrast, perhaps the greatest advantage of the aggregative/systems approach is that it easily addresses the simultaneity of demand and distribution, for example by applying IV or VAR methods (or vector error correction, VEC, in the presence of cointegration) to systems of equations like (5.23) and (5.24) or (5.31′) and (5.32). However, the aggregative approach does not readily distinguish which components of aggregate demand (consumption, investment or foreign trade) are driving the results.[51] Also, the prevalence of VAR or VEC methods in this literature has meant that few control variables (sometimes none) are usually included in the empirical models, which means that the estimates could be subject to omitted variable bias. Other problems in the typical empirical implementations of the aggregative approach will be discussed below.

### 5.4.2 Major recent studies: results and critiques

Table 5.1 lists some of the major recent studies that have estimated the impact of distribution on demand, grouped by methodology and findings.[52] Clearly, results for individual countries vary widely; some studies have found opposite results for the same countries (notably the US, Netherlands and Japan). But what is most striking is how the results are strongly correlated with the methodology employed. With only a few exceptions (to be discussed below), the aggregative/systems estimates almost always find that demand is profit-led demand, especially in the US case (which is the most studied). Although the structural estimates have more varied results, they are much more likely to generate findings of wage-led demand, especially for the larger countries or country blocs (US, Germany, euro zone and so on), while sometimes finding profit-led demand especially in more open (smaller or less developed) economies. The use of panel data does not eliminate this difference: Kiefer and Rada (2015) find profit-led demand using an aggregative/systems approach applied to panel data for 13 OECD countries, while Stockhammer and Wildauer (2016) find mostly wage-led demand using a structural model with a slightly larger (but overlapping) panel of 18 OECD countries. One therefore naturally wonders what could

**Table 5.1** Selected empirical studies of wage-led versus profit-led demand using alternative methodologies for countries shown

| Methodology | Findings | |
|---|---|---|
| | Wage-led demand | Profit-led demand |
| Structural estimates | Naastepad & Storm (2006), France, Germany, Italy, Netherlands, UK<br>Hein & Vogel (2008), France, Germany, UK, US<br>Stockhammer et al. (2009), euro area<br>Onaran et al. (2011), US<br>Stockhammer et al. (2011), Germany<br>Onaran & Galanis (2012), euro area, Germany, France, Italy, UK, US, Japan, Turkey, South Korea<br>Stockhammer & Wildauer (2016), panel of 18 OECD countries[b]<br>Onaran & Obst (2016), 11 European countries[c]<br>Cauvel (2018, Chapter 2), US (systems GMM)[d] | Naastepad & Storm (2006), Japan, US<br>Hein & Vogel (2008), Austria, Netherlands<br>Onaran & Galanis (2012), Australia, Canada, Argentina, China, India, Mexico, South Africa[a]<br>Onaran & Obst (2016), Austria, Belgium, Denmark, Ireland |
| Aggregative estimates | Vargas Sánchez & Luna (2014), Mexico, long run[e]<br>Charpe et al. (2018), US, UK, France, long run (growth)[f]<br>Cauvel (2018, Chapter 1), US (short run, controlling for productivity)[g]<br>Araujo & Costa Santos (2018), US (long run)[f] | Fernandez (2005), US<br>Barbosa-Filho & Taylor (2006), US<br>Tavani et al. (2011), US<br>Nikiforos & Foley (2012), US<br>Vargas Sánchez & Luna (2014), Mexico, short run<br>Kiefer & Rada (2015), panel of 13 OECD countries[h]<br>Carvalho & Rezai (2016), US<br>Barrales & von Arnim (2017), US[f]<br>Charpe et al. (2018), US, UK, France, short run (growth)[f]<br>Silva de Jesus et al. (2018), Brazil<br>Araujo & Costa Santos (2018), US (short run)[f] |
| Reduced form estimates | López et al. (2011), Mexico, short run[i] | |

Notes: For structural studies that give separate results for domestic demand and total demand (including net exports), only the results for total demand are shown here. Studies that only cover domestic demand and do not include net exports are not included. Some studies are shown more than once if they found different results for different countries or time horizons. Long run and short run are indicated only for studies that distinguish these.

**Table 5.1** (continued)

a Individual country results (some of these countries flip from profit-led to wage-led in response to a simultaneous redistribution of income in all countries).
b Marginal effects vary for individual countries; examples given are Netherlands (profit-led) and US, France, Germany, Austria (wage-led, in declining order of the magnitude of the effect). Other countries included in the panel estimates are Australia, Belgium, Canada, Switzerland, Denmark, Spain, Finland, Ireland, Italy, Japan, Norway, Sweden, UK.
c Finland, France, Germany, Greece, Italy, Luxembourg, Netherlands, Portugal, Spain, Sweden, UK.
d Using systems GMM to control for endogeneity of the wage share and GDP, also including a wage share equation.
e There is a typographical error in the cointegrating equation printed in this article, but I have verified by email from Gustavo Vargas Sánchez (14 October 2014) that the sign on the 'exploitation rate' (profit share) is negative indicating that output is wage-led in the long run. This is also consistent with what the authors state in the text.
f Uses a different methodology (wavelets) from most other studies shown.
g Using a cyclically adjusted measure of the wage share or treating the real wage and labour productivity as separate variables; results are sensitive to ordering in the impulse responses.
h Australia, Canada, Finland, France, Germany, Ireland, Italy, Japan, South Korea, Netherlands, Sweden, UK, US.
i Although this methodology does not directly identify whether demand is wage-led or profit-led, this study's finding that a depreciation of the peso is contractionary is consistent with the wage-led case (since a depreciation would be expected to lower the real wage). The results are referred to as 'short run' because the data are filtered and detrended.

account for this striking divergence in the results obtained using these two methodologies.

Of course, all econometric estimates can be highly sensitive to various aspects of their specifications, including data frequency and lag lengths,[53] measurement or transformation of the variables (for example, logs, differences, normalizations, filtering), functional forms (for example, linear or nonlinear), control variables included or omitted, and whether the methodology controls for potential endogeneity or simultaneity bias. Although many of these differences are idiosyncratic to particular studies in this literature, some general tendencies can be observed. Especially, most estimates of structural models to date have treated the wage (or profit) share and other right-hand-side variables as exogenous, thus creating possible simultaneity bias, while many of the aggregative/systems estimates (especially those based on the neo-Goodwinian approach) often lack control variables, which suggests the possibility of omitted variable bias.[54] Also, with a few notable exceptions, most studies using both approaches have paid scant attention to identifying common factors that could drive changes in both distribution and utilization or testing whether different causes of variations in distribution have different impacts on aggregate demand, as implied in the open economy model covered in section 4.4.3 in Chapter 4.[55]

One possible explanation for the different results is that the aggregative approach could be capturing some dynamic interactions that the structural approach potentially misses by estimating the impact of the wage share on the individual components of aggregate demand separately. For example, if a rise in profitability stimulates investment and this in turn boosts consumption via the multiplier effect on income, this will be captured by an aggregative model as a positive effect of profits on demand, whereas in a separate estimation of a consumption function the positive effect on aggregate consumption would be picked up by the income variable rather than the wage share. Also, in an open economy a rise in the wage share that increases consumption of imported goods will reduce net exports, but depending on the specification this may not be captured as a distributional effect in the net export part of the model (it could appear instead as an income effect in the import equation).

On the other hand, some interactions could make economies even more wage-led than they appear in a standard structural model. For example, if a rise in the wage share boosts consumption demand and this in turn stimulates investment via the accelerator effect, this would be incorporated as a distributional effect in an aggregative model but would be picked up by the utilization or accelerator term in a structural model, not by the distributional variable. Onaran and Obst (2016) address this issue by estimating the total effect of a distributional shift on investment, including the indirect effect via the multiplier–accelerator interaction as well as the direct effect on investment itself. They find that the total effect of the wage share on investment including this indirect effect is positive in most of the 15 European Union countries in their sample. Onaran and Obst also estimate the total effects of the wage share on net exports including multiplier–import interactions, but do not estimate possible investment-driven multiplier effects on consumption.

Most of the structural studies listed in the top portion of Table 5.1 have found robust evidence that consumption is wage-led (the coefficient measuring $\partial C/\partial \psi$ almost always turns out to be significant and positive) for the vast majority of countries studied.[56] Therefore, the results for different countries depend mainly on the strength of the distributional effects on investment (which are often small or insignificant, and may have either a negative or positive sign) and net exports: overall demand will be wage-led unless these last two effects combined are sufficiently negative. Except for the now somewhat old study by Naastepad and Storm (2006),[57] net negative effects on total demand are usually found (especially in the more recent studies) only for relatively small or highly open economies, such as Austria, Belgium, China, Ireland and Mexico.

A few studies (all using the structural methodology) have considered the impact of a simultaneous shift in income distribution across a sample of countries, a consideration that is especially important given the decline in the wage share in a large number of countries over the past two decades. Onaran and Galanis (2012) conduct such estimates for a wide range of advanced and developing or emerging economies, while Onaran and Obst (2016) do the same for their sample of 15 European nations. In both studies, since the competitive gains from a lower wage share (positive effects on net exports) tend to cancel out when all countries lower their labour costs simultaneously, many (although not all) countries that have profit-led demand by themselves 'flip' to having wage-led demand when simultaneous changes in all countries' wage shares are considered.[58]

### 5.4.3 Problems with aggregative estimates of profit-led demand

As discussed earlier, the aggregative/systems estimates almost uniformly find that demand is profit-led – and, in combination with finding a profit-squeeze in distribution, these same studies usually support the existence of neo-Goodwin cycles.[59] However, these findings are subject to three major problems: at best, the findings only pertain to short-run relationships, not to long-run behaviour; the most commonly used measures of the utilization rate or output gap are subject to statistical flaws that could make the estimates biased or spurious; and the failure to control for the cyclical behaviour of labour productivity biases the estimates towards finding profit-led demand. We will discuss each of these criticisms in turn.

*Short-run focus*

The short-run focus of most aggregative/systems estimates is seen not only in the theoretical framework of neo-Goodwin cycles, which most of these studies adopt, but also in how they specify the variables in these models empirically. For example, Barbosa-Filho and Taylor (2006) and Carvalho and Rezai (2016) measure $u$ by deviations of real GDP (in natural logs) from its Hodrick–Prescott (HP) filtered trend. In this method, the mean of the utilization index is forced to equal zero, so longer-term variations in utilization are ruled out by assumption. Hence, their findings of profit-led demand pertain *only* to short-run fluctuations in utilization around a mean that by construction must be constant in the long run, and the estimates are not informative about medium-run or long-run effects of distribution on output or growth.

As discussed earlier, Kiefer and Rada (2015) improve on this methodology by allowing the long-run equilibrium levels of $u$ and $\psi$ to vary, and discover that the equilibrium levels of both variables have shifted downward in the long run. This, of course, requires using a different measure of utilization, which is not restricted to having a constant mean; for this purpose, Kiefer and Rada used OECD estimates of output gaps. These authors identify one factor that could account for the long-term decline in the wage share (the race to the bottom, represented by the average wage share for all other countries having a negative effect on the wage share for each country in their panel), but do not test for other possible causes. Thus, the finding of profit-led demand in Kiefer and Rada pertains only to short-run fluctuations around those declining long-run trends; in fact, they identify a positive long-run correlation between the wage share and the output gap, but their methodology does not allow them to determine the direction of long-run causality.

Some economists have begun to suggest that distributional effects on demand or growth may vary over different time horizons. In a neoclassical framework, Halter et al. (2014, p. 81) provide both theory (of the NEGT variety, discussed in Chapter 1) and empirical evidence in support of the proposition that 'Higher inequality helps economic performance in the short term but reduces the growth rate of GDP per capita farther in the future.' In their theoretical model, increased inequality boosts private production in the short term, but reduces public investment with negative consequences for output in the long term. Applying GMM estimation to panel data for 106 countries for eight five-year periods between 1965 and 2005, Halter et al. find that inequality (measured by Gini coefficients) has a positive effect on average growth rates in the subsequent five-year period, but negative effects beyond that, with an overall negative long-run impact.

Blecker (2016c) hypothesizes that demand is more likely to be profit-led in the short run (if at all) and more likely to be wage-led in the long run, at least in the US case, based on the following propositions:

1) Positive effects of a higher wage share on consumption are more likely to prevail in the long run, because worker households are less constrained by income in the short run when they can finance consumption by borrowing.
2) Negative effects of a higher wage share on investment are more likely to prevail in the short run, because reductions in firms' profits mainly constrain their ability to finance planned investment projects, while in the long run investment plans are driven largely by accelerator effects.

3) Negative effects of a higher wage share on net exports are also more likely to prevail in the short run, partly because of well-known adjustment mechanisms that can offset higher labour costs (for example, a currency depreciation), but which may do so only with significant lags, and partly because of other possible long-run responses. For example, a high-wage country can specialize in high-quality, technologically innovative products that offer non-price competitive advantages.

A few studies have tested for the effects of the functional distribution of income on utilization or growth over different time horizons. Barrales and von Arnim (2017) use the methodology of wavelet decomposition to analyse cycles of different lengths in US macroeconomic data. For cycles of 4 to 8 and 8 to 16 years in length, they find a consistent neo-Goodwinian pattern in the data (counterclockwise rotations) using several alternative measures of economic activity (the output gap, income–capital ratio and employment rate) combined with the wage share. For the longest cycles they are able to identify (16 to 32 years), the neo-Goodwin pattern breaks down after 1980 as both economic activity and the wage share exhibit long-term declines thereafter.

Charpe et al. (2018) apply similar methods of wavelet analysis to much longer-term data for the US (1898–2010) as well as the UK (1856–2010) and France (1896–2010). Charpe et al. study output growth (rather than the utilization or employment rate) and generally conclude that 'An increase in the labor share reduces growth in the short-term but enhances growth in the long-term' (2018, p. 15). More specifically, they find that the labour share has a positive impact on output growth (even after controlling for possible endogeneity) at frequencies of 32 years and up in all three countries, and also at slightly higher frequencies (such as 16–32 years) in some countries in some periods, while at still higher frequencies (2–4 or 4–8 years) the impact of the labour share on growth is generally negative. Similarly, Araujo and Costa Santos (2018) conclude that the US exhibits profit-led demand and growth in the short run and wage-led demand and growth in the long run, also using wavelet methods.

### *Measurement of utilization rates or output gaps*

There is also a deeper problem with most commonly used measures of utilization (actual output relative to potential) in the empirical literature. That is, *ex post* estimates of potential output tend to adjust the trend of potential output downward in response to a persistently lower level of actual output following a recession, which has the effect of exaggerating utilization (or the output gap) prior to the downturn and underestimating the magnitude of the subsequent downturn (Cerra and Saxena, 2017). Contrary to what is assumed in most mainstream models of growth and cycles, economies do not generally return to predetermined trends after a major crisis or recession, but instead usually exhibit persistent declines in their growth trajectories (Cerra and Saxena, 2017). This tendency has been especially notorious since the Great Recession of 2007–09, after which the estimated trend of potential output is much lower than previous estimates (Ball, 2014), but it has also been observed in data on output growth across large samples of countries (both advanced and developing economies) and periods of time (see Cerra and Saxena, 2008; Cerra et al., 2013).

The measurement error embedded in typical measures of utilization (or the output gap) is most clear in the frequently used method of taking the trend of output estimated by an HP filter as a proxy for potential output. The use of an HP filter makes it impossible to accurately assess the timing of causal effects or to control for endogeneity by using lags, because the estimate of potential output (the trend) at any time $t$ incorporates information about both lags and leads of actual output at times before and after $t$. If the utilization rate (output gap) $u_t$ is then measured as the log difference between the actual output and its HP-filtered trend (both in logs), it also contains information from other time periods both before and after $t$. Hence, lagged values $u_{t-i}$ are not really predetermined at time $t$, and the estimated coefficients in any regression model or VAR system including the utilization rate do not reflect the true time phasing of the causal relationships between utilization, distribution and other variables.

Hamilton (2018) summarizes the statistical flaws of the HP approach in an article provocatively titled, 'Why you should never use the Hodrick-Prescott filter'.[60] Hamilton proves that, among other things, 'The Hodrick-Prescott (HP) filter introduces spurious dynamic relations that have no basis in the underlying data-generating process', and 'filtered values at the end of the sample are very different from those in the middle and are also characterized by spurious dynamics' (p. 831).[61] Hamilton proposes an alternative method for separating the cycle from the trend in macro time-series that does not suffer from these defects. For quarterly data, which are usually used in aggregative estimates of AD curves, Hamilton recommends measuring the cyclical component by the residual from a regression of any variable $y$ at time $t + h$ on its four most recent quarterly lags starting $h = 8$ quarters earlier

$$y_{t+h} = \beta_0 + \beta_1 y_t + \beta_2 y_{t-1} + \beta_3 y_{t-2} + \beta_4 y_{t-3} + v_{t+h} \qquad (5.36)$$

where the estimated residual $\hat{v}_{t+h}$ (here, the circumflex indicates an estimated value rather than a rate of change) is the cycle series. Note that, if $y$ is the natural log of real output, this residual series is the logarithmic output gap or utilization rate. Essentially, the cyclical component is measured as the deviation of the series from a forecast based on four quarters of realizations of the variable starting eight quarters earlier.

Hamilton's new method has many desirable statistical properties: it produces a cycle series that is stationary for variables with any degree of integration up to order four; it eliminates any issues of seasonality; it uses only information that is available to agents at time $t + h$ (no future information is used); and it does not bend the trend downward immediately before and after a recession. Nevertheless, the Hamilton method is not without some problems. Similar to an HP filter, it imposes the restriction of a zero mean on the cycle series by construction, so it cannot be used to detect long-term variations in utilization (although critics of neo-Kaleckian models who think that utilization cannot vary in the long run would not see this as a weakness – see Chapter 6 for more discussion). In addition, although the Hamilton cycle measure does not generate spurious spikes in utilization just before a recession and does not underestimate the depth of a recession, assuming the recession lasts less than two years, it does generate anomalously rapid declines in the trend (and correspondingly large hikes in estimated utilization) around two years (eight quarters) after a major crisis like the 2008–09 US recession. The only study we are aware of that has applied the Hamilton method to the estimation of distributional effects on demand is Cauvel (2018), whose work will be discussed below.

The problems with an HP filter are also found to varying degrees in other methods of estimating potential output, such as those based on a neoclassical model of what output would be at full employment of the 'factors of production'. As discussed in Chapter 1, a large amount of literature (both mainstream and heterodox) recognizes that the growth rates of physical and human capital as well as the adoption of technological innovations can be diminished by persistent shortfalls of aggregate demand such as occur in a major recession or depression (or during a prolonged period of stagnation). Since estimates of full-employment output are based on factor supplies and assumed rates of technical progress, they will naturally exhibit lower trends if factor supplies (physical and human capital) and productivity are growing more slowly in the aftermath of a crisis.

Hence, *ex post* estimates of the trend of potential output using a neoclassical method (as, for example, in the OECD output gaps) will also tend to bend downward before the onset of a major cyclical downturn. This makes capacity utilization look higher than it actually was before the recession or crisis, masks the severity of the depression of output during the downturn, and exaggerates the strength of the subsequent recovery. Like HP-filter-based measures, such estimates effectively embed future information into the estimated utilization rate (or output gap) at any time *t*, resulting in spurious estimates of contemporaneous and lagged effects.[62] Because the structural estimates generally do not normalize their variables by measures of potential output, they are not subject to the measurement error and biases introduced by the use of standard measures of utilization or output gaps in most of the aggregative estimates.

Finally, it is worth recalling what we discussed earlier in regard to the theoretical neo-Goodwinian model: estimates of the slope of AD using the utilization rate as the demand measure are estimating how the wage share affects actual output *relative* to how it affects potential output – in other words, the sign of this slope depends on the *difference* between the effects on demand and capacity, $v_2 = d_2 - b_2$ in equation (5.22) or (5.23).[63] Even if $v_2$ is found to be negative, we do not know if this is because demand is strongly profit-led ($d_2 << 0$) or rather because capacity is even more wage-led than output ($b_2 > d_2 > 0$). As noted earlier, some studies are now finding that output is more wage-led in the long run than in the short run, and if potential output follows persistent trends in actual output (as argued by Cerra and Saxena, 2017; see also Schoder, 2014) and such trends are projected backward in time by standard methods of estimating potential output, the latter case becomes plausible.

### *Endogeneity of labour productivity*

A third potential source of bias in the aggregative estimates that have found profit-led demand stems from their failure to correct for the impact of cyclical fluctuations in output on labour productivity and the wage share. As shown in Appendix 4.1 in Chapter 4, the profit share varies procyclically in response to demand shocks in the presence of overhead labour. The reason this occurs is that productivity also varies procyclically, because firms hoard overhead labour (they do not lay off executives, professionals, managers and office staff in proportion to the decline in output) in a recession. As a result, the wage share (which equals one minus the gross profit share used in Appendix 4.1) tends to vary countercyclically.[64] As argued recently by Lavoie (2017),[65] the finding of profit-led demand may therefore be based on a spurious correlation, because a demand-driven decline (recovery) in output will naturally be associated with a rise (fall) in the wage share – but with the causality running from output to the wage share (via productivity) instead of the reverse.

Cauvel (2018, Chapter 1) reports econometric evidence that confirms Lavoie's critique. When Cauvel separates out the real wage (hourly compensation) and labour productivity (output per hour) components of the wage share in a VAR model using quarterly US data for 1947Q1–2016Q4, the impulse responses show that productivity is positively affected by contemporaneous shocks to utilization (demand shocks) while demand becomes wage-led (in the sense that utilization rises in response to shocks to real hourly compensation).[66] Moreover, the profit-squeeze result disappears in these estimates: demand-side shocks (to utilization) have no significant effects on real compensation. In another set of estimates, Cauvel removes the cyclical component of productivity from the wage share using the same kind of filtering technique that he uses to measure utilization (either an HP or Hamilton filter), and again demand becomes wage-led (in the sense that shocks to the cyclically adjusted wage share have positive effects on the corresponding utilization rate). Taken together, these results suggest that the profit-led/profit-squeeze findings of previous studies have indeed been biased by a failure to control for the cyclicality of labour productivity.

Cauvel's estimates thus support Lavoie's argument that the short-run, cyclical behaviour of the wage share is driven by endogenous adjustments of labour productivity, not by movements in the real wage (which exhibits little cyclical variation). In this alternative view, the profit share declines in a recession because of decreased labour productivity, not because of a prior spike in the real wage, and rises in a recovery because labour productivity rebounds, not because wages fall. Hence, there is neither profit-led demand nor a profit-squeeze in distribution. Rather, demand-driven cycles (possibly sparked by financial instability or monetary policies) make the wage share rise and fall in ways that create a false impression of profit-led demand and a profit-squeeze if the endogenous adjustments of productivity are not controlled for. Demand is actually wage-led, once the wage share is adjusted to remove cyclical variations in productivity or else the real wage is treated separately from productivity, even in an aggregative analysis – which is more consistent with what has been found for most countries in the structural estimates.

## 5.5 Conclusions

This chapter has shown how a conflicting claims analysis can help to explain the functional distribution of income between labour and capital, and how such analysis can be combined with models of aggregate demand to understand the simultaneous determination of output (capacity utilization) and distribution (the wage or profit share). This chapter has also reviewed the empirical studies of the demand–distribution nexus. Overall, the evidence seems to support wage-led demand in most of the larger economies, but with a number of important exceptions especially for smaller and more open economies (including many emerging and developing nations). There is some mixed evidence for demand being more wage-led at long-run time horizons compared with short-run periods, and also some emerging evidence that empirical support for neo-Goodwin cycles (profit-led demand combined with a profit-squeeze in distribution) may be statistically biased for various reasons – especially the failure to control for cyclical variations in productivity.

These findings open up several directions for the alternative analyses of cyclical volatility that will be pursued in the next two chapters. Chapter 6 discusses neo-Harrodian models, which exploit Harrod's instability argument (introduced in Chapter 3) to construct theories of cyclical growth. Chapter 7 covers models of cyclical instability resulting from either the financial fragility of firms (Stockhammer and Michell, 2017) or the dynamics of household debt and expenditures (Setterfield and Kim, 2017; see also Fiebiger, 2018). These alternative approaches provide important alternatives to the neo-Goodwinian approach to business cycles, highlighting additional channels through which the appearance of profit-led demand and a profit-squeeze can arise from very different underlying dynamics.

## ? STUDY QUESTIONS

1) Show how distributive conflict can generate inflation. What relationship, if any, exists between the conflicting claims inflation process and the Phillips curve?
2) Derive the distributive curve. What does this curve add to the analytics of the basic neo-Kaleckian model outlined in Chapter 4? How do the dynamics vary depending on whether adjustments in output occur faster than, or at approximately the same rate as, adjustments in wages, prices and distributive shares?
3) Under what theoretical or empirical conditions do neo-Goodwin cycles emerge? What are the main criticisms of neo-Goodwinian models, and how do you think proponents of these models would respond?
4) Suppose one believes that growth is profit-led. How could one then explain the fact that, over the past few decades, many economies have experienced simultaneous reductions in their wage shares and slowdowns in their growth? Discuss.
5) Summarize the main methodologies and results associated with the empirical literature that investigates the relationship between distribution and growth. How do the results differ by methodology, time frame and type of country considered?
6) Given that the wage share is an endogenous variable in any model with both a distributive curve (DC) and an aggregate demand (AD) relationship, what meaning can be given to the idea that demand is 'led' by either wages or profits? Discuss.

7) In the model of Barbosa-Filho and Taylor (2006), what difficulties are associated with using the slope of the AD curve to determine whether an economy has wage-led or profit-led aggregate demand?
8) In recent decades, inflation rates have come down while wage shares have also decreased in many industrialized and emerging economies. Is this just a coincidence, or could there be a relationship, and what do you think could be the causality involved? Discuss with reference to the concepts and models covered in this chapter.

## NOTES

1 Modern conflicting claims models of inflation and distribution originated with Rowthorn (1977), Taylor (1985) and Dutt (1987). Burdekin and Burkett (1988) and Isaac (1990, 1991) analysed the interplay between conflicting claims explanations of inflation and monetary policy, while Isaac (2009) further explored the implications in relation to monetary and fiscal policies and long-run growth. Conflicting claims theories of inflation were also stated by earlier Latin American structuralists (for example, Noyola Vázquez, 1956; Sunkel, 1960; Furtado, 1963). See Lavoie (2014, Chapter 8) for a comprehensive account of post-Keynesian models of inflation, Cordero (2002) for an application to small open economies, and Vernengo (2006) and Pérez Caldentey (2018) on the structuralist tradition.

2 For alternative views on endogenous money, see Moore (1988), Wray (1990, 1998), Palley (1996c, 2013b), Rochon (1999) and Lavoie (2014), among many others. See Arestis and Sawyer (2007) for an overview and Taylor (1983, 1991) for structuralist models of inflation incorporating monetary policy for developing countries.

3 The latter quote is, of course, the famous statement of Friedman (1970); the former is a popular aphorism about inflation (which we would argue is generally mistaken as a causal claim, at least for modern advanced economies).

4 This is essentially the model of inflation in Dutt (1987), which he presented as an alternative to the hybrid neo-Marxian/neo-Keynesian model of Marglin (1984a) covered in section 3.5 of Chapter 3.

5 Below, we will sometimes use a circumflex to indicate an estimated econometric residual. Of course, labour contracts usually specify wage increases that are phased in over discrete units of time, but we use continuous time here for mathematical convenience.

6 See the discussion of factors that influence markups and alternative ways of modelling them in section 4.2 of Chapter 4 and Lavoie (2014) for a more in-depth discussion of firms' pricing behaviour.

7 Recall that the labour share equals $a_0w$ while the capital share equals $1 - a_0w$, so if both classes were to receive their targets for the real wage, the sum of their claims would be $a_0w_w + (1 - a_0w_f) = 1 + a_0(w_w - w_f) > 1$ if $w_w > w_f$.

8 This equilibrium is stable because the model can be described by a single first-order differential equation in one variable, the real wage, where $\hat{w} = \hat{W} - \hat{P}$. Substituting equations (5.1) and (5.2) and taking the derivative with respect to $w$, we can see that $\partial\hat{w}/\partial w = -(\varphi + \theta) < 0$.

9 This is the case effectively assumed in some Phillips curve models based on labour bargaining, where – in spite of additional complications involving unemployment – the real wage (or wage share) always converges to the firms' implicit target in equilibrium. See, for example, Hein and Stockhammer (2011a), Stockhammer (2011) and Blanchard (2017).

10 Wage indexation will be modelled more explicitly (and in a different manner) in section 5.2.3 below.

11 For a similar analysis of productivity growth, inflation and distribution see Lavoie (2014, pp. 561–4).

12 This point is emphasized by Storm and Naastepad (2012, 2017).

13 Since the wage share cannot fall forever, it is possible that these countries are in a gradual transition to new, lower equilibrium levels for the wage share. It is also possible that the underlying drivers of the falling wage share have been changing gradually over time, which would mean that the equilibrium itself has been shifting more or less continuously.

14 Bivens (2006) provides an econometric analysis of factors that have reduced labour's bargaining power (and ability to capture a portion of oligopolistic rents) in the US economy, focusing especially on how globalization has improved the 'fallback position of capital'. De Loecker and Eeckhout (2017) provide empirical estimates of increasing monopoly power and rising average profit markups in the US in recent decades.

**15** Logically, there is another possibility, which would be an acceleration of productivity growth in a situation in which $\beta + \gamma < 1$. This may be relevant for certain periods in certain countries, such as the late 1990s in the US, but there has not been a general acceleration of productivity growth in countries where wage shares have fallen (and in the US case, the wage share fell more in the 2000s than in the 1990s). Still another logical possibility would be a deceleration of productivity growth in a country where $\beta + \gamma > 1$.

**16** Although the empirical Phillips curve originated with Phillips (1958), the idea of an inverse relationship between the unemployment rate and nominal wage increases was stated by Robinson (1946 [1951]) in the course of explaining how a trade surplus would lead to rising money wages by increasing employment and reducing labour market slack.

**17** In reality, these two rates are not directly proportional to each other. Employment is often a 'lagging indicator' of output and utilization changes during short-run, cyclical upturns and downturns, while over longer periods employment growth can be diminished relative to output growth by increases in labour productivity. Nevertheless, $u$ and $e$ do typically have a strong positive correlation, especially if lags are taken into account, and many contemporary Phillips curve models use the output gap (similar to our utilization rate) in place of the unemployment rate, of course with the opposite sign.

**18** We do this to avoid further complicating the mathematics, but one could incorporate a $+\beta q$ term in (5.12) if desired.

**19** Raw material costs were included in the analysis of markup pricing in Kalecki (1954 [1968]), while costs of imported intermediate goods were incorporated in neo-Kaleckian 'structuralist' models by Taylor (1983).

**20** This connection was recognized by Ricardo (1821 [1951]) and Marx (1867 [1976]), as discussed in Chapter 2. For more recent analysis, see Storm and Naastepad (2012, 2017).

**21** The notion of efficiency wages originated with Smith (1776 [1976]). Bowles and Boyer (1990) argue that, to the contrary, high employment levels and rising wages could weaken worker discipline resulting in diminished work effort and reduced productivity.

**22** This is essentially the same productivity growth equation used by Barbosa-Filho and Taylor (2006, p. 395), but we make different assumptions about some of the parameter values as discussed below.

**23** This terminology is used by Taylor (2004). Marglin and Bhaduri (1990) referred to the supply-side relationship between the utilization rate and profit share as the 'producer equilibrium' (PE) curve.

**24** An exception could arise only if $q_2$ was very strongly negative, as suggested in the argument of Bowles and Boyer (1990) alluded to earlier (see note 21 above).

**25** Stockhammer (2013) found that the wage share was an increasing function of the output growth rate in a sample of Organisation for Economic Co-operation and Development (OECD) countries, although he did not test for effects of the utilization rate.

**26** This could occur if workers' ability to win wage increases in response to demand pressures is weak at low rates of utilization and employment but stronger when the latter are high (that is, when the economy is closer to 'full employment'), so that the parameter $\lambda_1$ is increasing in $u$. Alternatively, the U-shape could result if the cyclical sensitivity of productivity growth is stronger when utilization is lower (recession and recovery) and weaker when utilization is higher (expansion/boom), in which case $q_1$ would vary inversely with $u$. In the latter case, the movements in the wage share would be largely due to the cyclical behaviour of productivity, not wages – a proposition for which there is empirical support, as discussed in the next section.

**27** See Nikiforos and Foley (2012), Palley (2014), Nikiforos (2016a) and Skott (2017c) for further discussion of these ambiguities.

**28** A parallel analysis for the open economy case requires some rethinking of the modelling approach used in section 4.4.3 in Chapter 4. Instead of treating the markup as a static function of the real exchange rate, we can instead treat the markup and the associated profit and wage shares as state variables that are given in the short run and adjust in the medium run. Then, the effects of the real exchange rate on pricing and distribution can be introduced by modifying the wage and price reaction functions to include real exchange rate effects. In effect, this means that the two sources of distributional shifts (changes in monopoly power and labour costs) only affect the medium-run solution, not the short-run comparative statics. See Blecker (2011) for details.

**29** Note that this particular unstable case requires high bargaining strength of labour ($\varphi\lambda_1$ must be very large), so a 'structural' response of policy makers could be to try to weaken the labour movement in order to dampen wage responses to high rates of utilization and employment.

**30** A difference equation model in discrete time could also be developed, but we will rely on a continuous time formulation using differential equations for mathematical convenience, and for consistency with the original version in Barbosa-Filho and Taylor (2006).

**31** Although we follow the mathematical approach of Barbosa-Filho and Taylor (2006), we disagree with some of their claims about the signs of some of the parameters in these equations and we will present our own views here.

**32** Note that this is equivalent to $A_u < 1$ in the static model of the previous subsection.

**33** For empirical evidence that the total (direct and indirect) impact of a higher wage share on investment is positive in a sample of European countries, see Onaran and Obst (2016).

**34** Using long-run, historical data for the US, UK, France and Germany, Stockhammer et al. (2017) find generally positive net effects of the wage share on total investment in all four countries. For the two countries for which the requisite data are available (US and France), the authors confirm that the wage share has a negative effect on corporate investment.

**35** For a model that emphasizes this connection, see Marglin (2017).

**36** Barbosa-Filho and Taylor (2006) report that the dynamics of this system and how it produces 'moose–wolf' population cycles around an interior equilibrium were analysed in an unpublished working paper by Tu (1988), but they do not provide any details. This model differs from the Lotka–Volterra predator–prey model used by Goodwin (1967), in which (as discussed in section 2.8 of Chapter 2) the rate of change in each variable is a function of the level of the *other* variable only, thus generating a zero trace of the Jacobian which implies limit cycles (closed orbits). In the neo-Goodwin model, the rate of change in each variable depends on the levels of both variables, so there are 'own effects' as well as 'cross effects' and more varied dynamics can result. The neo-Goodwin model also differs from the original Goodwin model in using the capacity utilization rate instead of the employment rate as the indicator of economic activity. In this respect, the neo-Goodwinian approach takes aggregate demand as well as aggregate supply conditions into account and has a more neo-Kaleckian flavour.

**37** The saddle point equilibrium can be regarded as unstable, since a movement away from the equilibrium would be likely to carry the economy further and further away. There is, of course, one unique convergent path ('arm') to the saddle point equilibrium, but in this model the economy could reach that path only accidentally so the equilibrium cannot be regarded as stable (unlike in some neoclassical models, in which the assumption of perfect foresight or rational expectations guarantees that the economy operates on the convergent arm).

**38** Barbosa-Filho and Taylor (2006) claim that some empirical evidence supports such a positive effect, but it seems like a weak reed. Even if US workers briefly enjoyed such strong bargaining power, perhaps in the 1950s and 1960s, they have surely ceased to be so powerful since the advent of the neoliberal policy regime in the 1980s. The steady decline in the US labour share in the first two decades of the twenty-first century would seem to be prima facie evidence against making such an assumption for today's US economy.

**39** If demand is wage-led and distribution exhibits a wage-squeeze, so that AD slopes upward and DC slopes downward, the model will generate cycles but with a clockwise rather than a counterclockwise rotation (Kiefer and Rada, 2015).

**40** This section draws partly on Blecker (2016c).

**41** Here, we deliberately use output instead of the utilization rate for reasons that will become clear below. For expositional simplicity, we omit the role of intermediate imports and possible cyclical responses of *G* from the model of Stockhammer et al. (2011).

**42** In practice, empirical researchers often use different price indexes for exported and imported goods, and sometimes a real exchange rate variable is used instead of prices of domestic and foreign goods separately.

**43** The hypothesis that $NX_p < 0$ assumes that the Marshall–Lerner condition holds (see Appendix 9.1 in Chapter 9).

**44** Structural models in which the various equations are estimated separately are sometimes called 'single equation' estimates, but this terminology can be confusing because in fact several different equations (for consumption, investment and net exports) are actually estimated.

**45** If these effects are estimated as elasticities, then it is necessary to weight them by the shares of the various components of gross domestic product (GDP) in total GDP, usually measured at the sample means.

**46** This is especially true if autoregressive distributed lag (ARDL) or VAR/VEC methods are used, as they often are.

**47** Skott (2017c) emphasizes that if changes in exogenous factors shift both the AD and DC curves, the impact on equilibrium utilization and distribution will not depend on the slope of AD alone.

**48** Fernandez called his AD equation an IS (for investment–saving equilibrium) curve, following the usage of Marglin and Bhaduri (1990). Both the ordinary least squares (OLS) and 2SLS estimates in Fernandez (2005) showed profit-led demand in the US economy. Fernandez also found that the profit share was not significantly affected by the utilization rate after controlling for a measure of external competitiveness (the relative unit labour cost variable $z$ described in Chapter 4), so he concluded that there was no simultaneity bias in OLS estimates of the AD (IS) curve.

**49** Silva de Jesus et al. (2018) apply VAR methods to a similar model using annual data for Brazil.

**50** An exception is Cauvel (2018, Chapter 2), who finds that using general method of moments (GMM) to control for the endogeneity of the wage share and other variables does not alter the qualitative conclusion that the US economy has wage-led demand in two alternative models (for a third model, he could not find valid instruments).

**51** Decomposition methods can be used to try to assess the impact of underlying components of both output and the wage share, as in Barbosa-Filho and Taylor (2006), but the results may lack a causal interpretation.

**52** Some earlier studies that used weak econometric methods or found inconclusive results are omitted. Studies that analysed only domestic demand (consumption plus investment) are also omitted. Since the reduced form method does not yield explicit findings about wage-led versus profit-led demand, we locate the one study that has used this approach (López et al., 2011) by its findings for one key exogenous variable, the real exchange rate. Since a real depreciation (which lowers the real wage) causes a fall in output (utilization) in their estimates, we categorize this study as finding wage-led demand (but only in the short run, since the data are filtered and detrended).

**53** Stockhammer and Stehrer (2011) demonstrated that the results of a structural model of consumption and investment are highly sensitive to alternative lag lengths.

**54** Barrales and von Arnim (2017) find bidirectional causality between each of their three alternative measures of demand (discussed earlier) and the wage share, suggesting that any estimates that treat the wage (or profit) share as exogenous are subject to simultaneity bias. Palley (1994) and Kim (2013) have found significant effects of debt variables on US output, which implies that the omission of such variables could lead to omitted variable bias.

**55** One exception is López et al. (2011), who used a reduced form approach as discussed above. Another exception is Kiefer and Rada (2015), who found evidence for a 'race to the bottom' of many countries simultaneously seeking to drive their labour costs lower. A third exception is Stockhammer and Wildauer (2016), who found that the OECD countries mostly have weakly wage-led demand but financial variables were more important than distributional shifts in explaining their growth in the run-up to the 2008 crisis.

**56** Some authors, such as Onaran and Galanis (2012), instead estimate the marginal propensities to consume out of wage and profit income separately, and generally find that for the vast majority of countries, this propensity is significantly higher for wages than for profits.

**57** This study used a weak methodology for net exports: the authors estimated an export function, but treated imports as exogenous. They also found stronger positive effects of the profit share on investment in the US case than most other studies have found.

**58** In spite of these empirical findings, Razmi (2018) argues that the world economy as a whole – which is a closed system – need not be wage-led. See also von Arnim et al. (2014).

**59** Note, however, that as shown by Stockhammer and Michell (2017), a model based on real–financial interactions in which AD is wage-led can produce outcomes consistent with the patterns (in $u \times \psi$ space) produced by neo-Goodwinian models. The Stockhammer and Michell model is discussed in more detail in Chapter 7.

**60** Hamilton's critique builds on many earlier ones. Cogley and Nason (1995) and Canova (1998) demonstrated that HP filters can generate spurious cycles when no cycles exist in the underlying data. Comin and Gertler (2006) and Gordon and Krenn (2010) observed that conventional detrending methods (including, but not limited to, HP filters) put too much of the cycle into the trend. Blecker (2016c) shows that an HP filter applied to the log of US real GDP makes the Great Recession look like a relatively small fluctuation, while a survey-based index of capacity utilization from the Federal Reserve more accurately reveals how severe that recession was and also exhibits a downward long-term trend in utilization that cannot be seen in a utilization rate constructed using an HP filter.

**61** For the latter reason, analyses that find time-varying results using measures of utilization based on an HP filter, such as the threshold VAR model of Carvalho and Rezai (2016), should be interpreted with great caution (and may in fact be spurious).

**62** In principle, this problem could be solved by using 'real-time data', that is, estimates of potential output or output gaps as actually published by statistical agencies in each past time period.

**63** As explained earlier, the slope of AD has the same sign as $\upsilon_2$ under the assumption that $\upsilon_1 < 0$ for Keynesian stability.

**64** Definitionally, the wage share can be written as $\psi = WL/PY = (W/P)/(Y/L)$, so it can be seen as the ratio of the real wage to labour productivity (where the nominal wage is deflated by the same price index used in calculating real output). Thus, procyclical variations in $(Y/L)$ cause $\psi$ to vary countercyclically as long as $W/P$ does not have strongly offsetting procyclical variations (which it does not, as least in the US data).

**65** This same argument was stated previously by Lavoie (1995b, 2014). The same point was made much earlier (as Lavoie acknowledges) by Hahnel and Sherman (1982) and Sherman and Evans (1984), among others, in their critiques of empirical studies of the neo-Marxian profit-squeeze hypothesis, such as Weisskopf (1979).

**66** In these impulse responses, utilization falls in response to a shock to productivity, but that is also consistent with a rise in the wage share having a positive effect on utilization. Some of these results are sensitive to the ordering used in the impulse response functions; the results cited here are found when shocks to demand (utilization) come before productivity in the ordering so that they can have contemporaneous effects on productivity in the impulse responses. Using orderings in which utilization comes after productivity imposes the restriction that the former has no contemporaneous effect on the latter. Nevertheless, demand remains wage-led (with varying degrees of statistical significance) in some of the estimates using these other orderings as long as the real wage and productivity variables are included separately in the VAR. Cauvel's main qualitative results are not sensitive to using an HP filter or Hamilton's method to measure the utilization rate.

# Appendix 5.1 Conflicting claims and the Phillips curve

The reader who is familiar with mainstream macro theory will naturally wonder how the conflicting claims analysis of inflation in this chapter relates to the more standard treatment of inflation based on a Phillips curve. Our intention in this appendix is not to review the voluminous literature on (and debates about) the Phillips curve, but simply to show how a fairly standard type of Phillips curve can emerge from a modified version of the conflicting claims model presented in the text of this chapter.[a] We begin by using the same specifications of the distributional targets of workers and firms as given in section 5.2.3, which are reproduced here for convenience

$$\text{Workers' target wage share: } \psi_w = \lambda_0 + \lambda_1 u \tag{5.11}$$

$$\text{Firms' target profit share: } 1 - \psi_f = \eta_0 + \eta_1 u \tag{5.14}$$

For reasons that will become clear below, we must assume that $0 < \lambda_0 + \eta_0 < 1$ and $\lambda_1 + \eta_1 > 0$ to get intuitively plausible results for inflation and utilization in our solution for the Phillips curve.

At sufficiently high levels of utilization $u$, these two targets may become mutually incompatible in the sense that $\psi_w + (1 - \psi_f) > 1$, or the sum of the target wage and profit shares exceeds 100 per cent of the total social product. Note that this is equivalent to the workers demanding a higher wage share than the firms are willing to let them have: $\psi_w > \psi_f$. In such a situation, the inflation rate is assumed to increase so as to reduce the realized income of both workers and firms until their respective shares add up to only 100 per cent of the total. In this situation, each group's actual or realized income share is reduced to some extent by the increase in the inflation rate, $\Delta\hat{P}$

$$\text{Workers' realized wage share: } \psi = \lambda_0 + \lambda_1 u - \lambda_2 \Delta\hat{P} \tag{5A.1}$$

$$\text{Firms' realized profit share: } 1 - \psi = \eta_0 + \eta_1 u - \eta_2 \Delta\hat{P} \tag{5A.2}$$

where $\lambda_2, \eta_2 > 0$. The exact mechanisms through which realized wages and profits are reduced by increased inflation are not modelled explicitly here, but presumably nominal wages and prices are set without taking the increase in inflation into account and hence that increase must be regarded as unexpected.[b] In the short run, these two realized shares must add up to unity. Thus, by adding these two equations together (or, equivalently, substituting

5A.1 into 5A.2), we can solve for the short-run equilibrium increase in inflation

$$\Delta\hat{P} = \frac{\lambda_0 + \eta_0 - 1 + (\lambda_1 + \eta_1)u}{\lambda_2 + \eta_2} \tag{5A.3}$$

This equation represents a Phillips curve that is upward sloping in $u$ under the sign assumptions made above (recall that $u$ here is utilization, and is inversely related to unemployment). This is similar to many textbook presentations (for example, Carlin and Soskice, 2015; Jones, 2018) in which the output gap is used in place of the unemployment rate (so that the Phillips curve is upward sloping) and the Phillips curve is expressed in terms of the change in the inflation rate rather than the level of that rate.

In the medium run, inflation stabilizes at a constant (and indeterminate) rate, which means that the *change* in inflation must be zero.[c] Thus, setting $\Delta\hat{P} = 0$ and solving for $u$, we obtain the long-run (or medium-run) equilibrium utilization rate[d]

$$u_n = \frac{1 - (\lambda_0 + \eta_0)}{\lambda_1 + \eta_1} \tag{5A.4}$$

which (by analogy to the natural rate of unemployment, also known as the non-accelerating inflation rate of unemployment or NAIRU) can be called the NAICU (non-accelerating inflation rate of capacity utilization).[e] Assuming that this equilibrium utilization rate would become a norm or expectation for the private sector (especially if, say, monetary policy were directed towards maintaining it), we can consider it to constitute a 'normal' utilization rate, $u_n$. Under the sign restrictions stated earlier, $u_n > 0$, and to ensure also that $u_n \leq 1$, we must further assume $(\lambda_0 + \eta_0) + (\lambda_1 + \eta_1) \geq 1$. Implicitly, the model also solves for the medium-run equilibrium wage share, which can be found by substituting the solution for $u_n$ into either of the target income shares to obtain

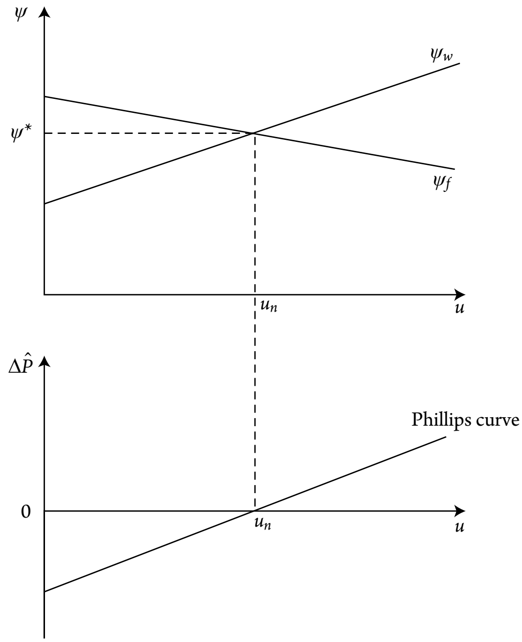
$$\psi^* = \frac{\lambda_0\eta_1 + \lambda_1(1 - \eta_0)}{\lambda_1 + \eta_1} \tag{5A.5}$$

In the special case in which the firms' target profit share is independent of the utilization rate (so $\eta_1 = 0$ and $1 - \psi_f = \eta_0$), this solution simplifies to $\psi^* = 1 - \eta_0$, which means that firms always get their target profit share (and their corresponding implicit target for the wage share) in the medium run – as assumed, for example, in Stockhammer (2011) and Hein and Stockhammer (2011a).[f]

Thus, the conflicting claims version of the Phillips curve depicts inflation as the mechanism for resolving distributional tensions (irreconcilable target

Note: This graph displays the case in which the firms' target wage share ($\psi_f$) is downward sloping; it could also be upward sloping (but flatter than the workers' target, $\psi_w$) or horizontal.

**Figure 5.7** Wage share, utilization rate and Phillips curve




shares) in the short run. In the medium run in this model, it is capacity utilization (and, implicitly, the unemployment rate) that adjusts in order to induce workers and firms to modify their distributional claims and accept targets that are mutually compatible. To visualize this intuition graphically, Figure 5.7 shows the wage share targets graphed against the utilization rate in the upper panel and the Phillips curve in the lower panel. Suppose that, in the short run, aggregate demand determines a utilization rate $u_0 > u_n$, which results in an increasing inflation rate $(\Delta\hat{P} > 0)$. In order to eliminate the (unexpected) rise in inflation, the government must adopt contractionary macro policies (a rise in the interest rate set by the monetary authority, or fiscal austerity via a tax increase or government spending cut) that reduce actual utilization, since there is no market mechanism that will automatically accomplish this. Once utilization falls to the NAICU $(u = u_n)$, inflation stabilizes $(\Delta\hat{P} = 0)$. The reduction in $u$ implies an increase in unemployment, which compels workers to moderate their wage demands and lower their target wage share. Note that this solution only guarantees a *stable* or *constant* inflation rate; the *level* of the inflation rate is indeterminate in the medium-run equilibrium, as in a standard expectations-augmented NAIRU model.

Whether a NAIRU or NAICU equilibrium exists and is stable has been the subject of much controversy. Criticisms include arguments that such equilibrium rates of unemployment or utilization may exhibit hysteresis (they can change over time in response to persistent high or low levels of actual unemployment or utilization). Another possibility is that they may be unstable if, for example, interest rate hikes induce firms to raise their target markups and profit shares, thus making the NAICU fall so that actual utilization may not fall fast enough to reach it. This series of events is discussed in more detail in section 6.4.6 in Chapter 6, where the influence of interest rates on the value and stability of the NAICU is analysed in the context of the Harrodian instability debate.[g]

The Phillips curve model presented here has some important differences from the conflicting claims model developed in section 5.2.3, even though both are grounded in the same specification of workers' and firms' distributional targets. First, the Phillips curve approach puts the main emphasis on demand-pull factors and inflationary expectations (or in this version, unexpected inflationary increases) in explaining actual inflation, while de-emphasizing cost-push factors (although the latter can be incorporated in a Phillips curve equation). As discussed earlier, the model in section 5.2.3 puts more emphasis on cost-push factors; it ignores inflationary expectations, but still takes demand-pull pressures into account.

Second, and perhaps more important, the two models tell different stories about the medium-run equilibrium distribution of income. In the conflicting claims solution (5.17), there is no presumption that the distributional targets of workers and firms are reconciled even in the medium-run equilibrium; in general, each class of agents is partially frustrated relative to its objectives, and this tension generates continued inflation in the medium run. In contrast, in the medium-run equilibrium solution given by equations (5A.4) and (5A.5) for the Phillips curve model, it is assumed that the *targets* of each class must be equalized ($\psi_w = \psi_f$ is achieved by the adjustment of utilization to $u_n$). This, then, is the real implication of a NAICU (or NAIRU) solution: it implies that distributional conflict is eliminated (or at least repressed) in the medium run, while inflation is maintained only by the inertial effects of expectations. Thus, actual inflation continues at a steady rate $\hat{P} > 0$ while there is no unexpected inflation (the change in inflation is zero, $\Delta\hat{P} = 0$) in the medium run of the Phillips curve model. The conflicting claims models covered in the body of this chapter, which do not assume the existence of a NAICU or NAIRU, imply instead that positive equilibrium inflation (at a determinate rate) results from the persistence of distributional conflict in the form of target shares that are never reconciled along with continued cost pressures (possibly, but not necessarily, aided by indexation) in the medium run.

Notes:

a The model presented here is adapted from models in Hein (2006), Stockhammer (2011), Hein and Stockhammer (2011a) and Hein et al. (2012), but with a different specification for firms. In this appendix, we do not consider interest rate effects on the firms' target profit share, as those authors do, but instead include demand effects on that share. This modification is intended to make the treatment of the target wage and profit shares more parallel and to allow workers' behaviour to influence the equilibrium profit share (in the Hein and Stockhammer version, the medium-run equilibrium profit share always equals the firms' target). Those authors also consider other important issues, such as hysteresis in the so-called natural rate of unemployment and the potential instability of this rate in response to inflation-targeting monetary policies. See section 6.4.6 in Chapter 6 for a discussion of the Hein–Stockhammer approach incorporating interest rate effects on the firms' target.

b Hence, Stockhammer (2011) and Hein and Stockhammer (2011a) also refer to the increase in inflation as the 'unexpected inflation rate', that is, $\Delta\hat{P} = \hat{P}^u$ (where the superscript '$u$' means unexpected). In effect, this assumes that nominal wages and prices are set in each period taking expected inflation into account, and any change in the inflation rate is unexpected if the expected inflation rate equals the one-period lag.

c This is equivalent to assuming that actual inflation equals expected inflation, where the latter is simply the one-period lag of actual inflation.

d Note that the same solution can be obtained by assuming that the two target income shares in equations (5.11) and (5.14) must add up to unity.

e We follow Hein (2014, p. 463) in using this terminology.

f Ball and Moffitt (2001) make essentially the same assumption in a somewhat different model of the Phillips curve by assuming that the firms' profit markup is constant.

g See also Stockhammer (2011) and Hein and Stockhammer (2011a) for further analysis and discussion.

# Appendix 5.2 Stability analysis for neo-Goodwin cycles and other cases

The linearized system of equations (5.23) and (5.24) can be written in matrix form as

$$\begin{bmatrix} \hat{u} \\ \hat{\psi} \end{bmatrix} = \begin{bmatrix} \upsilon_1 & \upsilon_2 \\ \omega_1 & \omega_2 \end{bmatrix} \begin{bmatrix} \ln u \\ \ln \psi \end{bmatrix} + \begin{bmatrix} \upsilon_0 \\ \omega_0 \end{bmatrix}$$

where we recall that $\hat{u} = d\ln u/dt$ and $\hat{\psi} = d\ln \psi/dt$. The Jacobian matrix for this system is

$$\mathbf{J} = \begin{bmatrix} \upsilon_1 & \upsilon_2 \\ \omega_1 & \omega_2 \end{bmatrix}$$

The equilibrium of this system, which can be found using Cramer's rule or other methods, is

$$\ln u^* = \frac{\upsilon_2\omega_0 - \upsilon_0\omega_2}{\upsilon_1\omega_2 - \upsilon_2\omega_1} = \frac{\upsilon_2\omega_0 - \upsilon_0\omega_2}{det(\mathbf{J})},$$

$$\ln \psi^* = \frac{\upsilon_0\omega_1 - \upsilon_1\omega_0}{\upsilon_1\omega_2 - \upsilon_2\omega_1} = \frac{\upsilon_0\omega_1 - \upsilon_1\omega_0}{det(\mathbf{J})}$$

We assume an equilibrium with economically meaningful values of the variables, which imposes restrictions on the parameter values. In theory, both $u$ and $\psi$ are positive fractions, so $u^*$ and $\psi^*$ must each lie within the interval (0, 1). In empirical applications, however, either or both of these variables may be measured by an index based on 100 in a base year or by deviations from a trend; in the latter case they would have both positive and negative values with means of zero.

The stability of this system can be analysed using the trace and determinant of the Jacobian, which are $Tr(\mathbf{J}) = \upsilon_1 + \omega_2$ and $Det(\mathbf{J}) = \upsilon_1\omega_2 - \upsilon_2\omega_1$, respectively (see Klein, 2002, pp. 474–84). Under the assumptions stated in the text, $Tr(\mathbf{J}) < 0$, because the 'own effects' of each variable on its own rate of change ($\upsilon_1$ and $\omega_2$) are both negative (self-stabilizing), while the sign of $Det(\mathbf{J})$ is ambiguous (because the 'cross effects' of each variable on the other one's rate of change, $\upsilon_2$ and $\omega_1$, can have either sign, and the magnitudes of all the partials may vary). Given a negative trace, the equilibrium is globally stable if, in addition, $Det(\mathbf{J}) > 0$, as occurs in panels (a) and (b) in Figure 5.6. However, if $Det(\mathbf{J}) < 0$, then the equilibrium is a saddle point, as occurs in panel (c) of that figure.[a]

For cycles (oscillatory behaviour) to occur, $\mathbf{J}$ must be non-diagonizable with complex characteristic roots; the equilibrium will be a stable focus if the real parts of the roots are negative and an unstable focus if they are positive (Klein, 2002, pp. 479–84). The condition for complex roots to occur is that $[Tr(\mathbf{J})]^2 < 4 \cdot Det(\mathbf{J})$, which is equivalent to $(\upsilon_1 - \omega_2)^2 < -4\upsilon_2\omega_1$. This condition can be met if the product of the cross effects $(\upsilon_2\omega_1)$ is sufficiently negative.

For example, in the case depicted in panel (a) of Figure 5.6, we must assume $\upsilon_2 < 0$ (wage-led demand) and $\omega_1 > 0$ (profit-squeeze in distribution) to get the slopes of the nullclines as shown, and if these two effects are large enough (in absolute value) then cycles of the sort drawn in that diagram will occur. It is important to emphasize, however, that this is *only* a possible and *not* a necessary outcome. If the cross effects are sufficiently weak so that $[Tr(\mathbf{J})]^2 > 4 \cdot Det(\mathbf{J})$, or, equivalently, $(\upsilon_1 - \omega_2)^2 > -4\upsilon_2\omega_1$, the equilibrium will be stable but non-oscillatory. Most of the empirical literature on neo-Goodwin cycles to date has focused on testing for the signs of these partial effects and hence the slopes of the AD and DC curves (nullclines), but has not paid much attention to whether the magnitudes of the cross effects are sufficient to imply oscillatory behaviour.

For panel (d) in Figure 5.6, we have to assume that $\omega_2 > 0$ so that the wage share is self-destabilizing, and we must also have $\omega_1 > 0$ (a profit-squeeze) so that DC slopes downward. Assuming Keynesian stability $(\upsilon_1 < 0)$ and profit-led demand $(\upsilon_2 < 0)$, AD also slopes downward. In this case, the determinant is definitely positive if DC is steeper (as shown) and negative if AD is steeper (not shown) – with the latter indicating a saddle point. If the determinant is positive, stability then hinges on the sign of the trace, which is ambiguous in this case. Instability (as shown in the diagram) will result if the trace is positive, which requires $\omega_2 > -\upsilon_1$, while stability will result in the opposite case. Again, cycles will emerge only if the condition $[Tr(\mathbf{J})]^2 < 4 \cdot Det(\mathbf{J})$ is satisfied, which requires a strongly negative product of the cross effects, $\upsilon_2\omega_1$. Thus, the type of cyclical behaviour portrayed in the diagram (in this case, unstable neo-Goodwin cycles) again depends on the particular parameter values.

Note:

a Assuming $Tr(\mathbf{J}) < 0$, a saddle point is only possible if the cross effects $\upsilon_2$ and $\omega_1$ have the same sign *and* are relatively large, so that $\upsilon_2\omega_1 > \upsilon_1\omega_2$. This in turn implies that AD and DC must slope the same way, and if they are both downward sloping then DC must be steeper in $u \times \psi$ space, as occurs in panel (c).

# 6

# Neo-Harrodian models and the Harrodian instability debate

## 6.1 Introduction

As previously noted in Chapter 3, although Harrod's thinking on macrodynamics pre-dates the publication of Keynes's *General Theory*,[1] it is commonly regarded as an early attempt to extend Keynes's short-period thinking into the long run (see, for example, Asimakopulos, 1991, Chapter 7). Harrod himself accepted this characterization, based on the fundamentally Keynesian property of his analysis: its treatment of investment by firms as independent of the saving decisions of households. As also previously noted, Harrod is sometimes thought of as the progenitor of modern growth theory, his analysis of macrodynamics marking the first step in the mid-twentieth-century renewal of interest in growth theory – a topic that had been central to the thinking of classical economists (especially Ricardo and Marx), but that garnered much less attention in the immediate aftermath of the late-nineteenth-century marginalist revolution.

Harrod's own approach to macrodynamics was discussed in detail in Chapter 3. The focus of this chapter is his influence on current theory and debate. There are two ways in which Harrod's contributions inform contemporary heterodox macroeconomics. First, Harrodian growth theory provides the basis of the 'corridor instability' view of post-Keynesian macrodynamics found in the work of neo-Harrodian growth theorists such as Skott (1989, 2010) and Fazzari et al. (2013). As discussed in Chapter 1, the predominant contemporary method of analysis in heterodox growth theory involves constructing stable, steady-state equilibrium models. The equilibria associated with these models are typically assumed to be (locally) stable – which is pedagogically convenient since, as evidenced by much of the analysis in preceding chapters, it facilitates discussion of the properties of the growth process based on the method of comparative dynamics. According to the neo-Harrodians, however, this assumption (of local stability) is inappropriate: capitalism is better viewed as a locally *unstable* dynamic process, in which instability is bounded, from above and below, by economic limits on the extent to which the system can move away from equilibrium. Together, the elements of this vision (bounded local instability) constitute the 'corridor instability' view referred to above. Local instability characterizes the motion of the system at any point in time, while the upper and lower bounds on this movement make up the 'corridor' within which macrodynamics are contained. As will become clear, these bounds also furnish explanations as to why the divergence of the system (associated with its local instability) can be 'checked' and eventually reversed – resulting in *fluctuations* in the pace of expansion as the growth rate diverges first this way, then that, from its (constant) steady-state value.

As can be inferred from this brief description, the corridor instability view rejects the notion of convergence towards a steady-state growth path – the vision of the growth process that is implicit in stable equilibrium growth models – and is instead consistent with the notion that it is best characterized as inherently *cyclical*. Hence, the neo-Harrodian approach provides an alternative account of cyclical growth dynamics, which contrasts with the neo-Goodwinian approach covered in Chapter 5. As we saw there, the neo-Goodwinian approach emphasizes profit-squeeze effects and profit-led demand, for which the empirical evidence is mixed at best. Fundamentally, the neo-Harrodian analysis offers an alternative vision in which endogenous forces of demand instability drive the cyclical behaviour of output, without relying on a neo-Marxian profit-squeeze mechanism. (In Chapter 7, we will examine models in which cyclical dynamics are driven by financial forces.)

Second, quite apart from neo-Harrodian growth theory and its development of the corridor instability view outlined above, Harrod's contributions continue to inform contemporary heterodox growth theory through the Harrodian instability debate – the question as to whether or not (local) instability of equilibrium growth outcomes associated with Harrodian dynamics can haunt *any* heterodox (post-Keynesian or classical-Marxian) growth model and, if so, whether or not this Harrodian instability can be 'tamed'. In other words, if Harrodian instability can plausibly arise in a model that otherwise appears to furnish a stable steady-state growth outcome, are there other dynamics that, once taken into consideration, eventually nullify the effects of Harrodian instability and in so doing restore the stability of the model's equilibrium growth outcomes?

The remainder of this chapter is organized as follows. Section 6.2 begins by briefly revisiting the substance of Harrod's macrodynamics first encountered, and discussed in detail, in Chapter 3. In section 6.3, we examine contemporary neo-Harrodian growth models that, by supplementing Harrod's original analysis with upper and lower bounds that contain the local instability associated with the second Harrod problem, produce the 'corridor instability' view of (cyclical) growth. Section 6.4 then turns to the possible emergence of Harrodian instability in models that are not (originally) of Harrodian pedigree, and considers various mechanisms that might be responsible for 'taming' this instability. Finally, section 6.5 concludes.

## 6.2 A review of Harrod's macrodynamics

As detailed in Chapter 3 (section 3.2), Harrod's chief concern, growing out of the experience of the 1930s depression, was that the economy would not necessarily grow at the natural rate of growth even in the long run (the first Harrod problem). To justify this concern, he observed that an alternative definition of an equilibrium growth rate was implied by Keynesian macro theory: the growth rate that would maintain equilibrium between realized saving and planned investment, the so-called warranted rate of growth. Since the natural and warranted growth rates are determined by different factors (the former by the growth of the labour force $n$ and labour productivity $q$, the latter by the saving propensity $s$ and capital to full-capacity output ratio $a_1$), the two rates are unequal in general and there is no obvious mechanism to bring them into equality with each other.

Harrod also argued that the warranted growth rate was itself an unstable equilibrium: any slight deviation of actual output growth from the warranted growth path would result in ever-greater divergences of actual from warranted growth, unless and until the economy hit upon certain ceilings or floors to output (such as full employment of labour or full utilization of capacity on the upside, and the need to replace depreciated capital or government action to boost demand on the downside). This instability of the warranted growth rate is the second Harrod problem. It states that if the economy gets away from the warranted rate of growth it will fall further away and not be able to get back into equilibrium easily or automatically. Instead, if (for example) output starts to grow at an actual growth rate that is greater than the warranted rate, firms would find that their fixed capital stocks (machinery and equipment) would be utilized at unusually high rates.[2] This would induce firms to invest in more fixed capital in an effort to reduce the rate of capacity utilization. But, this increased investment would only raise the (demand-led) actual growth rate so that it would diverge farther and farther from the warranted rate. This demonstrates the tension between microeconomic behaviour (at firm level) and macroeconomic outcomes that is responsible for propagating Harrodian instability. The same processes will unfold in reverse if actual output grows more slowly than the warranted rate to begin with: fixed capital will be underutilized, so firms will cut back on investment, leading to a further decrease in the actual rate of growth associated with movement away from (rather than back towards) the warranted rate.

## 6.3 Neo-Harrodian models

As their name suggests, neo-Harrodian models seek to build on Harrod's vision of the growth process, as recapped in the previous section. Various contemporary authors have furnished growth models of explicitly Harrodian inspiration (Skott, 1989, 2010; Fazzari et al., 2013; Ferri and Minsky, 1992; Ferri et al., 2011). While differing in the exact details of their construction, what these neo-Harrodian models have in common is their embrace of the second Harrod problem and hence their treatment of the steady-state expansion path of the economy as locally unstable. Drawing on the corridor instability view of capitalism, the challenge that they then confront is that of specifying upper and lower bounds to the divergence that Harrod's instability principle implies, the final result being models of cyclical growth.

### 6.3.1 Hicksian origins

The 'grandfather' of these neo-Harrodian models is, in fact, John Hicks. Hicks's theory of the business cycle, or what used to be called the 'trade cycle' (Hicks, 1950), explicitly acknowledges Harrod together with the fact that Harrod's instability principle alone suggests a more unstable variety of capitalism than that of ordinary experience. Hicks then develops a theory of the cycle by bounding Harrodian instability from above and from below. In fact, Hicks's contribution is all the more remarkable for anticipating the *two* key approaches to achieving this end that have informed the modern neo-Harrodian literature. The first involves adding independent ceiling and floor mechanisms that 'choke off' Harrodian instability and thereby prevent an indefinite explosion (or collapse) of the economy once the actual and warranted rates of growth differ. This approach is evident in the neo-Harrodian contributions of Fazzari et al. (2013) and Ferri et al. (2011). The second involves postulating endogenously self-limiting instability, by augmenting Harrod's dynamics so that they ultimately produce limit cycles and so recall the original neo-Marxian Goodwin model discussed in Chapter 2, section 2.8. This approach is evident in the neo-Harrodian contributions of Skott (1989, 2010).

Hicks (1950) himself puts particular weight on the first of these two approaches (although he does not neglect the second). His initial theory of the 'ceiling' is that it represents a supply-side limit imposed by the availability of productive resources – a limit beyond which it would be infeasible for the economy to expand. Before this limit is reached, however, it may be approximated by sectoral bottlenecks – supply limits in certain key sectors of the economy. Hicks was particularly concerned with the distinction between the investment goods and consumption goods sectors, and the possibility that testing the limits of production in the investment goods sector might induce a reduction in investment that, through multiplier effects, would then induce a reduction in consumption (even before the consumption goods sector reaches full-capacity output). Elsewhere, Hicks postulates that as the economy reaches its supply-determined ceiling, a turning point may be induced by the reaction of monetary policy to the onset of inflation. This introduction of monetary factors into the cycle is embellished by his emphasizing the role of bankruptcy and sudden spikes in liquidity preference as factors that exacerbate downturns – all of which can be interpreted as a precursor to the work of Hyman Minsky and the latter's eventual development of the financial instability hypothesis (Fazzari and Greenberg, 2015, p. 47). Hicks even considers the possibility that monetary policy may become active and reverse the cumulative expansion of the economy described by Harrod *before* the economy reaches its aggregate capacity constraint. This is because the central bank may be unable to accurately reckon the economy's productive capacity – an insight that anticipates the difficulties and controversy surrounding the efforts of contemporary, inflation-targeting central banks to estimate and act upon an output gap defined as the difference between the (unobserved) potential output path and the economy's actual output path.[3] As regards floor mechanisms, Hicks again gives some consideration to policy – both monetary and fiscal – staunching the dynamics of the sort of cumulative economic contraction to which Harrodian dynamics otherwise give rise. His main emphasis, however, is on an autonomous component of (investment) spending that grows at the same rate as the equilibrium rate of growth. As the rate of growth of endogenous sources of spending slows, the weight of this autonomous component in the determination of the overall growth rate increases, and this re-weighting of the (endogenous and exogenous) components of growth can arrest and eventually reverse the declining growth rate imposed on the economy by Harrod's instability principle.[4]

## 6.3.2 Informal models

As noted, contemporary neo-Harrodian models build on the theoretical and/or methodological insights of Hicks's early vision of corridor instability. Ferri and Minsky (1992) furnish an informal model that is chiefly designed to explain the absence from the post-war historical growth record (at least up to that point) of the sort of explosive or implosive growth associated with the second Harrod problem. This they ascribe to the existence of certain historically specific institutional conditions that together acted as 'thwarting mechanisms', preventing the onset of the sort of volatility that might otherwise arise from Harrodian dynamics. The first was the institutionalized growth of real wages in tandem with the rate of growth of productivity, which facilitated steady expansion of consumption spending (and reduced distributional conflict – a potential source of inflation). The second was the market power enjoyed by firms which, they argue, encouraged investment spending by creating confidence in profit expectations and so reducing both borrowers' and lenders' risk. Finally, Ferri and Minsky (1992) highlight the established role of the central bank as lender of last resort, bringing stability to the banking sector and so further solidifying the relationship between finance and the real economy. The authors warn, however, that institutions can and do change – not least as a result of the actions of those who come to see macroeconomic stabilizing mechanisms as fetters or constraints on their behaviour. Hence the Minskyan dictum that 'stability breeds instability':[5] as the pillars of macroeconomic tranquillity are eroded in complacent response to the very tranquillity they create, so 'thwarting mechanisms' (such as those outlined above) can dissolve, unleashing divergent forces (such as those associated with the second Harrod problem), bequeathing an episode or regime of greater macroeconomic volatility.

### 6.3.3 Ceilings and floors

Elsewhere in the neo-Harrodian literature, formal models of corridor instability have been developed, drawing on one or the other of the two mechanisms identified by Hicks (1950) for containing the local instability of the growth process. The 'limit cycle approach' – to which we return below – is adopted by Skott (1989, Chapter 6, 2010), while the contribution of Fazzari et al. (2013) exemplifies the 'independent ceiling/floor approach'.

Fazzari et al. (2013) set up a basic Harrodian model similar to that developed in section 3.2, wherein any departure from the warranted rate of growth ($y_w$) is associated with cumulative divergence of the actual growth rate from its steady-state value. They then set about augmenting this unstable growth model with auxiliary ceiling and floor mechanisms that contain or bound the divergence resulting from Harrodian dynamics, in such a way that the economy is seen to 'bounce off' its ceiling and floor, so that self-reinforcing contractions are eventually reversed and become self-reinforcing expansions and vice versa. The result is a model of cyclical growth.

The authors' floor mechanism involves appeal to an autonomous component of aggregate demand that is not subject to endogenous revision in response to the Harrodian dynamics that form the core of the model. The ceiling, meanwhile, is created by a supply constraint determined by the availability and productivity of resources that the demand-determined level of real output cannot logically exceed at any given point in time. Autonomous demand can also play a role in creating the ceiling of the growth corridor, however. In this case, the dynamics of the economy are entirely demand-driven, supply-side constraints playing no effective role in determining the trajectory of the economy.

As noted, the supply-determined ceiling mechanism is derived from the physical resource constraint placed on real economic activity at any point in time – or in other words, and in a growth context, Harrod's natural rate of growth. Fazzari et al. (2013) assume that the natural rate of growth is determined independently of the economy's demand-side dynamics, and that both the rate of growth of the population and labour productivity are given. These are precisely the conditions used to derive our expression for the natural rate of growth in section 1.3.2 of Chapter 1, which we can therefore reproduce here as:

$$y_N = \bar{q} + \bar{n}$$

As previously discussed, the actual rate of growth can exceed the natural rate in the short run. However, the (full-employment) potential output path traced out by the expression for the natural rate above cannot be exceeded as a matter of logic. As a result, the actual rate of growth realized, $y^a$, can be described as:

$$y^a = y \text{ if } Y < Y_N \quad (6.1)$$

$$y^a = \min[y, y_N] \text{ if } Y = Y_N$$

As can be seen from the structure of equation (6.1), the supply ceiling introduces a discontinuity in the model at the point at which the economy reaches the full-employment output path determined by the natural rate of growth: the actual rate of growth drops in this instant from $y^a = y > y_N$ to $y^a = y_N$.[6] Fazzari et al. (2013, pp. 10–11) show that for plausible parameter values, a simulation model that augments Harrodian instability with the mechanism in (6.1) will, in the instant at which the economy reaches its potential output path, reduce the actual and hence expected rates of growth to levels *below* the warranted rate. It will be recognized immediately that, per the dynamics of the second Harrod problem, these are precisely the conditions required for the onset of a process of cumulative decline in the rate of growth. What this means is that if we begin below the full-employment output path and with conditions where $y > y^e > y_w$ (where $y^e$ is the expected growth rate), the subsequent self-reinforcing upswing in the actual rate of growth brought about by Harrodian dynamics will be checked when the economy reaches its full-employment output path, at which point the actual rate of growth will be reduced to the natural rate and, with $y = y_N < y^e < y_w$, the cumulative expansion of the growth rate will be checked in a manner that involves the economy 'bouncing off' the ceiling, following which the rate of growth will begin a process of cumulative contraction. As Fazzari et al. (2013, p. 11) suggest, intuitively the full-employment output path – just like the warranted path – is unstable.

The question that we now confront is, what checks the cumulative contraction of the growth rate in such a way as to initiate another phase of its cumulative expansion (and so complete the description of a cyclical growth path)? As previously noted, the floor mechanism in Fazzari et al. (2013) is provided by an autonomous component of aggregate demand, $A$ (representing some part of household or government spending, for example, or possibly exports if we consider an open economy). It is first important to note that this modifies the warranted rate of growth, which ceases to be the constant previously described in Chapter 3 (see equation 3.6) and instead becomes *time-varying*. To see this, it is useful to begin by repeating the investment and saving equations (3.2) and (3.5) from Chapter 3:

$$I_t = a_1(Y_t^e - Y_{t-1}) \tag{6.2}$$

and

$$S_t = sY_t \tag{6.3}$$

where $a_1$ is the capital to full-capacity output ratio and $s$ is the marginal propensity to save. Now suppose that, in addition to these equations, we have:

$$A = \bar{A} \tag{6.4}$$

according to which the *level* of the autonomous component of demand $A$ is fixed. Equilibrium now requires:

$$S_t = I_t + A_t$$

and:

$$Y_t = Y_t^e$$

Combining equations (6.2), (6.3) and (6.4) under these conditions yields:

$$sY_t = a_1(Y_t - Y_{t-1}) + \bar{A} \tag{6.5}$$

$$\Rightarrow y_w = \frac{s - a_t^Y}{a_1}$$

where $a_t^Y = \bar{A}/Y_t$. Note that while $\bar{A}$ is constant, the ratio $a_t^Y$ is not, because of variation in $Y$ in the course of growth. Hence the warranted rate in (6.5) will vary with the value of $a_t^Y$ rather than remaining constant. The significance of this result will become apparent shortly.

In order to demonstrate the operation of the floor mechanism due to the introduction of $A$, suppose that the cumulative contraction of the growth rate that results from $y < y^e < y_w$ means that eventually we observe $y < 0$, and the economy begins to shrink in absolute terms. As the endogenous components of aggregate demand $E^D = C + I$ begin to shrink, so aggregate demand becomes increasingly dominated by the exogenously given component $\bar{A}$. In growth accounting terms:

$$Y_t \equiv E_t^D + A_t$$

$$\Rightarrow y \equiv (1 - a_t^Y)\hat{E}_t^D + a_t^Y \hat{A}_t \tag{6.6}$$

$$\Rightarrow y = (1 - a_t^Y)\hat{E}_t^D$$

given that $A = \bar{A}$ by assumption. As $Y_t$ declines continuously (in response to $y < 0$), so the value of $a_t^Y$ rises continuously, so that in the limit (with $a_t^Y = 1$) we will observe $y = 0$. In other words, in the limit the actual rate of growth will increase to zero (recall that we have been contemplating a stage of the growth cycle where $y < 0$). Meanwhile notice that with $a_t^Y = 1$, equation (6.5) indicates that in the limit we will have $y_w < 0$ (since $s < 1$ by assumption). What this demonstrates is that in the course of a cumulative contraction of the rate of growth, the presence of the exogenous component of spending will eventually bring about conditions where the (previously declining) actual rate of growth increases (bringing about an accompanying increase in the expected rate of growth), and increases above the level of the time-varying warranted rate.[7] This will create conditions where $y > y^e > y_w$ – conditions that will bring about the onset of a cumulative expansion in the rate of growth. As this expansion continues the economy will, of course, once again approach the ceiling imposed by its potential output path, at which point the cycle we have sketched will begin again.

One final feature of the Fazzari et al. (2013) model that is worth emphasizing is that the floor mechanism created by the exogenous component of expenditure $A$ can also create a ceiling mechanism. That is, before the economy reaches its potential output path, the decline in the value of $a_t^Y$ brought about by $y > 0$ in the expansion phase can elevate the value of the time-varying warranted rate in (6.5) above the actual rate of growth, and so create conditions for the onset of a cumulative contraction. Fazzari et al. (2013, pp. 16–17) show by means of simulations that the necessary conditions for this outcome are that $a_t^Y$ is sufficiently large initially, and that $A$ grows only slowly. Note that if these conditions are satisfied and the exogenous component of demand does, indeed, create both the ceiling and floor that contains divergent Harrodian dynamics, the economy will never test the limits imposed upon it (on the supply side) by its full-employment output path. As a result, the trajectory of the economy can be considered entirely demand-determined.

Not all neo-Harrodians are enthusiastic about the capacity of exogenous spending to contain Harrodian dynamics, however. Skott (2017a, 2017b), for example, argues that plausible values for the parameters in the saving and investment functions rule out the likelihood that an exogenous component of aggregate demand could provide an effective floor mechanism.[8] Note, however, that ceiling and floor models are not dependent upon exogenous demand to create a lower bound to contain divergent growth dynamics. Other mechanisms – and in particular, policy interventions – can also play this role, as originally discussed by Hicks (1950) and as emphasized by Fazzari and Greenberg (2015) with reference to the applicability of corridor instability models to the experience of the Great Recession.

### 6.3.4 Limit cycles

The limit cycle approach in the neo-Harrodian literature is exemplified by the work of Skott (1989, 2010).[9] Here, we begin by following Skott (2010, pp. 119–22), who presents a simplified version of the model originally found in Skott (1989, Chapter 6). The core behavioural structure of the model can be described by the following three equations:

$$y = y(\pi, e), y_\pi > 0, y_e < 0 \tag{6.7}$$

$$g = g(u), g' > 0 \tag{6.8}$$

$$\sigma = s_r \pi u \tag{6.9}$$

Equation (6.7) is the output growth function. Growth responds positively to the profit share because in the first instance the goods market adjusts in neo-Keynesian fashion: an increase in demand raises prices relative to nominal wages, increasing the profit share. According to equation (6.7), firms respond to these developments by increasing the growth of output.[10] Growth responds negatively to the employment rate, meanwhile, for two reasons. First, higher employment increases the costs of recruiting suitably qualified labour (extending search time, for example), creating a disincentive for firms to expand production. Second, higher employment changes the social relations of production: workers become more militant, which increases the costs of monitoring labour and extracting productive effort, which again creates a disincentive for firms to expand production. Equations (6.8) and (6.9), meanwhile, are investment and savings functions, respectively. The former allows for variation in the rate of accumulation in response to the rate of capacity utilization in the manner of the neo-Kaleckian models discussed in Chapter 4. The latter, meanwhile, is recognizable as the neo-Robinsonian saving equation (3.27) from Chapter 3, where (once again allowing for variation in the utilization rate) $r = \pi u / a_1$, but with $a_1 = 1$ for simplicity.[11] Using the standard goods market equilibrium condition $g = \sigma$, we can combine (6.8) and (6.9) and solve for the profit share to yield:

$$\pi = \frac{g(u)}{s_r u} = k(u) \tag{6.10}$$

Note that

$$\frac{d\pi}{du} = \frac{g' s_r u - s_r g(u)}{(s_r u)^2} = k' > 0$$

if

$$\frac{g' s_r u - s_r g(u)}{s_r u} = g' - s_r \frac{g(u)}{s_r u} = g' - s_r \pi > 0$$

This last condition is satisfied if $g' > s_r \pi$ – or in other words, if the responsiveness of investment to changes in capacity utilization in (6.8) exceeds the responsiveness of saving to changes in capacity utilization in (6.9). This violation of the Keynesian stability condition is a standard neo-Harrodian assumption.

Now note that it follows from the definition of the capacity utilization rate that:

$$u = \frac{Y}{Y_K} = \frac{a_1 Y}{K}$$
$$\Rightarrow \dot{u} = u(y - g) \qquad (6.11)$$

and from the definition of the employment rate that:

$$e = \frac{L}{N} = \frac{a_0 Y}{N}$$
$$\Rightarrow \dot{e} = e(y - n) \qquad (6.12)$$

Substituting equations (6.7), (6.8) and (6.10) into equations (6.11) and (6.12), we arrive at:

$$\dot{u} = u[y(k(u), e) - g(u)] \qquad (6.13)$$

$$\dot{e} = e[y(k(u), e) - n] \qquad (6.14)$$

If we now set $\dot{e} = \dot{u} = 0$ and ignore the trivial solutions where $e = 0$ and/or $u = 0$, it follows from (6.13) and (6.14) that

$$y^* = g(u^*) = n$$

from which (given that $g$ is monotonically increasing in $u$ in equation 6.8) it follows, in turn, that

$$u^* = g^{-1}(n)$$

and hence, recalling the relationship in equation (6.10),

$$\pi^* = k(u^*) = k[g^{-1}(n)]$$

Finally, given that output growth is monotonically decreasing in employment in (6.7), the steady-state values of $y$ and $\pi$ can be used to derive a unique steady-state value of the rate of employment, $e^*$ from equation (6.7). Note that the steady state must involve conditions of goods market clearing ($g = \sigma$) and hence the steady-state rate of capacity utilization identified above must be the normal rate of capacity utilization, consistent with the warranted rate of growth. Note also that the steady-state rate of growth is equal to the rate of growth of the labour force, which, in the absence of labour productivity growth (constancy of the labour coefficient $a_0$), is equivalent to the Harrodian natural rate of growth. Hence, in the steady-state outcome of Skott's neo-Harrodian model, there is no first Harrod problem (inequality of the actual and natural rates of growth). Finally, note that since both $y^*$ and $u^*$ are exogenously given, the steady-state rates of growth and capacity utilization are invariant with respect to both the saving rate and the profit share, and do not exhibit the post-Keynesian properties associated with either the paradox of thrift or paradox of costs.

The real purpose of Skott's model, however, is not to dwell on steady-state outcomes and their properties,[12] but instead to consider what happens in the *locale* of the steady state when the economy finds itself in disequilibrium. This issue can be addressed by considering whether or not the steady state is stable and, if not, exactly how the economy will behave in disequilibrium. Hence, we begin by writing the Jacobian of the system of differential equations in (6.13) and (6.14) evaluated at the system's steady state as:

$$\mathbf{J} = \begin{bmatrix} u(y_\pi k' - g') & uy_e \\ ey_\pi k' & ey_e \end{bmatrix} \tag{6.15}$$

It follows that:

$$Det(\mathbf{J}) = -ug'ey_e > 0$$

and:

$$Tr(\mathbf{J}) = u(y_\pi k' - g') + ey_e$$

Strictly speaking, the sign of $Tr(\mathbf{J})$ is ambiguous. Since $g' > 0$ and $y_e < 0$, it is possible that just the right constellation of parameters will give rise to the result $Tr(\mathbf{J}) = 0$. In this case, the Jacobian matrix in (6.15) will satisfy the conditions necessary for the underlying neo-Harrodian system to exhibit limit cycles around the steady-state outcomes identified earlier, the precise amplitude of which will be influenced by initial conditions.

Skott (1989, pp. 96–9) takes a different approach to obtaining the same (limit cycle) result, however. First, observe with reference back to equations (6.7) and (6.9) that we can rewrite $Tr(\mathbf{J})$ as:

$$Tr(\mathbf{J}) = u\left(\frac{\partial y}{\partial u} - g'\right) + ey_e$$

$$\Rightarrow Tr(\mathbf{J}) = u\left(\frac{\partial y}{\partial g}.\frac{\partial g}{\partial u} - g'\right) + ey_e \tag{6.16}$$

$$\Rightarrow Tr(\mathbf{J}) = u(\eta - 1)g' + ey_e$$

where $\eta = \partial y/\partial g > 1$ is consistent with the standard macroeconomic assumption that output adjusts faster than the capital stock. On this basis, we can state that the first term on the right-hand side of equation in (6.16) is unambiguously positive. Hence as long as $y_e < 0$ – the employment effect on output growth in equation (6.7) – is not too large, we will observe $Tr(\mathbf{J}) > 0$ in (6.16).[13] The system of equations (6.13) and (6.14) will, in characteristic Harrodian fashion, be *unstable*, in the sense that there will be no automatic convergence towards the steady-state solution of the system derived above.

It remains to be seen, however, exactly what the instability of the system entails – or more specifically, whether or not the system is still characterized by the continued *divergence* typical of the basic Harrodian dynamics explored in Chapter 3. Since $Tr(\mathbf{J})$ is non-zero, the dynamics of (6.13) and (6.14) do not lend themselves to easy identification of limit cycles.[14] Nevertheless, Skott (1989, pp. 97–9) is able to establish that these dynamics are, indeed, limited to a closed orbit of the steady state that involves strictly positive values of the state variables $e$ and $u$ as long as $g \leq y$ for small values of $u$, $g \geq y$ for very large values of $u$ and $e > 0$ when $\dot{u} = 0$ in (6.13). The first assumption, he argues, is plausible (and not strictly necessary); the second is actually required by virtue of the assumption of a fixed capital to full-capacity output ratio, $a_1$, which sets an upper limit on the ratio of the capital stock to the actual level of output; and the third is required in order to avoid a 'low-level equilibrium trap' that results in the model converging to a zero rate of employment (which outcome we have already dismissed as trivial in the process of analysing equations 6.13 and 6.14). Ultimately, then, the model produces self-perpetuating clockwise movements in $e \times u$ space as a result of (6.13) and (6.14) and hence (given the responsiveness of $y$ to $\pi = k(u)$ and $e$ in equation 6.7) self-perpetuating fluctuations in the rate of growth. The Harrodian instability of the warranted rate is thus transformed into a model of cyclical growth.

Since the Skott (1989, 2010) model produces clockwise movements in $e \times u$ space, and since $\pi = k(u)$ with $k' > 0$, it follows that the model will also produce clockwise movements in $e \times \pi$ space, or, in other words, *counter*-clockwise movements in $e \times \psi$ space – as in the original Goodwin model discussed in section 2.8 of Chapter 2 (and depicted in Figure 2.13) or (using *u* in place of *e*) the neo-Goodwin model in section 5.3.2 of Chapter 5 (and Figure 5.6(a)). This connection is embraced by Skott (1989, p. 85) himself, who suggests that in reduced form, his model uses class struggle effects to modify Harrodian instability in order to produce fluctuations around (rather than continuous divergence from) the warranted growth path, and in this sense owes an obvious debt to Goodwin. But as Skott (1989, pp. 85, 101–2) goes on to point out, there is no account of effective demand in Goodwin, and this is where important differences emerge between the original Goodwin model and Skott's neo-Harrodian model of cyclical growth. In the first place, the Goodwin model treats the output–capital ratio as fixed, whereas in Skott's model it is variable (thanks to variability in the rate of capacity utilization). Second, the mechanism responsible for generating cyclical growth is quite different in the two models, despite the similarities in their outcomes. In the (original) Goodwin model, the reserve army effect is *direct*: employment affects workers' bargaining power, which affects the profit share and hence profitability and hence the rate of accumulation, in a process that can be thought of as centred entirely on labour market outcomes. In the Skott model, meanwhile, the reserve army effect is *indirect*, involving labour market outcomes that are then routed through events in the goods market. Hence employment affects the rate of growth, which affects utilization rates and hence accumulation and hence profitability (through the neo-Keynesian adjustment of relative prices and hence income shares used to bring investment and saving into equilibrium). Skott (1989, p. 101) argues that the Goodwin model short-circuits this more complicated causal sequence because it assumes full capacity utilization and ignores the independence of investment from saving (in other words, it overlooks the principle of effective demand).

In view of all this, Skott's neo-Harrodian limit cycle model can be thought of as *quasi-Goodwinian* in its workings and outcomes. This characterization, in turn, gives us cause to reflect back on the contrast between neo-Harrodian and neo-Goodwinian conceptions of cyclical growth that was made in the introduction to this chapter, where it was suggested that neo-Harrodian models provide an alternative to neo-Goodwinian models. Of course, and unlike the original Goodwin model, both neo-Harrodian and neo-Goodwinian models are essentially Keynesian, as they incorporate a goods market that operates according to the principle of effective demand. In general, neo-Harrodian models do not need to rely upon the neo-Marxian profit-squeeze mechanism, which is central to neo-Goodwinian models, in order to generate growth cycles; the ceilings and floors models of Fazzari et al. (2013) and others are testimony to this fact. The Skott model blurs this distinction, however, by explicitly incorporating an (indirect) reserve army effect, through which a version of the profit-squeeze mechanism works. This suggests that, ultimately, a more nuanced interpretation of the relationship between neo-Harrodian and neo-Goodwinian models is required, which recognizes that the former may (but need not) rely on the same profit-squeeze mechanism that is central to the latter in order to generate macrodynamics characterized by cyclical growth.

### 6.3.5 Other contributions

Several models build on the 'independent ceiling/floor approach' to neo-Harrodian dynamics exemplified by Fazzari et al. (2013). Botte (2019) shows that Harrodian dynamics can arise as a generative property within an agent-based model. He then uses a supply-determined ceiling and a floor determined by autonomous demand to contain these divergent dynamics, and so produce growth cycles (in the manner of Fazzari et al. 2013). Ferri et al. (2011), meanwhile, re-think the investment function associated with the model outlined in section 6.3.3. In Fazzari et al. (2013), the 'core' model of divergent Harrodian dynamics is based, in part, on an accelerator-type investment function similar to that found in equation (6.2), where investment is undertaken to keep pace with the expected rate of expansion of real output.[15] Ferri et al. (2011) modify this investment function to include a term that allows investment to respond also to any gap between the desired and actual capital stock. The rate at which this adjustment occurs is then described as varying inversely with the real rate of interest, which, following Hicks (1965), itself varies directly with the actual rate of growth. Hence as the rate of growth increases (decreases) as a result of the model's Harrodian dynamics, the rate of interest rises (falls) and in so doing decreases (increases) investment spending by firms. *Ceteris paribus*, this last development arrests the increase (decrease) in the rate of growth emanating from the model's divergent dynamics. Ferri et al. (2011, pp. 216–17) show that under certain conditions this mechanism is capable of reversing a cumulative expansion (contraction) in the rate of growth, thus producing a pattern of cyclical growth.[16] The model thus produces results that are in keeping with the archetypal corridor instability view of growth: the rate of growth does not converge to its steady state; growth rates do not rise or fall indefinitely; and aggregate fluctuations are an endemic feature of the growth process (not the result of unexplained extraneous shocks).

The Ferri et al. (2011) model is interesting chiefly because it represents something of a 'halfway house' between the neo-Harrodian 'ceiling and floor' and 'limit cycle' models described earlier. On the one hand, while utilizing essentially the same Harrodian instability dynamic found in Fazzari et al.

(2013), turning points in Ferri et al. (2011) arise from an endogenous mechanism (growth-induced variations in interest rates that offset the changes in investment driven by the model's Harrodian dynamics). Hence the model produces aggregate fluctuations as a result of the operation of this endogenous mechanism and without the addition of independent ceilings and floors (as in Fazzari et al., 2013). In this respect, the model appears to bear comparison to the approach taken by Skott (1989, 2010). But on the other hand, and unlike the Skott (1989, 2010) models, aggregate fluctuations do not conform to a limit cycle. As Ferri et al. (2011, p. 218) demonstrate, the time path produced by their model is only apparent by means of a simulation exercise, and the results associated with this exercise reveals that neither the amplitude nor the period of fluctuations in the rate of growth is regular.

In a similar vein – and further cementing the notion that the distinction within the neo-Harrodian literature between 'ceiling and floor' and 'limit cycle' thinking is not absolute – Ryoo and Skott (2017) develop a model that dispenses with the reserve army mechanism that generates limit cycles in Skott (1989, 2010), and instead consider an economy characterized by Harrodian instability that is augmented by monetary *and* fiscal policy rules. The question the authors then explore is whether these policy rules amend or reinforce the innately divergent dynamics of the economy.

Ryoo and Skott (2017, pp. 504–18) begin by specifying a simple benchmark model in which a fixed warranted rate of growth (explicitly specified in terms of the normal rate of capacity utilization) is augmented by a differential equation that describes the actual rate of growth as increasing in the difference between the actual and normal rates of capacity utilization. This is sufficient to ensure the onset of Harrodian instability whenever the economy departs from its normal rate of capacity utilization.[17] This benchmark model is then extended to include monetary and fiscal effects. The former are incorporated by making the normal rate of capacity utilization vary positively with the difference between the actual and steady-state real interest rate, on the grounds that higher real interest rates raise the cost of carrying productive capacity and thus reduce the willingness of firms to carry excess capacity as a buffer against unexpected fluctuations in demand. The latter are incorporated by introducing a public sector that, by virtue of current and historical taxation and expenditure decisions, accumulates debt, which in turn accrues as assets in the private sector. Finally, a Phillips curve links real activity to inflation outcomes (which are assumed to be a general matter of concern to monetary policy makers).

The model is then completed by the addition of a Taylor rule and a fiscal policy rule, describing the conduct of monetary and fiscal policy, respectively.

The Taylor rule nests two special cases: *active monetary policy*, where the central bank adjusts the real interest rate in response to variations in inflation and outcomes in the real economy, and *passive monetary policy*, where the central bank ignores the state of the economy and sets a constant real interest rate (consistent with steady-state conditions). The fiscal policy rule also nests two special cases: *functional finance*, where the public sector pursues full employment, and *sound finance*, where the policy target is public debt. Ryoo and Skott (2017, pp. 518–35) demonstrate that working individually, both active monetary policy and functional finance can be stabilizing. Significantly for our present purposes, the combination of functional finance and passive monetary policy results in *growth cycles*. These are damped – so unlike the Skott (1989, 2010) model, the Ryoo and Skott (2017) model does not produce self-perpetuating limit cycles. Nevertheless, numerical simulations demonstrate that the cycles are highly persistent (Ryoo and Skott, 2017, p. 532).

Numerical simulations also show that the addition of a Taylor rule to a fiscal policy based on functional finance can further dampen the system's cycles, but need not always do so. On the contrary, the interaction of these individually stabilizing policy rules is complicated, and – somewhat paradoxically – *instability* (in the form of explosive cycles) may re-emerge in some cases.[18] At the same time, the interaction of two policy rules that, individually, fail to overcome the innate Harrodian instability of the system can suffice to meet this objective when made to work in tandem, and will do so, once again, by producing damped but persistent growth cycles.

In sum, neo-Harrodians associated with both the 'ceiling and floor' and 'limit cycle' approaches identified earlier have implicitly 'joined forces' through their revisiting of a basic theme from Hicks (1965): that *policy interventions* may play a key role in defining the 'corridor' associated with a 'corridor instability' view of capitalist macrodynamics and the occurrence of growth cycles therein.

## 6.4 The Harrodian instability debate

As noted in Chapter 1, one of the crucial debates among heterodox growth theorists concerns the treatment of the rate of capacity utilization in the long run. Those more inclined to a classical-Marxian or neo-Keynesian viewpoint argue that the rate of capacity utilization is invariant in the long run, tied to a fixed normal rate. Nevertheless, Marxists and neo-Keynesians make some allowance for the possibility of variations in the capacity utilization rate in the short run, around its long-run (normal) rate, in response to fluctuations in aggregate demand. Neo-Kaleckians, meanwhile, are inclined to treat capacity utilization as variable in the long run. Indeed, flexibility of the capacity utilization rate is necessary for the derivation of the key results of the Kalecki–Steindl model covered in Chapter 4, in which both the equilibrium growth rate and the realized profit rate are decreasing functions of the markup rate and profit share.

As we have already seen in this chapter, departure of the actual from the normal rate of capacity utilization can be associated with the conditions that give rise to Harrodian instability. Suppose that $y = y^e = y_w$ and $u = u_n$ initially, following which we observe $y^e > y_w$ (due, say, to a change in animal spirits). Recall from Chapter 3 (section 3.2.4) that consistent with this last inequality:

$$\Delta K_u > \Delta K \tag{6.17}$$

In other words, firms will find they are creating too little new capital ($\Delta K$) relative to the capital required for production to keep pace with the economy's actual rate of expansion ($\Delta K_u$), as a result of which the rate of capacity utilization will rise. The investment response this triggers (designed to correct the apparent shortfall of new capacity creation) will, as previously demonstrated, cause the expected and actual rates of growth to diverge ever further from the warranted rate, and so exacerbate (rather than correct) the supply-side imbalance between the rates of increase of new capacity creation and capacity utilization. In short, we will find that $u > u_n$ is accompanied by the onset of Harrodian instability.

It can therefore be argued that, regardless of the model at hand, *any* departure of the actual from the normal rate of capacity utilization might trigger Harrodian instability. This outcome would, of course, be destructive of the stable, steady-state equilibrium methodology that many (although not all) contemporary economists in the classical-Marxian and post-Keynesian traditions favour for the analysis of long-run growth. It is perhaps not surprising, then, that authors associated with these traditions have proposed various mechanisms designed to 'tame' Harrodian instability and, in the process, restore the stability properties of the steady-state growth rate. In this section, following Hein et al. (2011, 2012), we construct a composite heterodox growth model to explore these mechanisms in greater detail.

### 6.4.1 A generic model

To fix ideas, consider the following generic Kalecki–Robinson model, which is based on the models previously developed in Chapters 3 and 4:

$$g = f_0 + f_1 r \tag{6.18}$$

$$r = \frac{\pi u}{a_1} \tag{6.19}$$

$$\sigma = s_r r \tag{6.20}$$

Equation (6.18) is the linear neo-Robinsonian investment function from Chapter 3, with $r^e = r$ for simplicity (see equation 3.31). The expressions for the rate of profit and the saving function (in equations 6.19 and 6.20 respectively), meanwhile, are identical to equations (4.17) and (4.18) from Chapter 4. Taken together, the system of equations (6.18)–(6.20) can be considered a generic Kalecki–Robinson model in the sense that it embodies Kalecki's two-way interaction between investment and profit, assumes a neo-Robinsonian investment function, and also assumes (in the first instance) a variable rate of capacity utilization. As we will see, this simple hybrid model provides a sufficient vehicle for demonstrating both the potential onset of Harrodian instability in a (seemingly) non-Harrodian context, and both classical-Marxian and neo-Kaleckian mechanisms for taming this Harrodian instability, should it arise.

Substituting equation (6.19) into both (6.18) and (6.20), our generic Kalecki–Robinson model can be summarized as:

$$g = f_0 + \frac{f_1}{a_1}\pi u \tag{6.21}$$

$$\sigma = \frac{s_r}{a_1}\pi u \tag{6.22}$$

Equations (6.21) and (6.22) are depicted in Figure 6.1, intersecting (where $g = \sigma$) at the initial steady-state equilibrium $g^*$, $u^*$. Note that, by assumption, $u^* = u_n$. Figure 6.1 therefore depicts an outcome that would be acceptable as a characterization of long-run equilibrium to all heterodox growth theorists. This is because it *both* satisfies certain general equilibrium conditions (expectations are realized – recall that $r^e = r$ by assumption – and saving equals investment) *and* constitutes a fully adjusted position (capacity utilization is at its normal rate).[19]

### 6.4.2 Keynesian (in)stability versus Harrodian instability

Before using this model to discuss the debate over Harrodian instability, it is useful to clearly distinguish the latter from *Keynesian* stability (or instability). As will quickly become clear, we have already twice encountered the conditions for Keynesian stability in the course of this book, by virtue of

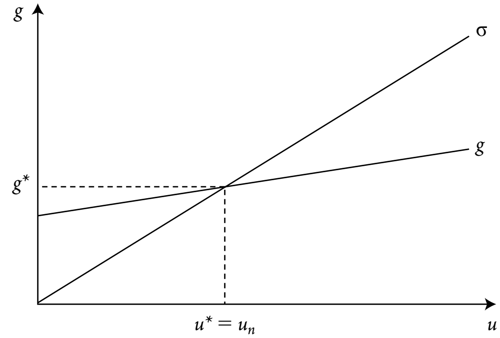



**Figure 6.1** A fully adjusted equilibrium position

the fact that some variant of the so-called *Keynesian stability condition* is a recurrent feature of all (stable) steady-state equilibrium growth models in the Kalecki–Robinson tradition.

Consider again Figure 6.1. It is clear by inspection of equations (6.21) and (6.22) that both $g$ and $\sigma$ are increasing in the rate of capacity utilization, and the depiction of the $g$ and $\sigma$ curves in Figure 6.1 is faithful to this. However, the depiction of (6.21) and (6.22) in Figure 6.1 'smuggles in' the auxiliary assumption that the $g$ curve is flatter than the $\sigma$ curve. Referring back to equations (6.21) and (6.22), this implies that:

$$\frac{d\sigma}{du} > \frac{dg}{du}$$

$$\Rightarrow \frac{s_r\pi}{a_1} > \frac{f_1\pi}{a_1}$$

$$\Rightarrow s_r > f_1$$

This is the Keynesian stability condition. We can verify that the condition above does, indeed, imply that the equilibrium depicted in Figure 6.1 is stable by once again using the technique used in Chapters 4 and 5, which identifies the responsiveness of the excess demand for goods to the adjusting variable in the system (in this case, the rate of capacity utilization). In the generic Kalecki–Robinson model:

$$EDG = g - \sigma = f_0 + \frac{f_1}{a_1}\pi u - \frac{s_r}{a_1}\pi u$$

It follows that:

$$\frac{dEDG}{du} = \frac{f_1\pi}{a_1} - \frac{s_r\pi}{a_1} = \frac{\pi}{a_1}(f_1 - s_r)$$

and hence:

$$\frac{dEDG}{du} < 0 \Leftrightarrow f_1 - s_r < 0$$

Notice that as stated above, the Keynesian stability condition is identical to the condition identified with the stability of equilibrium in the neo-Robinsonian model in Chapter 3, using the linear investment function specified in equation (3.23). The Keynesian stability condition has also made a previous appearance in this book in Chapter 4, in the guise of the condition stated in equation (4.24) for stability of the equilibrium in the Kalecki–Steindl model (and analogous conditions for the other neo-Kaleckian models covered in Chapter 4). Observe that the precise form of the stability condition in equation (4.24) differs from its appearance above and in Chapter 3. This difference arises in part because of the different specific assumptions about saving and investment behaviour in the neo-Robinsonian and neo-Kaleckian models. More fundamentally, these two types of models differ on what are the underlying adjusting variables: prices and distributive shares in the former *versus* output and capacity utilization in the latter. Nevertheless, the *behavioural substance* of the Keynesian stability condition is always the same in any model in the Kalecki–Robinson tradition. It states that the responsiveness of household saving behaviour to variations in the rate of profit exceeds the responsiveness of corporate investment behaviour to variations in the rate of profit.[20]

If the Keynesian stability condition does not hold, what results is Keynesian instability. Hence if $s_r < f_1$ in equations (6.21) and (6.22), then we will observe $dEDG/du > 0$ in our generic Kalecki–Robinson model. Since excess demand raises the capacity utilization rate and an increase in capacity utilization now raises excess demand, excess demand and the rate of capacity utilization will be subject to self-reinforcing change as a result of which the model will move ever further away from its equilibrium (where $g = \sigma$) in the event that it is displaced from this equilibrium in the first place. The onset of Keynesian instability is not necessarily fatal to models in the Kalecki–Robinson tradition, however. Neo-Kaleckians such as Lavoie (1992, pp. 288–90; 2014, pp. 351–2) have argued that periodic violations of the Keynesian stability condition can be used to characterize different growth regimes with different dynamic properties, the idea being that in this way (and consistent with the evolutionary view of the long run associated with post-Keynesian thinking in Chapter 1) Kalecki–Robinson models can be used to characterize capitalism as a sequence of discrete and historically specific episodes of growth.[21] The onset of Keynesian instability is not without cost, however.

First, the properties of the model are altered. Equating (6.21) and (6.22) using the equilibrium condition $g = \sigma$ and solving for $u$ yields:

$$u^* = \frac{a_1 f_0}{(s_r - f_1)\pi}$$

from which it follows that:

$$\frac{\partial u^*}{\partial \pi} = \frac{-a_1 f_0 (s_r - f_1)}{[(s_r - f_1)\pi]^2}$$

and:

$$\frac{\partial u^*}{\partial s_r} = \frac{-a_1 f_0 \pi}{[(s_r - f_1)\pi]^2}$$

If the Keynesian stability condition holds so that $s_r - f_1 > 0$, both of these derivatives are negative: the equilibrium capacity utilization rate responds negatively to redistribution towards profit and an increase in the propensity to save. These are recognizably neo-Kaleckian results. If, however, the Keynesian stability condition is violated so that $s_r - f_1 < 0$ (and assuming that, in this case, $f_0 < 0$, so that we continue to observe $u^* > 0$), the derivatives turn *positive*. Redistribution towards profit and an increase in the propensity to save will now raise the equilibrium capacity utilization rate. The first of these results is akin to Bhaduri and Marglin's (1990) exhilarationism or profit-led demand (as discussed in Chapter 4). The second – which posits the absence of the post-Keynesian paradox of thrift – can be likened to results associated with classical-Marxian or, given the acceptance of some classical-Marxian scholars of post-Keynesian results in the short run (see, for example, Duménil and Lévy, 1999), neoclassical macroeconomics.

Second, embracing Keynesian instability thwarts the use of the method of comparative dynamics. In the discussion above, we have been careful to point out that Keynesian instability reverses the effects of redistribution and changes in the propensity to save on *equilibrium* outcomes. But Keynesian instability renders these equilibrium outcomes unstable, complicating our interpretation of events. For example, even as Keynesian instability means that an increase in the propensity to save will increase the equilibrium rate of utilization, if the economy began in equilibrium (so that the higher propensity to save creates an initial state of disequilibrium), movement will now be *away* from the new equilibrium outcome, as a result of which the actual utilization rate will fall – as would be expected in a neo-Kaleckian model in which Keynesian stability is assumed to hold.[22] As recalled at the start of this chapter, the motivation for constructing steady-state equilibrium models in heterodox macroeconomics is partly methodological. If the equilibria associated with these models are (locally) stable, the properties of the growth process can effectively be discussed by appeal to comparative dynamics. The reconfiguration of the steady-state equilibrium induced by parametric change doubles as a sensible approximation of the outcomes that will be observed in reality, because the equilibrium is stable (that is, the economy will tend to move towards it).[23] As the example just recounted illustrates, relaxing Keynesian stability serves little purpose by way of clarifying the essential properties of the underlying model.

The upshot of these considerations is that in much heterodox growth modelling, satisfaction of the Keynesian stability condition has become a default assumption. It turns out, however, that assuming Keynesian stability does not suffice to eliminate the spectre of Harrodian instability – the main focus of this section, and to which theme we now return.

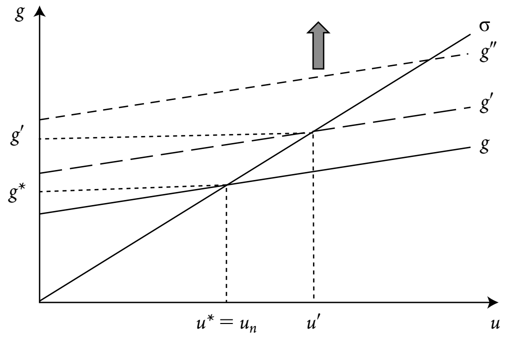
### 6.4.3 Harrodian instability

Figure 6.2 depicts the possible consequences of an initial improvement in animal spirits in the generic Kalecki–Robinson model, which can be captured by an increase in $f_0$ in equation (6.21) causing an upward shift in the investment curve to $g'$ in Figure 6.1. For neo-Kaleckians, these developments would terminate in the establishment of a new equilibrium at $g'$, $u'$ with correspondingly higher rates of growth and capacity utilization. But note that this new equilibrium is not a fully adjusted position: $u' > u_n$, and following the argument outlined earlier, this threatens the onset of Harrodian instability. If the investment response to $u' > u_n$ that is responsible for Harrodian instability does, indeed, materialize, this suggests further increases in $g$ without any further increase in $u$. This possibility is captured in our generic Kalecki–Robinson model by further increases in $f_0$ in equation (6.21). In other words, suppose that:

$$\dot{f}_0 = \alpha_0(u - u_n) \tag{6.23}$$

where $\alpha_0 > 0$ is an (assumed constant) speed of adjustment parameter that determines the exact rate of change of $f_0$ in response to departures of the actual capacity utilization rate from its normal value. Equation (6.23) ensures that $g'$, $u'$ is not, in fact, an equilibrium. It will cause further upward shifts in the investment curve in Figure 6.2, as exemplified by the shift to $g''$ and the directional arrow immediately above this curve. Since these developments only serve to widen the inequality between the actual and normal rates of capacity that fuels equation (6.23), we can now expect (other things being equal) both the rate of growth and the rate of capacity utilization to

**Figure 6.2** The onset of Harrodian instability



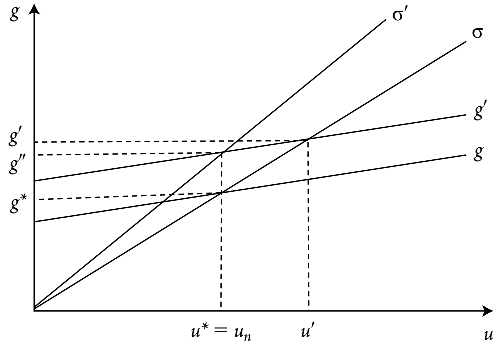

increase continually as the economy finds itself in the grip of Harrodian instability.

Other things, however, may not be equal. Both classical-Marxian and neo-Kaleckian authors have proposed a variety of auxiliary mechanisms that can 'tame' the Harrodian instability unleashed in Figure 6.2. These mechanisms are alike in so far as they restore stability to the steady-state equilibrium growth outcome of our generic Kalecki–Robinson model. They differ, however, by virtue of both the precise channels of adjustment to which they appeal and, as a result, the *properties* of the stable steady-state equilibrium outcome that is recovered. Classical-Marxian mechanisms posit adjustments that restore the rate of capacity utilization to its original value in Figure 6.2 ($u^* = u_n$), thus maintaining the central classical-Marxian postulate (previously discussed in Chapters 1 and 2) that there can be no variation in the capacity utilization rate in long-run (or 'long-period') analysis. Neo-Kaleckian solutions, meanwhile, either challenge the strict interpretation of what constitutes a fully adjusted position, or suggest that the normal rate of utilization is subject to hysteresis effects, and thus itself adjusts to accommodate changes in the actual rate of capacity utilization in the long run. Either of the neo-Kaleckian solutions maintains the long-run variability of the capacity utilization rate that is characteristic of neo-Kaleckian analysis (and hence the possibility that the growth process is wage-led).

## 6.4.4 Classical-Marxian solutions

One mechanism for taming Harrodian instability, proposed by Shaikh (2007, 2009), appeals to corporate retained earnings as an important component of total savings in the economy and, in classical-Marxian fashion, the importance of prior savings as a source of funding for planned investment expenditures. Suppose, then, that we return to the original equilibrium configuration depicted in Figure 6.1 and assume, once again, an initial increase

**Figure 6.3** The Shaikh mechanism




in the value of $f_0$ that shifts the investment curve upward from $g$ to $g'$, thus (apparently) creating a new equilibrium at $g'$, $u'$. These initial developments are now depicted in Figure 6.3. According to equation (6.23), with $u' > u_n$ in Figure 6.3, we will now witness the onset of Harrodian instability (as previously depicted in Figure 6.2). Suppose, however, that the developments discussed thus far are accompanied by:

$$\dot{s}_r = \alpha_1(u - u_n), \quad \alpha_1 > 0 \tag{6.24}$$

In keeping with Shaikh (2007, 2009) and the classical-Marxian tradition, equation (6.20) posits that realizing an increase in planned investment necessitates an increase in *savings* sufficient to fund this increased investment. In equation (6.24), we can think of the increase in the economy-wide saving rate out of profit, $s_r$, being driven by an increase in the corporate retention rate – the rate at which firms save out of profits before they are distributed to capitalist households – motivated by the rise in $g$ brought about by the initial increase in $f_0$. The result is seen in Figure 6.3, where in accordance with equation (6.24), there is a counterclockwise rotation of the saving curve as long as $u > u_n$. This rotation ceases only when capacity utilization is restored to its original (normal) value, at which point the economy is restored to steady-state equilibrium with $g' = g''$ and $u^* = u_n$. Note that the threat of Harrodian instability is eliminated because capacity utilization has returned to its normal rate, while the equilibrium growth rate has risen. Reflecting back on the contents of Chapter 2, it will be noticed that these outcomes are quite consistent with a classical-Marxian vision of macrodynamics: a higher rate of saving by capitalists boosts the rate of growth while leaving the long-run rate of capacity utilization unchanged.

The astute reader will notice that Figure 6.3 makes no allowance for the effects of Harrodian instability as previously depicted in Figure 6.2, which (in accordance with equation 6.23) would cause further upward shifts in the investment curve in Figure 6.3 as long as $u > u_n$. One possibility is that the upward shift in the investment curve and counterclockwise rotation in the saving curve depicted in Figure 6.3 are simultaneous and instantaneous, if an increase in planned investment can only be realized by a prior increase in savings – so that the shift to $g$ and the outcome $u > u_n$ are conjectural, realized adjustments (consistent with movement from the original equilibrium $g^*$, $u^*$ to the new equilibrium $g''$, $u^*$) occurring only once the rotation of the saving curve to $\sigma'$ has occurred. In this case, the dynamics in (6.23) that produce Harrodian instability are never given a chance to set in. Alternatively, if adjustments are gradual, Figure 6.3 can be thought of as a special case where $\alpha_1 > \alpha_0 = 0$. In general, the inequality in this last statement, which says that the speed of adjustment in (6.24) is greater than the speed of adjustment in (6.23), will suffice to mean that Harrodian instability materializes only temporarily and eventually disappears, as the 'taming' mechanism in (6.24) restores steady-state equilibrium conditions consistent with the normal rate of capacity utilization (see Appendix 6.1 for a formal demonstration). In other words, as long as the relative speeds of adjustment in (6.23) and (6.24) are such that Harrodian instability is eventually tamed, the special case $\alpha_1 > \alpha_0 = 0$ can be thought of as no more than a simplifying assumption. We will make simplifying assumptions similar to those made in Figure 6.3 throughout the remainder of our discussion of Harrodian instability.

A second classical-Marxian 'taming' mechanism has been proposed by Duménil and Lévy (1999). As noted by Hein et al. (2011, p. 598), Duménil and Lévy (1999) do not explicitly address Harrodian instability, but their overarching vision of capitalist macrodynamics – which posits that capitalism is unstable in levels but stable in proportions – lends itself to the interpretation that follows. Certainly, 'stability in proportions' seems consistent with the notion of a constant (normal) rate of capacity utilization in the long run, which, as we have seen, would eventually shut down the dynamics of Harrodian instability should the latter arise in the short run.

The Duménil and Lévy (1999) mechanism is illustrated in Figure 6.4. We begin with a decrease in the saving rate, $s_r$, in equation (6.22), which creates clockwise rotation in the saving curve in Figure 6.4 to $\sigma'$. Other things being equal, this creates a recognizably post-Keynesian (indeed, neo-Kaleckian!) result, increasing both the rate of growth (to $g'$) and the rate of capacity utilization (to $u'$). But as we have seen, other things may not be equal: $u' > u_n$ will provoke the onset of Harrodian instability. Suppose, however, that a second effect of $u' > u_n$ is that its tightening of the goods market provokes inflation, to which the monetary authorities respond by raising interest rates.

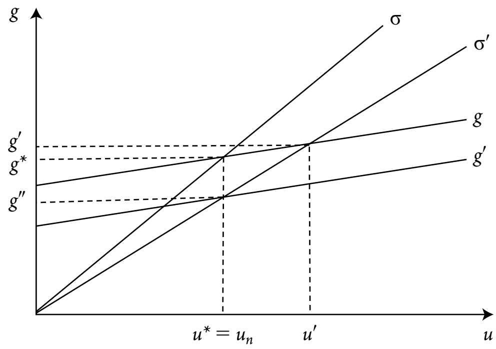



**Figure 6.4** The Duménil and Lévy mechanism

Suppose further that investment spending varies inversely with the interest rate. These auxiliary hypotheses can be summarized by the addition to our generic Kalecki–Robinson model of the equation:

$$\dot{f}_0 = \alpha_2(u - u_n), \alpha_2 < 0 \tag{6.25}$$

The operation of (6.25) causes the investment curve in Figure 6.4 to shift down. Ultimately (and again assuming for simplicity that $\alpha_0 = 0$), the investment curve will shift down to $g$ in Figure 6.4, with steady-state equilibrium restored at $g = g''$, $u^* = u_n$. Despite the post-Keynesian features of the short run, this long-run result has a distinctly classical-Marxian flavour, a decline in the saving rate ultimately being associated with slower growth at a constant rate of capacity utilization. Just as significantly for our purposes, Harrodian instability is revealed as a transitory phenomenon, the conditions from which it derives ($u > u_n$) being eliminated by the adjustments summarized by equation (6.25).

As is clear from the above, the Duménil and Lévy (1999) mechanism rests on a monetary policy reaction function similar to the Taylor rule that is associated with neoclassical New Consensus macroeconomics. Hein et al. (2011, pp. 600–601) argue that the mechanism involves numerous implicit assumptions that have neoclassical overtones – including the notions that the normal rate of capacity utilization is analogous to the natural or non-accelerating inflation rate of unemployment (NAIRU), so that $u > u_n$ is necessarily inflationary; investment spending is interest-elastic, so that an increase in interest rates suffices to curtail 'excess' demand; and rising interest rates (which increase the debt-servicing costs of firms) are not, in and of themselves, inflationary (the so-called cost-channel of monetary policy – see Lima and Setterfield, 2014).[24] At the same time, the authors note that the mechanism bears a striking resemblance to the workings of Joan Robinson's 'inflation barrier' (which, as demonstrated in Chapter 3, arises when there is real wage resistance by workers) – including both the response and effects of monetary policy to the onset of inflation that she describes (Robinson, 1956, p. 238; Robinson, 1962, p. 60).

### 6.4.5 Neo-Kaleckian solutions

The 'taming' mechanisms considered thus far are of a distinctly classical-Marxian character. Not only do they eliminate Harrodian instability and restore the stability of the steady-state growth equilibrium, they do so in a manner that gives the economy's macrodynamics identifiably classical-Marxian features. It is also possible to 'tame' Harrodian instability and restore the stability of steady-state equilibrium in our model in a post-Keynesian fashion, however – specifically, in ways that imbue our generic Kalecki–Robinson model with more neo-Kaleckian features.[25] As was briefly observed in the conclusions to Chapter 4, advocates of the Kalecki–Steindl vision of the long run have variously argued that: competing objectives within firms may mean that there is no single value of the capacity utilization rate towards which it is possible for firms to converge in the long run; the capacity utilization rate may be variable in the long run within a certain range of values rather than tending towards a unique normal rate; and what is accepted as a 'normal' rate of capacity utilization may, in fact, vary endogenously in response to variation in the actual rate of capacity utilization itself. In what follows, we will explore this thinking in more detail.

The idea that firms are complex organizations within which conflicting (and therefore irreconcilable) expectations and targets arise has been suggested by Lavoie (1992, pp. 417–21; 2002, 2003) and developed at length by Dallery and van Treeck (2011). On this view, heterogeneity of decision making *within* firms – as between different departments, for example, or on the part of shareholders on the one hand and managers on the other – may give rise to different expectations and/or targets that are ultimately irreconcilable, so that it is impossible to conceive of firms moving towards (for example) a unique value of the capacity utilization rate.

To see how such conflicts may occur, consider again the structure of our generic Kalecki–Robinson model. Suppose we reinstate the original neo-Robinsonian expectational structure of the investment function by replacing equation (6.18) with (3.31) from Chapter 3, rewritten here as

$$g = f_0 + f_1 r^e \qquad (6.18')$$

We can no longer use equation (6.19) to rewrite equation (6.18′) in terms of the rate of capacity utilization unless we rewrite the former as

$$r^e = \frac{\pi u^e}{a_1} \qquad (6.19')$$

Equation (6.19′) requires that expectations – in this case, $u^e$ and $r^e$ – are formed consistently by firms – that is, in accordance with the underlying structural relationship between the rates of profit and capacity utilization in equation (6.19). This is a reasonable requirement, but it is not inevitable that it will be satisfied. For example, in an environment of uncertainty, the structure of the firm itself may only be imperfectly understood by those who run it, with the result that $u^e$ and $r^e$ are formulated independently of the relationship between profit and capacity utilization in (6.19). In effect, this will give rise to two different values of $u^e$. Alternatively, conflict between constituents of the firm – such as managers and shareholders – may result in different expectations and targets. For example, managers planning investment in fixed capital may build into equation (6.18′) a value of $u^e$ different from that consistent with (6.19′) given the profit expectations of shareholders who are focused on the short term. If firms are *resilient* to such inconsistencies – that is, if they are not driven out of business by, in this case, failure to recognize an underlying accounting relationship between the rate of profit and the rate of capacity utilization – then by hypothesis such practices will survive. As such, it will not be possible to characterize firms in terms of uniquely valued target values of variables and/or expected values of variables.

The idea that firms are willing to tolerate variation in the capacity utilization rate within a range around the normal rate – that, in effect, the normal rate of capacity utilization is an *interval* rather than being single-valued – appears in Dutt (1990, pp. 58–9) and Lavoie (1992, pp. 327–32, 417–22). The idea is developed in detail by Dutt (2010a), who appeals to G.L.S. Shackle's concept of 'potential surprise'. Shackle's potential surprise function posits that in an environment of uncertainty, decision makers will regard disappointment of their expectations (formulated on the basis of limited information) as a normal state of affairs, so that such disappointment – at least within a 'normal' range – will be considered unexceptional and invite no behavioural response. Only larger deviations from forecasted outcomes will have the effect of provoking behavioural change. Recall that deviation of the actual from the normal rate of capacity utilization can be associated, in Harrodian macrodynamics, with deviations of the expected from the warranted rate of growth – and hence deviations of the actual from the expected rate of growth. What Dutt's appeal to the potential surprise function amounts to, then, is the claim that there is a range of expectational disappointment (and hence a corresponding range of discrepancy between the actual and normal rates of capacity utilization) that will be tolerated by firms, without it eliciting the changes in investment behaviour characteristic of the dynamics of Harrodian instability.

Similar arguments are made by Setterfield (2018), drawing upon Harrod's own work. As noted in Chapter 3, Harrod consistently opposed characterization of his instability principle as a knife edge, activated by the slightest departure from the warranted rate of growth. According to Harrod (1939), the operation of the instability principle involves a *reaction time* (of about six months) before firms adjust their investment behaviour, while in his later work, the *size* of departures from the warranted rate (or in other words, departures of the actual rate of capacity utilization from its normal rate) are instrumental in instigating Harrodian instability (Harrod, 1970, 1973). Developing this second theme, Setterfield (2018) argues that rather than resting on a 'knife edge' at a specific normal rate, firms rest on a 'shallow dome' based on rules of thumb that specify tolerable intervals of variation in actual utilization rates around the normal rate. These tolerable intervals are, themselves, based on recent past observation of fluctuations in capacity utilization. Only 'extreme' fluctuations that push actual capacity utilization outside the tolerable interval will attract attention and provoke firms to change their investment behaviour in a manner consistent with the onset of Harrodian instability.

In sum, both Dutt (2010a) and Setterfield (2018) suggest there is a range of variation in $u$ about $u_n$ that is deemed unremarkable, so that variations in capacity utilization within this range will not provoke a behavioural response. This suggests there are discontinuities in the investment function that equation (6.23) fails to capture, and that a better representation of firms' behaviour would be:

$$\dot{f}_0 = 0 \text{ if } |u - u_n| < \bar{c} \qquad (6.23')$$

$$\dot{f}_0 = \alpha_0(u - u_n) \text{ otherwise}$$

where $2\bar{c}$ is the size of the interval within which the rate of utilization can vary without causing a change in investment behaviour.[26]

The consequences of these first two neo-Kaleckian taming mechanisms are illustrated in Figure 6.5. As in previous exercises, we begin at the fully adjusted, steady-state equilibrium $g^*$, $u^* = u_n$ and contemplate an increase in $f_0$ that shifts the investment curve upwards to $g'$, establishing a new equilibrium at $g'$, $u' > u_n$. With the actual exceeding the normal rate of capacity uti-

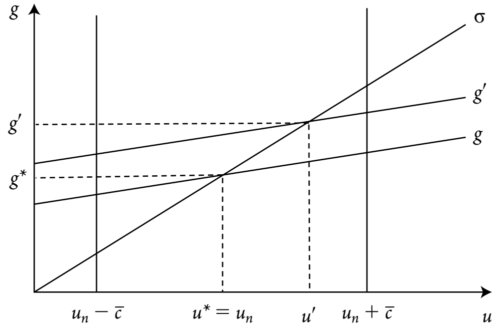



**Figure 6.5** Variation in the rate of capacity utilization around a fixed normal rate

lization, we are seemingly now inviting the onset of Harrodian instability. But suppose that, following Lavoie (1992, pp. 417–21; 2002, 2003) and Dallery and van Treeck (2011), $u^*$ and $u'$ are but two examples of the conflicting target values of variables that characterize firms. Alternatively, and following Dutt (2010a) and Setterfield (2018), note that in Figure 6.5, $u' < u_n + \bar{c}$. In other words, $u'$ falls within the interval of values that firms will deem an acceptable approximation of $u_n$ itself, and will therefore not induce any effort to accumulate additional capacity with the express aim of reducing the actual rate of capacity utilization. In either case, the economy will remain at the new equilibrium $g'$, $u'$ established by the initial shift upward in the investment curve to $g'$ and without experiencing the onset of Harrodian instability, so that this new equilibrium is effectively (if not literally) a fully adjusted position. Note also that $g'$, $u'$ is exactly the sort of long-run equilibrium outcome – and with the same properties (for example, displaying the paradoxes of thrift and costs) – that would emerge from the Kalecki–Steindl model developed and discussed in Chapter 4 (section 4.3).

By way of a response to the above, it might be argued (for example) that while the arguments advanced by Dutt (2010a) and Setterfield (2018) are all very well in theory, if $\bar{c}$ is small and is therefore well approximated (in the first instance) by a value of zero, the notion that the normal rate of utilization is an interval collapses. In this case, the original neo-Harrodian interpretation of affairs – as illustrated in Figure 6.2 – is restored. Another way of looking at this debate over the size of $\bar{c}$ is that it cannot be resolved in theory but must instead be considered an empirical matter.

The third and final neo-Kaleckian taming mechanism rests on the notion that the 'normal' rate of capacity utilization may, in fact, vary endogenously in response to changes in the actual rate of capacity utilization. This idea begins with the observation that classical-Marxian, neo-Keynesian and neo-Harrodian models all involve an implicit closure of the form:

$$u_n = \bar{u}_n \tag{6.26}$$

Equation (6.26) posits that the normal rate is *fixed* independently of variations in the actual rate. But neo-Kaleckians contend that rather than being truly parametric in this fashion, the normal rate of capacity utilization exhibits *deep endogeneity* (Setterfield, 1993) and is thus subject to *hysteresis effects*. Hysteresis is a form of path dependency which posits that variations in the actual values of variables can affect the underlying structure of the systems from which they derive, and hence their long-run equilibrium values. Such hysteresis effects apply to (among other things) behavioural norms, such as the normal rate of capacity utilization.

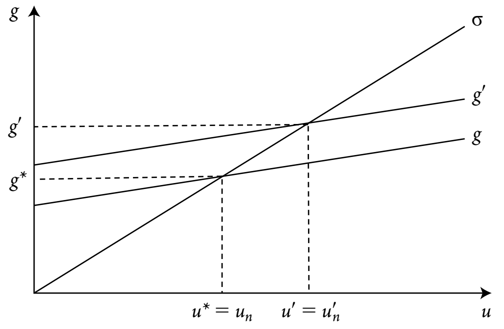
In general, the idea is that norms anchor behaviour so that once established, the normal value of a variable will influence behaviour in such a way as to bring the actual (observed) value of the variable into conformity with the normal value itself. But actual outcomes can be expected to deviate from their normal values from time to time. Much of the time these deviations will be small and unremarkable, having no effect on the norm itself, towards which actual outcomes will eventually revert. But if the deviations are persistent or extraordinary, decision makers – who are the ultimate architects of the norms – may be induced to rethink what is 'normal'. This process of adjustment draws on the idea that in an environment of uncertainty, norms are social inventions (designed to guide behaviour in the absence of an optimal course of action) that derive from the objective circumstances of actual experience – rather than just being 'made up'. In the circumstances described, the norm becomes the adjusting variable, changing in response to actual outcomes (rather than vice versa), so giving rise to a type of hysteresis effect. Or, in colloquial terms, a change in economic circumstances (such as a prolonged period of chronically depressed demand) can lead to the acceptance of a 'new normal' for the rate of capacity utilization.

Neo-Kaleckian applications of these principles to the normal rate of capacity utilization are numerous (see, for example, Dutt, 1997, 2009, 2010a; Lavoie, 1995b, 1996, 2010; Nikiforos, 2013; Setterfield and Avritzer, 2019). They can be summarized by replacement of equation (6.22) with the alternative closure:

$$\dot{u}_n = \alpha_3(u - u_n), \;\; \alpha_3 > 0 \tag{6.27}$$

Equation (6.27) states that the normal rate of capacity utilization will increase (decrease) whenever the actual rate rises above (falls below) the normal rate itself.[27]

**Figure 6.6** Hysteresis in the normal rate of capacity utilization




The effects of adding equation (6.27) to our generic Kalecki–Robinson model are captured in Figure 6.6. Suppose that, once again, we begin at the steady-state equilibrium (and fully adjusted position) $g^*$, $u^* = u_n$, and that an increase in $f_0$ then shifts the investment curve upward to $g'$. This will establish a new equilibrium at $g'$, $u' > u_n$, the latter result provoking the onset of Harrodian instability. But suppose that the dynamics of the system now involve equation (6.27). With $u' > u_n$, the normal rate of utilization will begin to rise, as persistent experience of a higher actual rate of utilization shapes firms' perceptions of what is normal and hence the value of the normal rate of capacity utilization. Ultimately (and again assuming for simplicity that $\alpha_0 = 0$), the normal rate of utilization adjusts to $u_n' = u'$, at which point the steady-state equilibrium $g'$, $u'$ is rendered a fully adjusted position and there is no further prospect of Harrodian instability. Once again, this is exactly the sort of long-run equilibrium outcome (with the same properties) anticipated by the Kalecki–Steindl model presented in Chapter 4.

According to Skott (2012), however, while its mathematical implications are clear, equation (6.27) is mechanical and lacks adequate behavioural foundations. The preceding discussion suggests there *can be* a behavioural basis for thinking of norms as enduring but ultimately transmutable and, in particular, that actual outcomes affect the formulation and reformulation of these norms. Whether or not this applies to the normal rate of capacity utilization, however, and whether or not neo-Kaleckians have adequately articulated how and why this is so, are potentially different matters. In this regard the contribution of Setterfield and Avritzer (2019) is useful, since it purports to address the behavioural basis for equation (6.27) directly. The authors argue that this behavioural basis is found in one of the main purposes of having a normal rate of utilization less than unity – because excess capacity serves as a device that insulates firms against loss of market share in the face of unforeseen increases in final demand.[28] Setterfield and Avritzer (2019) argue that there are episodes or regimes of macroeconomic performance during which lower actual rates of capacity utilization are accompanied by higher volatility in the capacity utilization rate – the latter suggesting the need for a lower normal rate of capacity utilization to create the desired insulation against shocks. On this view, lower actual capacity utilization and lower normal capacity utilization go hand in hand, the 'missing behavioural link' between the two being the volatility of the capacity utilization rate, which both *correlates* with lower average values of the actual utilization rate and *causes* lower values of the normal rate.

Part of the debate between neo-Kaleckians and neo-Harrodians with respect to hysteresis in the normal rate of utilization may also centre on the functional form of (6.27), which, as previously noted, describes *continuous* change that is not entirely faithful to the concept of hysteresis. While this has previously been justified as a simplifying assumption, it should be noted that many proponents of hysteresis are highly critical of continuous approximations of the process such as that found in equation (6.27) (see, for example, Amable et al., 1993, 1994, 1995).

### 6.4.6 Other contributions to the debate

Despite their earlier criticisms of the model (Hein et al. 2011, pp. 600–601) as discussed above, Hein et al. (2012, pp. 158–65) contemplate an extension of the Duménil and Lévy (1999) model in which monetary policy, motivated by fear (or the actual onset) of inflation, is responsible for taming Harrodian instability. They incorporate an explicit (negative) effect of the real interest rate into the investment function so that, à la Duménil and Lévy (1999), an increase in the rate of interest reduces the rate of growth (holding other factors constant). However, this is accompanied by a negative real interest rate effect on *saving*, as higher interest rates redistribute gross profit income from firms (who retain all earnings for the sake of investment financing) to rentiers (who save only part of their income). Hence in the Hein et al. (2012, pp. 158–65) model an increase in the interest rate has an ambiguous effect on goods market outcomes, simultaneously reducing investment *and* the propensity to save, so that the equilibrium rate of capacity utilization may rise or fall.

At the same time, the authors re-conceptualize the normal rate of capacity utilization as the non-accelerating inflation rate of capacity utilization (NAICU). The NAICU was first introduced in Appendix 5.1 in Chapter 5, where it was remarked that it can come to be regarded as the normal rate of capacity utilization by firms if, for example, it is consistently targeted and maintained by monetary policy. The version of the NAICU model developed here, however, following Hein et al. (2012), is articulated in terms of real wage targets rather than wage share targets (as in equations 5.11 and 5.14).[29] Recall that the function of the NAICU is to reconcile competing claims in a conflicting claims inflation model, so as to bring about a constant rate of inflation. Based on Hein et al. (2012, p. 160), but modified to make the notation more similar to that used in Chapter 5, we can describe the real wage targets of workers and firms as

$$w_w = \lambda_0 + \lambda_1 u \tag{6.28}$$

$$w_f = (1 - \eta_0) - \eta_2 id \tag{6.29}$$

Note that we have omitted the $\eta_1 u$ term found in equation (5.14) from (6.29), since (as noted in Chapter 5) it is not included in Hein's original model. Instead, we restore Hein's negative effect of debt service burdens ($-\eta_2 id$), where $\eta_2 > 0$ is a coefficient, $i$ is the interest rate and $d$ is the debt-to-capital ratio, so that $id$ measures the interest payments that indebted firms are obliged to make to rentiers (normalized by the capital stock). Equation (6.28) describes the workers' real wage target as increasing in the rate of capacity utilization. The idea here is that the tightening of the goods market denoted by a rise in $u$ will (by assumption) be accompanied by a tightening of the labour market which will increase the bargaining power and/or aspirations of workers, and hence the value of their real wage target. Equation (6.29), meanwhile, states that the implicit real wage target of firms (corresponding to their underlying target for the markup) is decreasing in the real interest rate. This is because a rise in the interest rate increases firms' overhead costs – specifically, their debt-servicing payments, $id$. Consistent with the theory of markup pricing, firms will seek to cover these higher overheads by raising the value of the markup and hence the gross profit share – or in other words, by reducing their implicit target for the real wage.

Equilibrium requires that we find $u$ such that $w_w = w_f$, thereby eliminating conflict over the real wage so as to make the rate of inflation constant.[30] Imposing the condition $w_w = w_f$ on equations (6.28) and (6.29) and solving for the resulting (normal) rate of capacity utilization (now understood to represent the NAICU) yields:

$$u_n = \frac{1 - (\lambda_0 + \eta_0) - \eta_2 id}{\lambda_1} \tag{6.30}$$

where a positive numerator is required for an economically meaningful (positive) solution. It is clear by inspection of (6.30) that an increase in the interest rate *i* will, for any given debt ratio *d*, *reduce* the normal rate of capacity utilization (NAICU). The intuition is straightforward. If firms react to a rise in interest rates by increasing their markups and hence *reducing* their target real wage, the resulting difference between workers' and firms' wage targets will create distributional conflict, giving rise to an increase in the inflation rate ($\Delta\hat{P} > 0$). Stabilizing the inflation rate once again (restoring $\Delta\hat{P} = 0$) then requires a lower rate of capacity utilization (and hence employment) to reduce the workers' target real wage to the new (lower) level targeted by firms. This will eliminate the conflicting income claims that are otherwise a source of rising (unexpected) inflation, but at the cost of depressed economic activity and employment.

Recall that in the Duménil and Lévy (1999) model, an increase in the interest rate will reduce the rate of capacity utilization and, in so doing, move the equilibrium rate of capacity utilization back towards its (fixed) normal rate. In the Hein et al. (2012) model, however, an increase in the rate of interest can either raise or lower the equilibrium rate of capacity utilization (as determined by its effects on aggregate demand), but will *necessarily* lower the *normal* rate of capacity utilization (now understood to represent a NAICU). The range of possible outcomes is bewildering! Hein (2006, 2008, pp. 153–67) shows that the interactions of the model can produce equality between the equilibrium and normal rates of utilization, but can also produce constant, damped or divergent cycles in, or even monotonic decline of, both capacity utilization rates. Furthermore, the fact that the value of the normal rate of capacity utilization is affected by monetary policy means that any final (fully adjusted) equilibrium position that does emerge can have neo-Kaleckian properties, rather than possessing the strict classical-Marxian properties associated with long-run equilibrium in Duménil and Lévy (1999).

Franke (2018) also revisits the Duménil and Lévy (1999) model, although the focus of his analysis is exclusively on the potential role of monetary policy as a stabilizing device that eliminates Harrodian instability, rather than the (neo-Kaleckian or classical-Marxian) properties of the long-run equilibrium itself (should it prove to be stable). According to Franke (2018, p. 595), 'the possibly stabilizing effects of monetary policy are less obvious than the [Duménil and Lévy (1999) model] suggests'. This, he argues, is because Duménil and Lévy (1999), following the neoclassical New Consensus with which their thinking has been compared, posit an unambiguously negative effect of a rise in the real interest rate on aggregate demand, thus overlooking two important considerations: first, the fact that the effects of monetary policy take time to work themselves out, so that any sizeable effect of changes in monetary policy on investment will likely materialize only after a delay; and second, the fact that a rise in the interest rate that reduces the real rate of return will only be of concern to firms if the rate of return is reduced below some desired or hurdle rate (measured by the risk-free rate of interest).

Suppose, then, that we replace equation (6.25) from the Duménil and Lévy (1999) model with:

$$\dot{f}_0 = \alpha_4(r - i - i^*), \alpha_4 > 0$$

where $i^*$ denotes the risk-free rate of interest. According to this expression, if the net rate of return $r - i$ exceeds the risk-free interest rate $i^*$, investment is encouraged, whereas if this difference turns negative, investment will decline. On the basis of equation (6.19), we can rewrite this Franke mechanism as:[31]

$$\dot{f}_0 = \alpha_4\left(\frac{\pi u}{a_1} - i - i^*\right) \tag{6.25'}$$

Now suppose that $u$ rises above its normal rate bringing about the onset of Harrodian instability *and* provoking a monetary policy response (specifically, an increase in the interest rate). Will the latter act as a stabilizing force in the face of the former? According to (6.25′) the answer is 'maybe, but not necessarily'. First, note that under the hypothesized conditions (an increase in the rate of capacity utilization and the rate of interest) the net rate of return $(\pi u/a_1) - i$ in equation (6.25′) may actually *increase*: it may take a succession of interest rate hikes in order for the net rate of return to decline, giving expression to Franke's concern with the possibility that monetary policy will have its desired effects on the economy only with some delay.

Furthermore, note that as long as the net rate of return continues to exceed the risk-free rate in (6.25′), we will have $\dot{f}_0 > 0$, so that (6.25′) will *contribute to* (rather than redress) Harrodian instability. This gives expression to Franke's concern that firms' investment behaviour depends not only on the net rate of return per se, but also on a comparison of the net rate of return with an alternative (the risk-free interest rate). Of course, a sufficiently large increase in interest rates may eventually reduce the net rate of return below the risk-free rate, and at that point, with $\dot{f}_0 < 0$ in (6.25′), and provided this is sufficient to offset the dynamics of Harrodian instability in equation (6.23), monetary policy will play the stabilizing role originally attributed to it by Duménil and Lévy (1999) (and the neoclassical New Consensus). The point remains, however, that in Franke's model, monetary policy may or may *not* be stabilizing in the face of Harrodian instability. 'Thus, instead of virtually suffocating the Harrodian mechanics by slipping the standard stability effects of the New Consensus over their head, so to speak, stability becomes an open issue again' (Franke, 2018, p. 613).

Another recent development has involved the incorporation of an exogenous component of aggregate demand into the neo-Kaleckian growth model, which, in a manner analogous to appeal to exogenous spending as a floor mechanism in the neo-Harrodian models of Fazzari et al. (2013) and Ferri et al. (2011), can act as a stabilizing force in the face of Harrodian instability (Allain, 2015, 2019; Lavoie, 2016). But unlike the neo-Harrodian models, where exogenous spending contributes to the creation of a corridor within which growth cycles associated with Harrodian instability occur, appeal to exogenous spending in neo-Kaleckian models follows the lead of Sraffian authors such as Serrano (1995) and Freitas and Serrano (2015), where it creates a new adjusting variable that, in the manner of the various neo-Kaleckian mechanisms discussed earlier, restores stability to the steady-state growth rate.[32]

The basis for this result can be demonstrated by introducing a new variable $a^K = A/K$ into our generic Kalecki–Robinson model, where $A$ once again denotes an autonomous component of expenditure, as in the neo-Harrodian model of Fazzari et al. (2013) discussed in section 6.3.3. Following Freitas and Serrano (2015) and Lavoie (2016), we assume here that $A$ specifically represents autonomous consumption spending by capitalists.[33] Referring back to equation (6.22), which is derived from capitalists' saving behaviour, incorporating this new variable into our model requires that we rewrite the savings function as:

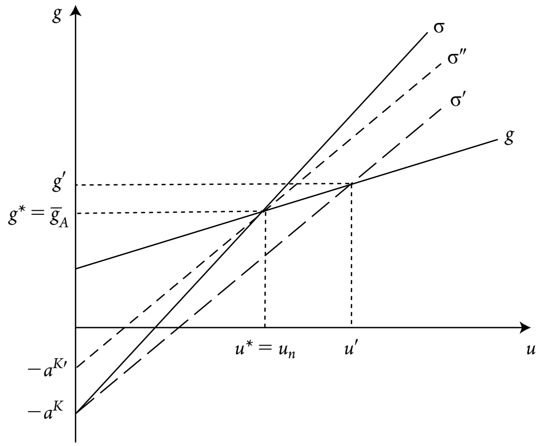
$$\sigma = \frac{s_r}{a_1}\pi u - a^K \tag{6.22$'$}$$

Equation (6.22′) says that if there were no profit income ($\pi u/a_1 = 0$), capitalists would have to be dissaving ($\sigma = -a^K$) in order to finance the autonomous consumption spending in which they are now assumed to engage regardless of income. Furthermore, note that:

$$\begin{aligned} a^K &= \frac{A}{K} \\ \Rightarrow \dot{a}^K &= a^K(\hat{A} - \hat{K}) \qquad (6.31) \\ \Rightarrow \dot{a}^K &= a^K(\bar{g}_A - g) \end{aligned}$$

where $\bar{g}_A$ denotes the exogenously given rate of growth of the autonomous component of spending, $A$. It is now clear that in addition to modifying our

**Figure 6.7**
Autonomous demand and Harrodian instability




saving relation (to become equation 6.22′), the introduction of autonomous spending creates a new dynamic adjustment process (described by equation 6.31) as a result of which the variable $a^K$ adjusts endogenously to any difference between the exogenously given rate of growth of autonomous spending and the actual rate of growth. This, in turn, means that any steady-state equilibrium we derive must satisfy the condition $\bar{g}_A = g$, which is necessary for $\dot{a}^K = 0$ in (6.31).

Bearing all this in mind, Figure 6.7 depicts an initial steady-state equilibrium of our model denoted as $g^* = \bar{g}_A$, $u^* = u_n$. Now consider a reduction in the propensity to save out of profits $s_r$, which causes the saving curve to rotate to $\sigma'$. This establishes a new equilibrium at $g'$, $u'$, and since $u' > u_n$, we are once again primed for the onset of Harrodian instability thanks to the operation of equation (6.23). But at the new equilibrium, we also have $g' > \bar{g}_A$, as a result of which the value of $a^K$ must now begin to fall in equation (6.31), causing the saving curve to shift upward. If we once again assume for simplicity that $\alpha_0 = 0$ in equation (6.23) in order to focus attention exclusively on the workings of the new dynamic in equation (6.31), it can be seen in Figure 6.7 that the adjustments of the model will only be complete when the saving curve has shifted up to $\sigma''$, at which point, with $g = g^* = \bar{g}_A$, we will observe $\dot{a}^K = 0$ in equation (6.31). In addition to this last result, we have once again restored equality between the actual and normal rates of capacity utilization, as a result of which the spectre of Harrodian instability has been removed.

In short, the introduction of an autonomous component of spending ($A$) introduces a new adjusting variable into the system ($a^K$) which is capable of taming Harrodian instability. Note, however, that in our final steady-state equilibrium, the rate of capacity utilization has returned to its original (normal) value: capacity utilization is no longer variable, so the properties of the equilibrium will be classical-Marxian rather than neo-Kaleckian. Lavoie (2016) argues that the properties of the growth process are nevertheless neo-Kaleckian *on average*, if we consider outcomes over the course of the traverse from the initial equilibrium configuration to the final equilibrium configuration in Figure 6.7. This is because neo-Kaleckian properties are evident in the disequilibrium states of the model depicted in Figure 6.7. For example, following the initial reduction in the propensity to save in Figure 6.7, the rate of growth rises in accordance with the paradox of thrift and will remain elevated during the discrete period that it takes for $a^K$ to adjust to its new equilibrium value of $a^{K\prime}$. Whether or not this argument is compelling is a matter that we will set aside for now, and take up again when we revisit the Sraffian-inspired developments discussed here in Chapter 7, as part of our account of contemporary supermultiplier analysis and the associated turn towards exogenous growth theory in heterodox macrodynamics.

Finally, it is important to draw attention to the contribution of Franke (2019). None of our analysis thus far suggests that the various mechanisms capable of taming Harrodian instability are mutually exclusive; several of them may operate at once. Intuitively, this would seem only to reinforce the possibility that Harrodian instability, should it materialize, will be tamed. But Franke cautions against such inference, by demonstrating that two of the mechanisms considered above that, when operating individually, are capable of taming Harrodian instability, actually become *jointly destabilizing* when they operate simultaneously![34]

The two mechanisms in question are monetary policy and the existence of an exogenous component of aggregate demand. Franke (2019) considers a system in which both equation (6.25′) and equation (6.31) are operative. Recall that, per Franke (2018), equation (6.25′) is not *necessarily* stabilizing. A necessary condition for stability is that any shock to the system that increases the rate of capacity utilization induces a sufficiently large increase in the real rate of interest to offset the inevitable increase in $\pi u/a_1$ and so reduce the net rate of return $(\pi u/a_1) - i$. Let us assume, then, that this necessary condition is satisfied. Now suppose we contemplate a shock that raises the value of $f_0$ and so increases the equilibrium rates of growth and capacity utilization (in the manner originally contemplated in Figure 6.2). This increase in the growth rate will activate equation (6.31), bringing about a reduction in $a^K$.

In and of itself this development is stabilizing: a fall in $a^K$ will, as demonstrated in Figure 6.7, reduce the rates of growth and capacity utilization.

But as we have just assumed (in order to satisfy the necessary condition for monetary policy to be stabilizing), the net rate of return varies inversely with the rate of capacity utilization. As such, the fall in $u$ brought about by the decline in $a^K$ will now stimulate (rather than retard) the rate of change of $f_0$ in equation (6.25′) – which development is *destabilizing*. In short, the *interaction* of the two (individually stabilizing) mechanisms in (6.25′) and (6.31) sets up a destabilizing self-reinforcing mechanism, whereby an increase in $f_0$ causes a decrease in $a^K$, which in turn causes an increase in $f_0$. Franke (2019) demonstrates that the resulting system is saddle-path unstable. In other words, the destabilizing interaction effect will, in general, dominate the otherwise stabilizing tendencies of each mechanism considered individually.

Franke's (2019) result is timely at this stage of the discussion. Between them, classical-Marxian and neo-Kaleckian authors have conjured a plethora of mechanisms that appear capable of taming Harrodian instability. This may create the impression that some combination of these mechanisms must surely be active enough to thwart divergent dynamics, and so maintain the stability of the steady-state growth rate. But the work of Franke (2019) serves warning that Harrodian instability should not be written off – it may reassert itself in the most surprising fashion!

### 6.4.7 Empirical controversies

Neo-Harrodian models share with neo-Kaleckian models a commitment to the fundamentally post-Keynesian principle that investment by firms and saving by households are separate activities such that saving does not automatically create investment. Instead, investment – undertaken by firms independently of the saving behaviour of households – generates aggregate demand and so spurs output. This generates additional income from which arise changes in the quantity saved. One of the crucial differences between these models, however, is the *relative responsiveness* of investment and saving to disequilibrium conditions once (for whatever reason) the economy has been dislodged from a steady-state growth equilibrium consistent with Harrod's warranted rate. This issue is at the heart of the various 'stability controversies' (concerning Keynesian stability and Harrodian instability) discussed above.

To see this more clearly, consider again Figures 6.5 and 6.6, depicting neo-Kaleckian solutions to the problem of Harrodian instability. In either case, following some initial disturbance, the economy moves towards a new steady-state equilibrium consistent (for one reason or another) with the normal rate of capacity utilization, so that the new equilibrium is a fully adjusted position and there is no onset of Harrodian instability. This result requires the responsiveness of investment to disequilibrium conditions to be smaller than the responsiveness of saving. More specifically, as the actual capacity utilization rate rises following the initial increase in $f_0$ (which causes the upward displacement of the investment curve from $g$ to $g'$ in both Figures 6.5 and 6.6), the induced response of saving

$$\frac{d\sigma}{du} = \frac{s_r \pi}{a_1}$$

is greater than the induced response in investment spending. Because of the reconciliation of the now-higher rate of capacity utilization with the normal rate (due either to hysteresis effects or the tolerance interval around the normal rate of size $2\bar{c}$), and hence the absence of Harrodian instability, which would otherwise involve further changes in investment spending induced by the mechanism in equation (6.23), this induced response in investment spending reduces to:[35]

$$\frac{dg}{du} = \frac{f_1 \pi}{a_1}$$

According to the neo-Harrodians, even if we overlook questions regarding the mechanisms used to suppress the onset of Harrodian instability, the relative responses of investment and saving just described are, in and of themselves, empirically implausible. To demonstrate this point, Skott (2010, p. 111) performs a simple comparative static exercise using plausible numerical values for parameters such as $s_r$, $\pi$, $a_1$ and $f_1$ in a model in which the Keynesian stability condition is assumed to hold.[36] He then considers the effects on growth and the rate of capacity utilization of a small initial change in the saving rate – an exercise analogous to the increase in $s_r$ that produces the clockwise rotation in the saving curve in Figure 6.4 to $\sigma'$, followed (at least initially) by an increase in the rate of growth (from $g^*$ to $g'$) and an accompanying increase in the rate of capacity utilization (from $u^*$ to $u'$). Skott calculates that the change in the capacity utilization rate is approximately ten times larger than the change in the rate of accumulation – so that, for example, an increase in the growth rate from 2 to 4 per cent will be accompanied by an increase in the capacity utilization rate of about 20 percentage points. He then argues that this relative order of magnitude is inconsistent with the data, since economies such as the US display only *modest* long-run fluctuations in the rate capacity utilization accompanying fluctuations in the long-run rate of growth.

Neo-Kaleckian authors have furnished various responses to this line of criticism. Lavoie (2010, pp. 136–7) notes that consideration of other sources of saving would alter Skott's calculations. These might include saving by workers, for example, from which the neo-Kaleckian model typically abstracts for simplicity (see also Lima, 2010). Saving behaviour might also be modified by making it sensitive to financial variables (Franke, 2017), or the inclusion of a rentier class on which capitalists depend to finance investment (Hein, 2014, Chapter 9; Lima and Meirelles, 2007). Alternatively, Franke (2015) proposes a model in which a fiscally active public sector levies proportional taxes on corporate and personal income. This creates additional withdrawals from the circular flow of income, thus increasing the likelihood that the Keynesian stability condition will hold. Finally, Lima and Setterfield (2016) argue that modifying the neo-Kaleckian investment function so that it embodies a more fully developed post-Keynesian theory of how expectations are formed provides a better behavioural foundation for the relative unresponsiveness of investment to changes in capacity utilization. Collectively, these arguments suggest that elaborating the saving and/or investment relations typical of neo-Kaleckian models provides better foundations for the Keynesian stability condition while (potentially or demonstrably) altering the relative orders of magnitude of the changes in capacity utilization and growth that Skott calculates. In effect, these responses suggest that Skott applied an overly simplified model to the data, whereas a more complete model would not generate such extreme results.

An altogether different response involves *relaxing* the Keynesian stability condition altogether – and hence reversing the relative responsiveness of investment and saving to disequilibrium conditions typical of the neo-Robinsonian and neo-Kaleckian models in general – while showing that the stability (and key comparative static results) of the neo-Kaleckian model can nevertheless be retained. This is achieved by entertaining a second channel of adjustment in response to disequilibria between investment and saving of the sort created initially by changes in the investment or saving curves. As originally demonstrated by Bruno (1999) and Bhaduri (2006, 2008), if the quantity adjustments (reflected in changes in capacity utilization) implicit in the Keynesian stability mechanism operate simultaneously with a price adjustment mechanism (similar to that found in various of the neo-Keynesian models discussed in Chapter 3), the neo-Kaleckian model may be stable even if the Keynesian stability condition is violated.[37] Note that this extension involves relaxing the exogenously given markup assumption implicit in the constant profit share assumed in the basic Kalecki–Steindl model.[38]

Another empirical controversy germane to the discussion in this section concerns the relationship between the actual and normal rates of capacity utilization. As we have seen, conditions ripe for the onset of Harrodian instability can be expressed in terms of a discrepancy between the actual and normal rates of capacity utilization. But how do firms react to such discrepancies: by adjusting their investment spending in an effort to restore the actual rate of capacity utilization to its normal rate (resulting in the onset of Harrodian instability, as depicted in Figure 6.2); or by revising their sense of what constitutes a normal rate of capacity utilization in response to the 'new normal' suggested by actual experience (which will eliminate Harrodian instability, as depicted in Figure 6.6)?

This question is addressed in an as yet small empirical literature that, motivated by the hysteresis mechanism found in equation (6.27), has utilized a variety of empirical techniques in an effort to test whether or not the normal rate of utilization can be regarded as endogenous to the actual rate (potentially thwarting the onset of Harrodian instability). The approach in Lavoie et al. (2004) involves estimating and assessing the empirical performance of four different investment functions using Canadian data for 1960–98. Two of their investment functions are classical-Marxian in character, while two are neo-Kaleckian – one of which incorporates the sort of mechanism found in equation (6.27). They find that the latter function outperforms all others, and interpret this finding as evidence favourable to the neo-Kaleckian treatment of the relationship between investment and the rate of capacity utilization. This interpretation is disputed by Skott (2012), however, who argues that the estimating equation in Lavoie et al. (2004) does not, in fact, capture the hysteresis effect in (6.27), but instead represents a quintessentially *Harrodian* investment function in which the normal rate of capacity utilization is time-varying (but not endogenous to the actual rate). Skott also re-estimates Lavoie et al.'s (2004) preferred investment function using a Koyck transformation to eliminate the need to proxy unobservable variables (in this case, the normal rate of capacity utilization). He shows that the coefficients obtained from this estimation do not satisfy the Keynesian stability condition.

Skott's (2012) results are, in turn, disputed by Schoder (2012), who shows that Skott's parameter estimates may be biased due to serial correlation in the residuals. Schoder also provides evidence of his own, using an approach that explicitly entertains as its null hypothesis the sort of hysteresis mechanism found in equation (6.27). He postulates that the dynamics of the normal rate of capacity utilization follow a random walk augmented by an adjustment term, where the latter is a function of the difference between the actual and normal rates of capacity utilization lagged one period. This relationship is estimated using a Kalman filter, a multivariate filter technique that is well suited to estimating unobservable variables and their dynamics.[39] Using US manufacturing sector data from 1984–2007, Schoder finds evidence of a positive and robust adjustment term, as would be predicted by a neo-Kaleckian model featuring hysteresis in the normal rate of capacity utilization.

Nikiforos (2016b) adopts a different approach again. Having first constructed an alternative measure of capacity utilization based on the average workweek of capital, he then uses an autoregressive distributive lag (ARDL) model to regress changes in the trend value of the average workweek of capital, estimated using both an Hodrick–Prescott (HP) filter and a locally weighted least squares (LOWESS) method, and interpreted as the normal rate of capacity utilization, on its own lagged values, the Federal Reserve's measure of capacity utilization, and changes in the trend value of the Federal Reserve's measure of capacity utilization. The last is included to control for variations in the dependent variable that can be associated with the process of extracting a trend from data per se, which make trend values subject to a process similar to that described in (6.27) for purely statistical (rather than behavioural) reasons.[40] Regardless of the method used to extract the trend value of the average workweek of capital, the null hypothesis that the actual value of the Federal Reserve's measure of capacity utilization has no effect on the trend value of the average workweek of capital is rejected. Nikiforos argues that this provides evidence consistent with equation (6.27), suggesting that the normal rate of capacity utilization varies in response to deviations of the actual rate of capacity utilization from the normal rate.

Finally, Schoder (2014) reintroduces a note of caution. Schoder's project is not directly concerned with the mechanism in (6.27), but is instead designed to reflect on the possibility of reconciling the principle of effective demand with a stationary long-run rate of capacity utilization. Using a cointegrated vector autoregression (VAR) model applied to US data for 1960–2012, he shows that shocks to demand have persistent, long-term effects on output levels, and that full-capacity (potential) output also adjusts to actual output in the long run (while the reverse is not true). Schoder finds that these adjustments of capacity to output make the capacity utilization mean-reverting, so that while output levels – both actual and potential – do exhibit long-run effects of demand shocks, the utilization rate does *not* vary in the long run. These findings are consistent with Harrodian dynamics and, indeed, more generally with an enduring theme in classical-Marxian scholarship mentioned earlier in this chapter, that capitalism is stable in *proportions* (as represented by ratios such as the capacity utilization rate) but not in *dimension* (as represented by the absolute values of individual variables) (Duménil and Lévy, 1993).

## 6.4.8 A summing up

This section has highlighted the possible onset of Harrodian instability in models that are not intrinsically Harrodian by design, and the various mechanisms that classical-Marxian and neo-Kaleckian authors have postulated in an effort to counteract this destabilizing force in order to restore the stability of the steady-state equilibrium growth rate. The key difference between these mechanisms is whether or not they allow for long-run variation in the rate of capacity utilization. According to classical-Marxian economists, the capacity utilization rate does not vary between long-run (or long-period) equilibrium positions: any short-term departure of the actual rate of capacity utilization from its normal rate must eventually be resolved by convergence of the actual rate back towards its (fixed) normal rate. Neo-Kaleckians, meanwhile, argue that a fully adjusted position of this type is compatible with long-run variability of the capacity utilization rate – either because the normal rate is, itself, an endogenous variable, or because there is some basis for thinking that discrepancies between the actual and normal rates of capacity utilization are to be expected and will not provoke behavioural responses from firms (at least, within certain wide limits). This is significant because long-run variability of the capacity utilization rate can imbue the stable, steady-state growth rate with particular properties – for example, growth can become subject to the paradoxes of thrift and costs.

Ultimately, the debate between these classical-Marxian and neo-Kaleckian positions seems to rest on their fundamentally different conceptualizations of the long run. Recall from Chapter 1 that in the classical-Marxian view, a long-run position conforms to certain specific conditions (such as the realization of a specific normal rate of capacity utilization) that can be identified a priori, with the economy displaying only tendential gravitation towards such conditions (which may never actually be realized in practice). Post-Keynesians, meanwhile – including neo-Kaleckians – are more likely to view the 'long run' as the path-dependent product of a series of short- or medium-term episodes of growth, with long-run equilibrium positions (which may or may not be realized in practice) best conceived as the evolutionary result of processes that play out in historical time. As previously noted in Chapter 4, Kalecki himself argued that 'the long-run trend is but a slowly changing component of a chain of short-period situations; it has no independent entity' (Kalecki, 1971b, p. 165).

It is important to remember, however, that the debate between these positions is not the be-all and end-all of the debate about long-term dynamics in heterodox macroeconomics. What both the classical-Marxian and neo-Kaleckian positions have in common is their preference for conceiving the steady-state equilibrium growth rate as *stable*. For neo-Harrodians, this is fundamentally wrong-headed: following Harrod, the steady-state or warranted growth rate should be thought of as unstable, and this instability of the steady state should be embraced in the context of a 'corridor instability' view of macrodynamics, in which the instability of the warranted rate is ultimately local (as a result of independent ceilings and floors, or endogenous mechanisms that alter the direction of an otherwise self-reinforcing process of divergence), and the long-run growth process is thereby rendered fundamentally cyclical. As demonstrated earlier in this chapter, the idea here is to embrace instability as a fact of capitalist macrodynamics, not to look for additional dynamics designed to restore the stability of steady-state equilibrium. Neo-Harrodians then augment Harrod's original divergent disequilibrium dynamics with auxiliary mechanisms that bound it both from above and from below, so that the growth path of the economy oscillates within bounds. Quite apart from important differences in matters of theory, then, it is important to bear in mind these important differences in methodology and even ontology – or what Schumpeter called 'pre-analytic vision' (see Heilbroner and Milberg, 1996) – that inform debate within and between adherents of heterodox macroeconomics.

## 6.5 Conclusions

This chapter has focused on neo-Harrodian thinking in contemporary heterodox macroeconomics. Two main issues have been explored. The first concerns formalizations of the corridor instability view of macrodynamics that modify the divergent tendencies found in Harrod's work by adding mechanisms that create turning points and thus give rise to cyclical growth outcomes. Such contributions are found in the work of neo-Harrodian theorists such as Skott (1989, 2010), Ferri et al. (2011), Fazzari et al. (2013), and Ryoo and Skott (2017). The second concerns the potential occurrence and implications of Harrodian instability in heterodox growth models that are *not*, by their nature, Harrodian in inspiration.

According to neo-Harrodians, capitalist dynamics are characterized by local instability at any point in time. Per Harrod, departures from the equilibrium (warranted) rate of growth, consistent with the normal rate of capacity utilization, are self-reinforcing rather than self-correcting – and these tendencies are likely to assert themselves in *any* growth model, whether or not the model is (initially) Harrodian in inspiration. This claim constitutes a challenge to the stability of long-run equilibrium growth dynamics that is posited in models associated (variously) with the classical-Marxian, neo-Keynesian and neo-Kaleckian traditions. As demonstrated in this chapter, various authors associated with these traditions have sought to explicitly admit and then subsequently tame Harrodian instability through a variety of mechanisms, including (among others) policy response functions, tolerance for local (if not global) departures from the normal rate of capacity utilization, and hysteresis effects. The ambition of these contributions is to show that even if Harrodian instability arises in a model that will otherwise converge towards its steady-state growth outcome, there are other dynamics that, once taken into consideration, cancel out the effects of Harrodian instability and restore the stability of equilibrium.

Given their conviction that local instability is ubiquitous, neo-Harrodians regard this quest to restore the stability of equilibrium as fundamentally wrong-headed and at variance with empirical evidence. However, the empirical evidence related to the debate over Harrodian instability reviewed in this chapter is mixed and therefore inconclusive. But local instability does not mean that capitalism is explosive – a claim that, itself, would be at variance with even the most casual observation of the last two centuries (or more) of history. Instead, neo-Harrodians postulate that local instability is contained by a variety of mechanisms. Some of these – such as supply-side determinants of the potential output path, or the trajectory of any genuinely exogenous component of aggregate demand – place upper and lower bounds on the divergent trajectory of a capitalist economy. Others – such as reserve army mechanisms or policy reaction functions – set up self-correcting feedback mechanisms that, when they interact with Harrodian instability, can create self-perpetuating (or at least highly persistent) cycles. Either way, the 'true' Harrodian project is revealed as providing an account of fluctuations in growth dynamics that, given the essentially post-Keynesian nature of Harrod's dynamics, provides an alternative to the neo-Marxian Goodwin model discussed in Chapter 2, and differs also from the neo-Goodwinian account of business cycles discussed in Chapter 5, both of which are based on cyclical fluctuations in wages that in turn drive 'profit-squeeze' dynamics operating through the labour market and investment behaviour.

Whatever the merits of neo-Harrodian models, they make a valuable contribution to heterodox macrodynamics by drawing attention to an alternative *vision* of the capitalist growth process, as unsteady and subject to fluctuations. As we will see in the next chapter, moreover, we are not yet finished with the topic of cyclical growth. Some recent contributions to heterodox macrodynamics have produced models capable of generating cyclical growth and, in the process, the Goodwin pattern of counterclockwise movements in wage share–capacity utilization (or employment) space, and that are neither Goodwinian nor neo-Goodwinian (in the sense of Chapter 5), nor even pseudo-Goodwinian in the sense of the Skott (1989, 2010) limit cycle model discussed in this chapter. It is to these and other recent advances and developments that we now turn.

## ? STUDY QUESTIONS

1) What might create ceilings and floors capable of containing the instability associated with the second Harrod problem?
2) How does the profit-squeeze mechanism contribute to neo-Harrodian growth dynamics?
3) Outline two processes capable of taming Harrodian instability – one classical-Marxian and one neo-Kaleckian. What significance, if any, attaches to the difference between these mechanisms?
4) Is it possible to reconcile long-run variation in the utilization rate with the concept of a normal rate of capacity utilization?
5) Can monetary policy contain Harrodian instability and, if so, is the result necessarily a stable, steady-state equilibrium (warranted) growth rate?
6) Outline the debate between neo-Kaleckians and neo-Harrodians over the responsiveness of investment and saving behaviour to variations in the rate of profit/utilization.

## NOTES

**1** See Harrod (1936). Harrod (1939) is regarded as the classic statement of his dynamic theory, a theory that he continued to work on and refine throughout his life (for example, Harrod, 1948, 1973).

**2** Adjustment may also occur through changes in inventories of finished and unfinished goods, but we focus here on changes in the rate of capacity utilization consistent with the earlier exposition of Harrod's dynamics in Chapter 3.

**3** These difficulties are, of course, compounded by local bottlenecks occurring at sectoral or even industry level, which can set off inflationary pressures in some sectors (that show up in aggregate statistics) even as excess capacity persists elsewhere in the economy.

**4** The notion that there exists an autonomous component of aggregate demand that plays a decisive role in the determination of aggregate growth outcomes is also an important feature of contemporary super-multiplier analysis, a topic that is taken up in Chapter 7. Hicks is also considered the 'grandfather' of this literature – it was, in fact, Hicks (1950) who first coined the term 'supermultiplier'.

**5** According to Minsky (1975, p. 125), 'Stability – even of an expansion – is destabilizing in that the more adventuresome financing of investment pays off to the leaders and others follow'.

**6** Note that in order to approach its potential output path from below, the economy must be growing at a rate that exceeds the natural rate of growth – hence the claim above that $y^a > y_N$ until the ceiling is reached.

**7** Fazzari et al. (2013, p. 13) show that the level of output at the lower turning point of the cycle is a multiple of the exogenously given component of demand, $A$, and that for reasonable parameter values this lower turning point is much higher than the value of $A$ itself.

**8** Indeed, Skott is sceptical as to whether there *are* components of aggregate demand that are genuinely exogenous. We will return to this issue in Chapter 7, when we analyse contemporary supermultiplier models and their emphasis on the exogeneity of the growth rate.

**9** Other neo-Harrodians who have built on this approach include Ryoo (2013, 2016), who extends the long-run (steady-state) growth implications of Skott's framework (discussed below) by introducing financial fragility (Ryoo, 2013) and variations in the saving rate generated by changes in income inequality (Ryoo, 2016).

**10** Recall that in the neo-Keynesian models surveyed in Chapter 3, an increase in the profit share raises the rate of profit and hence the rate of *capital accumulation*. Because the capital to full-capacity output ratio $a_1$ is fixed, this is equivalent to an increase in the rate of growth of output.

**11** Since $a_1$ is a constant, its normalization to a value of one here has no effect on the subsequent analysis

– much like our treatment of the normal rate of capacity utilization in the analysis of classical models in Chapter 2, where we set $u_n = 1$.

12 Having said this, however, Skott (2010, p. 122) argues that 'the comparative statics of the steady-state growth solution will give a good approximation to changes in the average values' of the model's endogenous variables. This is reflected in analysis of neo-Harrodian inspiration that takes its lead from the steady-state implications of Skott's model (rather than its cyclical growth outcomes), such as the work of Ryoo (2013, 2016) cited earlier.

13 According to Skott (1989, p. 97), this requirement is empirically plausible. Note also that with $y_e < 0$, the simplified case mentioned earlier, where $Tr(\mathbf{J}) = 0$, is still possible.

14 The system is nonlinear, so $Tr(\mathbf{J}) > 0$ does not automatically eliminate the possibility of limit cycles.

15 The investment function in Fazzari et al. (2013) also includes a term that allows for a positive rate of depreciation of capital. In equation (6.2), the rate of depreciation is implicitly assumed to be zero.

16 Ferri et al. (2011, pp. 217–18) also elaborate on the adaptive expectations found in Fazzari et al. (2013). First, they posit two groups of agents, optimists and pessimists, whose expectations are quantitatively different. They then allow the populations of optimists and pessimists, and hence the average sentiment in the economy as a whole (which is a weighted average of the expectations of the two groups), to evolve endogenously in response to past forecasting success/failure.

17 The simple model of Harrodian instability so described is consistent with the discussion in section 3.2.4 of Chapter 3, where departure from the warranted rate of growth and the onset of Harrodian instability was shown to entail departure from the normal rate of capacity utilization. The mechanism used by Ryoo and Skott (2017) to capture Harrodian instability is formally modelled in section 6.4.3 below, as part of a more general discussion of Harrodian instability in heterodox macrodynamics.

18 See also the discussion of Franke (2019) in section 6.4.6 below. Franke shows that the interaction of a stabilizing monetary policy rule and a stabilizing rule for the growth of autonomous consumption spending by households may be destabilizing.

19 This outcome is consistent with the classical-Marxian and neo-Keynesian views that the capacity utilization rate is constant, at its normal rate, in long-run equilibrium.

20 The interested reader can check this claim by referring to the structure of the models used in Chapters 3 and 4, in tandem with that of the generic model developed above, and noting (in each case) what the precise form of the Keynesian stability condition implies in terms of the way each model describes the response of saving and investment behaviour to variations in the profit rate.

21 See also Chapter 8 for discussion of a similar view in the context of Kaldorian growth theory.

22 The reader will recall similar complications associated with interpreting the effects of parametric change in the Harrod model developed in Chapter 3.

23 Note that this reasoning does also assume that the speed of convergence towards the stable steady-state equilibrium is sufficiently swift. A slow speed of adjustment relative to the frequency with which exogenously given parameters are, themselves, inclined to undergo change means that a system will never be 'in equilibrium', as a result of which its equilibrium outcomes will never serve as an accurate description of the actual outcomes observed in the system (Harcourt, 1981 [1982], p. 218; Fisher, 1983, p. 3; Cornwall, 1991, p. 107; Halevi and Kriesler, 1992, p. 229).

24 See also Lavoie (2003) and Lavoie and Kriesler (2007) for criticism of the Duménil and Lévy (1999) model as reminiscent of neoclassical New Consensus macroeconomics, in which monetary policy conducted by an inflation-averse central bank plays the role of macroeconomic servo-mechanism, ensuring the stability of (in that model) the supply-determined natural rate of unemployment.

25 Note that it is not possible for *all* features of the growth process associated with neo-Kaleckian models to emerge from the generic Kalecki–Robinson model utilized here, because the precise form of the investment function used in our model does not admit the possibility of wage-led growth. The point to be made, however, is that as we will see, the introduction of neo-Kaleckian taming mechanisms into our generic Kalecki–Robinson model rehabilitates the possibility that the rate of capacity utilization is variable in the long run – a necessary condition for some of the results associated with neo-Kaleckian macrodynamics.

26 Setterfield (2018) postulates that $\bar{c}$ is a function of the standard deviation of past observations of the actual rate of capacity utilization.

27 Equation (6.27) is continuous whereas the account of hysteresis effects just given is suggestive of discontinuities in behaviour. Indeed, discontinuities are an important feature of formal models of 'strong' or 'true'

hysteresis – but continuous functions such as (6.27) can sometimes be considered appropriate simplified first approximations of such processes. See Setterfield (2009).

**28** Chapter 4 (section 4.2) discusses two other reasons why firms have desired excess capacity: to serve as an entry barrier in oligopolistic markets and because of indivisibilities in the efficient scale of plant and equipment.

**29** Recall from Chapter 5 (section 5.2.2) that a conflicting claims inflation model must be couched in terms of wage share targets in order to take labour productivity growth into account. An implicit simplifying assumption in the analysis here, then, is that the labour coefficient $a_0$ is constant or, in other words, there is no labour productivity growth.

**30** This is similar to Appendix 5.1, but expressed in terms of the real wage instead of the wage share. Note that in equations (6.28) and (6.29), only $w_w$ adjusts in response to variations in $u$, while the value of $w_f$ remains constant. In other words, the equilibrium condition $w_w = w_f$ is achieved entirely by the adjustment of $w_w$. What this draws to attention is that, by omitting the utilization effect previously included in equation (5.14) from (6.29), the Hein specification of the NAICU model used here guarantees that the equilibrium real wage will equal firms' target real wage – that is, workers have to accept the real wage that firms desire to pay them. This is a feature of basic Kaleckian macro models in which the markup (and hence, given the value of the nominal wage, the real wage) is established exclusively by firms, and also some conflicting claims inflation models that claim to characterize 'neoliberal' capitalism (Setterfield, 2006a, 2007).

**31** As Franke (2018, p. 601) notes, since steady-state conditions now require $u = u_n$ in (6.19) *and*, from (6.25′):

$$\frac{\pi u}{a_1} - i - i^* = 0$$

$$\Rightarrow u = \frac{a_1(i + i^*)}{\pi}$$

then unless, by chance, we observe:

$$u_n = \frac{a_1(i + i^*)}{\pi}$$

this model is consistent with the neo-Kaleckian claim that firms may have multiple, conflicting targets that cannot be reconciled.

**32** Sraffian supermultiplier models based on the work of Serrano (1995) are discussed in their own right in Chapter 7.

**33** In contrast, the source of exogenous spending in Allain (2015) is government expenditure. To avoid raising issues associated with the accumulation of public debt, Allain (2015) assumes an exogenous income tax rate that adjusts so as to maintain a balanced budget, and hence avoid public debt accumulation altogether. The modification of the saving rate demonstrated below by the introduction of autonomous consumption spending is thus achieved in Allain (2015) by variations in the tax rate that are, in turn, driven by the size of $A/K$ now interpreted to denote the scale of autonomous government spending relative to the capital stock. The basic workings of Allain's model are otherwise similar to those of the model outlined below.

**34** The reader is reminded of the similar finding of Ryoo and Skott (2017), emanating from the interaction of two (individually stabilizing) policy rules, discussed earlier in section 6.3.5.

**35** The astute reader will recognize that, as stated above:

$$\frac{d\sigma}{du} > \frac{dg}{du} \Rightarrow s_r > f_1$$

This is simply a restatement of the Keynesian stability condition associated with our generic Kalecki–Robinson model that was identified earlier. Note, however, that as the preceding discussion demonstrates, the controversy between neo-Harrodians and neo-Kaleckians concerning the responsiveness of investment and saving to disequilibrium conditions concerns not just the Keynesian stability condition per se, but also the question as to whether or not departure from the normal rate of capacity utilization will induce the investment response by firms that is the proximate driver of *Harrodian* instability.

**36** See also Dallery (2007) for a more extensive set of exercises that results in similar scepticism regarding the empirical plausibility of the neo-Kaleckian model.

**37** See Lavoie (2010, pp. 137–40) for a summary discussion of this dual adjustment process and its consequences.

**38** As such, a model of this sort is of a piece with those discussed in section 5.3.2 of Chapter 5, where both the rate of capacity utilization *and* the profit (or wage) share are endogenous and vary simultaneously in response to conditions of disequilibrium. Models of this type are important because they go some way to addressing another common empirical criticism of neo-Kaleckian models – namely, that they have nothing to say about profit share dynamics, despite the observation of variations in the profit share in the data (Skott and Zipperer, 2012). See Ohno (2014) for a neo-Kaleckian model in which the price adjustment mechanism depends on firm entry and exit, and the resulting influence of industrial concentration on the size of the markup.

**39** The estimating model is completed by a neo-Kaleckian investment equation, an equation describing serial correlation in the error term of the investment equation, and an expression for the expected rate of sales growth.

**40** See also Skott (2012), who makes the same point.

# Appendix 6.1 'Taming' Harrodian instability

In general we can write:

$$\frac{du}{dt} = \frac{\partial u}{\partial s_r}\frac{ds_r}{dt} + \frac{\partial u}{\partial f_1}\frac{df_1}{dt}$$

Drawing on equations (6.23) and (6.24) and the equilibrium solution of equations (6.21) and (6.22), this total derivative can be evaluated as:

$$\frac{du}{dt} = -\frac{a_1 f_0 \pi}{(s_r - f_1)\pi}\alpha_2(u - u_n) + \frac{a_1 f_0 \pi}{(s_r - f_1)\pi}\alpha_1(u - u_n)$$

$$\Rightarrow \dot{u} = (\alpha_1 - \alpha_2)\frac{a_1 f_0 \pi}{(s_r - f_1)\pi}(u - u_n)$$

Clearly, equilibrium ($\dot{u} = 0$) requires $u = u_n$. Note also that if $\alpha_2 > \alpha_1$:

$$\frac{d\dot{u}}{d(u - u_n)} = (\alpha_1 - \alpha_2)\frac{a_1 f_0 \pi}{(s_r - f_1)\pi} < 0$$

In other words, an increase in $u$, and hence an increase in $u - u_n$ above its equilibrium value of zero, will cause a reduction in $\dot{u}$ below its corresponding equilibrium value of zero, so that $\dot{u} < 0$. This means that $u$ will be falling back towards $u_n$ whenever it is above its normal rate. By the same token, a decrease in $u$, and hence a decrease in $u - u_n$ below its equilibrium value of zero, will cause an increase in $\dot{u}$ above its corresponding equilibrium value of zero, so that $\dot{u} > 0$. This means that $u$ will be rising back towards $u_n$ whenever it is below its normal rate. In short, in the extended generic Kalecki–Robinson model that includes both equation (6.23) and equation (6.24), $\alpha_2 > \alpha_1$ is sufficient to ensure that the system is once again self-equilibrating. Any manifestation of Harrodian instability will be strictly temporary.

# 7

# New directions: wage inequality, rentier income, financial dynamics and supermultiplier models

## 7.1 Introduction

This chapter investigates several new directions in heterodox macro modelling, with two principal themes. First, we will examine models that treat different dimensions of income distribution besides the simple wage–profit divide. Many, although not all, of these models can be classified as 'three-class' models, because they incorporate some kind of (upper) middle class that lies in-between production workers and capitalists or firms in the social hierarchy. This middle class may take the form of highly paid workers (managers, professionals and/or executives) who receive higher 'wages' than ordinary workers, or else a 'rentier' class that lives off income on financial assets (bonds or equity issued by the firms). In addition, we will consider gender wage gaps, which are an example of a situation in which one branch of the labour force has a privileged position over another (other examples would include wage gaps based on race, ethnicity or national origin). Second, we will cover models that feature 'drivers' of long-term growth and short-term cycles other than the distributional factors emphasized in earlier chapters. One type of alternative driver of cyclical fluctuations is financial dynamics, which may entail the accumulation of debt on the part of either firms or households. Models featuring financial dynamics have become increasingly common in response to the process of financialization – the increasing dominance of financial interests over industrial interests in contemporary capitalist economies – first identified in Chapter 1. Another possibility is that long-run growth may be driven by exogenous components of aggregate demand (autonomous consumption, government expenditures or exports), as contemplated in so-called supermultiplier models. This chapter will explore models in which long-run growth adjusts to an exogenous component of domestic demand; models emphasizing exogenously growing external demand (exports) will be covered in Chapters 8–10.

In spite of this multiplicity of themes, there are several threads that tie this chapter together. One is that most of the models covered here highlight different ways in which income distribution (along multiple dimensions) can interact – both as a cause *and* as an effect – with aggregate demand, cyclical instability and long-term growth. Indeed, financial models by their nature involve income distribution beyond the simple functional distribution between wages and profits, while supermultiplier models highlight alternative ways in which distribution and demand can affect long-run outcomes (for example, by influencing the level but not the growth rate of output). Another commonality is that, in many of the models presented here, the distinction between wage-led and profit-led demand regimes – already covered extensively in Chapters 4–6 – becomes endogenous and is shown to depend on various structural aspects and parameters of an economy and its evolution. Chapter 4 already revealed several factors that determine whether an economy has wage-led or profit-led demand, such as the propensity to save out of wages and the degree of openness to international trade, but this chapter will introduce others including wage inequality among different classes of workers and the amount of financial payouts (dividends or interest) by firms to rentier households. Finally, some of the models presented in this chapter will show ways in which an economy can appear to exhibit Goodwin-type cycles, even though the underlying causality and dynamics are completely different from those in the Goodwin models (old and new) discussed in Chapters 2 and 5, and do not reflect either profit-led demand or a profit-squeeze in distribution.

All of the topics covered in this chapter are major focuses of current research; this chapter will only seek to lay some foundations for studying these issues and will only cover a few examples of each type of analysis. Apart from the usual differences in notation, where ours is intended to be more standardized across the various models presented, we deliberately simplify some of the models for expositional purposes and in some cases we present only the comparative statics of the models and not their dynamic properties. Also, we try to emphasize intuition over complete mathematics where the latter might be too tedious (but full mathematics are given for some models). This has been done in part to facilitate comparisons among the different models, and in part simply because of space constraints. One particular omission here is that we do not give a complete 'stock-flow-consistent' analysis of the income–expenditure and balance sheet relations among the variables, but fortunately other sources are available to teach those methods.[1] Thus, the discussion in this chapter in many respects reflects only the tips of several very large icebergs, but our hope is that this discussion will spark the reader's interest in pursuing more of the details and alternatives in the original studies in these unfolding literatures.

The rest of the chapter is organized as follows. Section 7.2 presents short-run models of wage inequality in models largely of neo-Kaleckian inspiration (but with some variations). Section 7.3 discusses short-run models incorporating the financial income of rentiers, such as interest on bonds and dividends on equity. In several of the models in both sections 7.2 and 7.3, we will see that the degree to which demand is wage-led or profit-led varies depending on the distribution of income between the middle class (managers or rentiers) and the top and bottom strata (corporations or capitalists versus production workers). Section 7.4 presents two models of debt dynamics, one focused on borrowing by firms that creates Minskyan financial fragility and the other on borrowing by worker households for consumption purposes. In both cases, the models generate cyclical dynamics that can resemble neo-Goodwin cycles but are driven by entirely different causal mechanisms. Section 7.5 then discusses supermultiplier models in which long-run growth is driven by an exogenous component of domestic demand. Section 7.6 concludes.

## 7.2 Models of wage inequality

All the models of income distribution so far in this book have emphasized the functional distribution of income between wages and profits, broadly defined as the total compensation of labour and total returns to capital, respectively. However, there are many other dimensions of income distribution in any given society, one of which is inequality between different strata of the labour force. In micro-level studies, workers can of course be divided in numerous ways (and in great detail) according to their age, gender, race or ethnicity, immigration status, industries, occupations, and even the particular firms that employ them or subnational regions in which these firms are located. But for macro modelling purposes, workers can be divided broadly into production workers, who actually produce the output, managers and supervisors, who oversee the production process and perform administrative tasks, and top corporate executives, who may identify more with the stockholders who own the firm's capital than with other employees. This last point is most important in a system of corporate governance marked by 'shareholder value orientation', in which the top management is obliged to maximize the returns to shareholders rather than (as in older models of corporate governance) focus on the long-term growth of the firm. Under many contemporary compensation schemes, a large part of the top executives' remuneration is linked directly or indirectly to the firm's current profits, and hence these executives have incentives to behave more like capitalists (firm owners) than like workers, managers or other employees.

Even if we omit top executives from the labour force and count them simply as claimants on firms' profits (along with the stockholders or owners of equity in the firms), there are still more dimensions of inequality in compensation among other types of labour. In general, managers earn more than the production workers they supervise. More highly educated and 'skilled' labour, such as professional and technical workers, generally earn more than less educated or less 'skilled' labour, such as assembly line workers or others who perform repetitive, manual tasks. Some workers may suffer discrimination because of their gender, race, ethnicity, age or national origin (for example, immigrants or guest workers), in which case they are paid wages below those of comparable workers with similar education and skills performing similar productive functions. Such discrimination involves power relations such as patriarchal domination in the gender arena or the dominance of particular racial, ethnic and national groups. On the other hand, unionized workers may be able to win higher wages and greater benefits than comparable non-union workers, regardless of their other characteristics, and government policies can have a strong impact on the ability of workers to form independent unions as well as the non-wage social benefits that workers receive.

While space precludes full coverage of all these possible distinctions, this section will briefly cover three examples from the literature of models that try to take some of these dimensions of wage inequality into account. To facilitate the exposition and comparison of these models, as well as for the sake of brevity, they will be presented only in comparative static form, omitting the long-term dynamics also considered by the original authors. All of these models are situated largely in the neo-Kaleckian tradition, covered in Chapter 4, but some of them break out of the standard neo-Kaleckian mould in one way or another.

### 7.2.1 A two-class model with capitalist-managers

One way of introducing wage inequality into a neo-Kaleckian model is the approach of Palley (2017). Palley assumes an economy in which the labour force is divided into two types of households: working-class and middle-class production workers ($L$) and wealthy 'capitalist-managers' ($K$). In effect, Palley keeps the top executives in the labour force in his model and groups them together with highly paid managers and wealthy shareholders in the firms. Since the capitalist-managers provide managerial labour in Palley's model, they receive a combination of wage (labour) and profit (capital) income, but of course they receive much higher remuneration for their labour than they pay to the workers they employ.

The main driving factor in Palley's model is the difference in the saving behaviour of the worker and capitalist-manager households, where the latter are assumed to have a higher propensity to save: $s_K > s_L$. More like in Pasinetti's (1962) model, covered in Chapter 3, than in the other models covered in Chapters 4–6, these saving propensities apply equally to all types of income received by each type of household, and do not depend on the functional source of the income (capital or labour). Since workers do save, they must own part of the society's capital stock (for example, via pension funds) and hence they also receive a portion of the total profits, but their marginal propensity to save $s_L$ is the same for their profit income as it is for their wage income. Similarly, capitalist-managers receive compensation for their labour, and they apply the same marginal propensity to save $s_K$ to their labour income (wages) as to their capital income (profits). Implicitly, all profits are paid out to the two types of households combined; Palley does not consider profits retained by corporations.

The results of Palley's model then hinge on the shares of the worker and capitalist-manager households in each type of income, wages and profits. For wages, the shares are $\varphi_L$ and $\varphi_K$, where $\varphi_L > \varphi_K$ (so worker households receive relatively more of the total wage income) and of course $\varphi_L + \varphi_K = 1$. Similarly, for profits, the shares are $\delta_L$ and $\delta_K$, where $\delta_L < \delta_K$ (so capitalist-manager households receive relatively more of the profits), and $\delta_L + \delta_K = 1$.[2] For the sake of realism, Palley also assumes that worker households have a higher share of labour income than of profit income ($\varphi_L > \delta_L$), and the converse is true for capitalist-manager households ($\varphi_K < \delta_K$).

Based on these assumptions, the saving functions for the two classes (normalized by the capital stock) are

$$\sigma_L = s_L[\varphi_L(1 - \pi) + \delta_L\pi]u/a_1 \tag{7.1}$$

$$\sigma_K = s_K[(1 - \varphi_L)(1 - \pi) + (1 - \delta_L)\pi]u/a_1 \tag{7.2}$$

where (as in previous chapters) $u$ is the utilization rate (the ratio of actual to full-capacity output), $a_1$ is the ratio of capital to full-capacity output, and $\pi$ is the profit share of national income ($1 - \pi$ is the wage share). Thus, each class of households applies its marginal propensity to save to its share of the total wage and profit income. The profit share is taken as exogenously given, presumably because of an underlying fixed markup rate (and this assumption is crucial to the results, as we shall see).

For an investment function (also normalized by the capital stock), we will use a linearized version of the Bhaduri–Marglin investment function (equation 4.37):

$$g = h_0 + h_1\pi + h_2(u/a_1) \tag{7.3}$$

where $h_1, h_2 > 0$ (the sign of $h_0$ will be discussed below). This is a simplified version of what Palley (2017) uses, but the simplification does not affect the key results we will show here so we prefer the simpler version.[3] Assuming no government or foreign trade for simplicity, the equilibrium condition (saving equals investment) is

$$\sigma_L + \sigma_K = g \tag{7.4}$$

Substituting (7.1)–(7.3) into (7.4), we can obtain an explicit solution for the equilibrium utilization rate

$$u^* = \frac{h_0 + h_1\pi}{\Sigma} \tag{7.5}$$

where $\Sigma = \{s_L[\varphi_L(1-\pi) + \delta_L\pi] + s_K[(1-\varphi_L)(1-\pi) + (1-\delta_L)\pi] - h_2\}(1/a_1) > 0$ assuming the Keynesian stability condition holds, in which case $h_0 > -h_1\pi$ is required for an economically meaningful (positive) solution.

Several important results emerge from this solution. First, *regardless* of whether demand (utilization) is wage-led or profit-led overall (which we will analyse below), a redistribution of either type of income towards production workers is always expansionary in Palley's model. To see this, note that the effect of a rise in the workers' share of wage income (holding the profit share $\pi$ constant) is

$$\frac{\partial u^*}{\partial \varphi_L} = \frac{(s_K - s_L)(h_0 + h_1\pi)(1 - \pi)(1/a_1)}{\Sigma^2} > 0 \tag{7.6}$$

since $s_K > s_L$. This positive effect thus stems from the fact that worker households have a higher marginal propensity to consume (lower propensity to save) than capitalist-manager households, so higher wages for the former relative to the latter will increase consumer demand and, via the multiplier–accelerator interaction, investment spending as well (for any given profit share). Similarly (and for the same reason), an increase in the workers' share of profit income is also expansionary

$$\frac{\partial u^*}{\partial \delta_L} = \frac{(s_K - s_L)(h_0 + h_1\pi)(\pi/a_1)}{\Sigma^2} > 0 \tag{7.7}$$

Thus, a more equitable distribution of either labour income (total wages) or capital ownership (and total profits) increases aggregate demand and employment, for any given profit share, and regardless of whether the latter variable has a positive or negative impact on demand. However, these results depend critically on the strong assumption that changes in relative shares of workers and capitalist-managers in total wage and profit income would leave the overall profit share unchanged.

Second, as we would expect for any neo-Kaleckian model with positive workers' saving and a Bhaduri–Marglin investment function (as covered in Chapter 4), it is ambiguous whether demand is wage-led or profit-led in the Palley (2017) model. The relevant partial derivative of equation (7.5), holding the workers' and capitalist-managers' shares of each type of income constant, is (after much cancellation and rearrangement in the numerator)

$$\frac{\partial u^*}{\partial \pi} = \frac{\{h_1[s_K(1 - \varphi_L) + s_L\varphi_L] - h_0(s_K - s_L)(\varphi_L - \delta_L) - h_2h_1\}(1/a_1)}{\Sigma^2} \tag{7.8}$$

In this solution, the degree of wage- or profit-ledness is endogenous and depends on the workers' shares of the two kinds of income, $\varphi_L$ for wages and $\delta_L$ for profits. Since the denominator of equation (7.8) is obviously positive, the sign of this derivative depends only on the sign of the numerator.[4] It can easily be seen that a higher worker share of wage income $\varphi_L$ necessarily reduces the numerator (assuming $s_K > s_L$), and hence makes it more likely that the economy has wage-led rather than profit-led demand. In contrast, a higher worker share of profit income $\delta_L$ definitely raises the numerator and makes it more likely that demand is profit-led instead of wage-led. These results make intuitive sense, because when wage income is redistributed to workers (who have a higher marginal propensity to consume, or lower marginal propensity to save) then a redistribution of overall income towards wages is more likely to be expansionary, while if profit income is redistributed to workers, then for the same reason a redistribution of overall income towards profits would tend to be more expansionary.

### 7.2.2 A three-class model with manager-supervisors

A different approach to modelling wage inequality is that of Tavani and Vasudevan (2014). Their model has three distinct social classes: capitalists, managers and production workers. The capitalists are the owners of the firm's capital (presumably including top executives who, as noted earlier, may be closer in outlook and behaviour to shareholders); they receive the profits, which are the residual income of the firm after the costs of both kinds of labour (managers and workers) are paid. Unlike in Palley's model, the profit share $\pi$ in the Tavani–Vasudevan model is not exogenously given. Although the model is demand-driven and has many neo-Kaleckian features, profits are a residual (and $\pi$ is endogenous), more like in some of the classical-Marxian models covered in Chapter 2. Importantly, the middle-class managers play a key supervisory role in overseeing the production workers. The managers' labour is a cost to the firm that cuts into the capitalists' profits.

For reasons of space, and for ease of comparison with Palley's model, what we will present here is a simplified version of Tavani and Vasudevan's short-run model, omitting their analysis of medium-run dynamics, and translated into our own notation. National income in real terms is given by

$$Y = w_L L + w_M M + rK \tag{7.9}$$

where $L$ signifies hours of production workers' labour, $M$ represents hours of managerial labour, and $w_L$ and $w_M$ are the real wage rates of workers and managers, respectively. The three inputs are used with fixed coefficients: $a_0$ is workers' labour per unit of output, $a_1$ is capital per unit of full-capacity output and $a_2$ is managers' labour per unit of output. As in Chapter 4 (equation 4.17), the profit rate can be written as $r = \pi u/a_1$, but now the profit share is the residual left over after the two kinds of labour are paid their respective shares: $\pi = 1 - w_L a_0 - w_M a_2$.

To represent the managers' supervisory role, it is assumed that they are employed in a fixed proportion to workers, $1/\theta$, where $\theta = a_0/a_2 > 1$. A rise in $\theta$ can be interpreted as an increase in 'managerial efficiency' or the ease of 'surplus-extraction', since it means that fewer managers are required to get the same amount of production worker labour performed. Wage inequality is represented by the ratio $\omega = w_M/w_L$, and for the sake of realism we can assume $\omega > 1$. Using the definitions of $\theta$ and $\omega$, and defining the share of production workers as $\psi_L = w_L a_0$, we can express the profit share as

$$\pi = 1 - \psi_L\left(1 + \frac{\omega}{\theta}\right) \tag{7.10}$$

where higher costs of managerial labour (either a rise in $\omega$ or a fall in $\theta$) can be seen to reduce the profit share, for any given workers' share $\psi_L$.

Turning to saving and investment, Tavani and Vasudevan use the Kalecki–Steindl investment function (equation 4.19 in Chapter 4), which we reproduce here as

$$g = g_0 + g_1 r + g_2 u \tag{7.11}$$

where $g_0, g_1, g_2 > 0$ – although none of the qualitative results reported below would be affected if we used the linearized Bhaduri–Marglin investment function (7.3) adopted by Palley instead. Tavani and Vasudevan assume (for simplicity) that the capitalists save 100 per cent of their profits, production workers do not save and managers have the saving propensity $s_M$, where $0 < s_M < 1$. Using the definitions of $\psi_L$, $\pi$, $\omega$ and $\theta$ given above, we can express the workers' share as $\psi_L = (1 - \pi)\theta / (\theta + \omega)$, and the saving rate (saving relative to capital) can be written as

$$\sigma = [1 - \psi_L - (\omega/\theta)\psi_L(1 - s_M)](u/a_1) \tag{7.12}$$

Using those same definitions as well as the definition of $r$, the investment function (7.11) can be rewritten analogously as

$$g = g_0 + \left\{\frac{g_1}{a_1}\left[1 - \psi_L\left(1 + \frac{\omega}{\theta}\right)\right] + g_2\right\}u \tag{7.11'}$$

Substituting (7.12) and (7.11′) into the equilibrium condition $\sigma = g$, we can find the reduced form solution for the equilibrium utilization rate $u^*$ as

$$u^* = \frac{g_0}{\tilde{\Sigma}} \tag{7.13}$$

where $\tilde{\Sigma} = (1/a_1)\{1 - \psi_L[1 + (\omega/\theta)(1 - s_M)] - g_1[1 - \psi_L(1 + (\omega/\theta))]\} - g_2 > 0$ assuming the Keynesian stability condition.[5] Note that we cannot analyse the effect of a rise in the overall profit share $\pi$ in this model, because as noted earlier, $\pi$ is an endogenous variable (profits are a residual). Thus, the Tavani–Vasudevan model does not permit a characterization of aggregate demand as either wage-led or profit-led in general. Like Tavani and Vasudevan, however, we can solve for the comparative static effects of changes in two key distributional parameters: the wage inequality ratio $\omega$ and the workers' share of national income $\psi_L$.

The effect of increased wage inequality, holding all other factors constant, is found by totally differentiating the equilibrium condition with respect to $u$ and $\omega$, which yields (after some simplification)

$$\frac{\partial u^*}{\partial \omega} = \frac{\left(\dfrac{\psi_L u^*}{a_1 \theta}\right)(1 - s_M - g_1)}{\tilde{\Sigma}} \tag{7.14}$$

where again $\tilde{\Sigma} > 0$ assuming Keynesian stability. To interpret the numerator of (7.14), it helps to observe that $1 - s_M$ is the managers' marginal propensity to consume, while an increase in managerial labour costs reduces profits and hence has a negative effect on investment through the $g_1$ coefficient. Thus, increased wage inequality has a positive, expansionary effect if the managers' marginal propensity to consume exceeds the firms' marginal propensity to invest out of profits $(1 - s_M > g_1)$, and a negative, contractionary effect in the opposite case $(1 - s_M < g_1)$. These results depend crucially on the assumption that managers are always employed in a fixed proportion to workers, $1/\theta$, and the fact that the cost of employing managers reduces firms' profits and hence has a negative impact on investment.

The effect of a rise in the production workers' share $\psi_L$ is given by

$$\frac{\partial u^*}{\partial \psi_L} = \frac{\dfrac{1}{a_1}\left\{\left[1 + \dfrac{\omega}{\theta}(1 - s_M)\right] - g_1\left(1 + \dfrac{\omega}{\theta}\right)\right\}}{\tilde{\Sigma}} \tag{7.15}$$

where again the denominator is positive and the numerator is ambiguous in sign. The term in brackets $[\cdot]$ in the numerator is the positive effect of higher workers' wages on consumption (including induced extra consumption of managers, who are hired in a fixed proportion to production workers), while the term $-g_1(\cdot)$ is the negative effect on investment (which occurs because higher labour costs reduce profits). Thus, a redistribution of income towards production workers is expansionary if the positive effect on consumption dominates the negative effect on investment, and is contractionary in the opposite case.[6] Except for the inclusion of managerial consumption, this is very much like the conditions for demand to be wage-led or profit-led in any standard Bhaduri–Marglin model.

To understand the differences in the results of the Palley and Tavani–Vasudevan models, it is important to understand the differences in their specifications of the social class structure and the determination of income distribution. Palley's capitalist-managers are essentially top corporate managers and major stockholders, who have very high incomes, assets and saving propensities. Moreover, Palley's firms maintain constant profit markups and profit shares regardless of how the wage and profit income is parcelled out between the two broad classes of workers and (combined) manager-capitalists. By making these assumptions, Palley is able to consider the 'pure' effects of the relative distribution of income between workers' households and the wealthy households of capitalist-managers, and in this setting a redistribution towards the former is always expansionary.

In contrast, Tavani and Vasudevan depict a three-class structure in which the managers are simply intermediaries and supervisors, and neither they nor the workers receive any share of the firms' profits.[7] Moreover, instead of assuming a fixed markup rate and profit share, Tavani and Vasudevan assume that the wages of both workers and managers are costs to the firms, and increases in both types of labour costs diminish the profits of the firms and the income of the true capitalists who own them. The capitalists in this model derive all their income from their ownership of the firms' capital; they do not perform any labour and do not receive wages or salaries. Therefore, in our static version of Tavani and Vasudevan's model, increases in the wages of either workers or managers can have either expansionary or contractionary demand-side effects, depending on whether the resulting increase in consumption outweighs the reduction in investment. In effect, Tavani and Vasudevan's managers are really just middle-class supervisors, and except for having a higher (positive) saving propensity, their impact on profits is more similar to that of workers rather than capitalists.

### 7.2.3 The gender wage gap in an export-oriented, semi-industrialized economy

Another important form of wage inequality is gender wage gaps, based on patriarchal power that facilitates discrimination against women workers. That is, women workers are often paid less than men workers of similar qualifications and experience, and some firms exploit their ability to underpay women to lower their labour costs. Of course, wage gaps are only one of many areas in which gender relations can matter for macroeconomic and development outcomes (see surveys by Duflo, 2012; Benería et al., 2015; Seguino, 2019). Studies of gender relations have also highlighted the burdens of women's unpaid work in the household sector as an obstacle to their advancement in the market economy and an implicit subsidy to the hiring of male labour (for example, Floro, 1995). In addition, a new line of research emphasizes the importance of the 'care economy' more broadly, including the provision of care services through either social or market mechanisms as well as within the home (Braunstein et al., 2011). However, in keeping with the theme of wage inequality in this section, we will restrict our attention here to a model of the gender wage gap.

As an example, this subsection will present a model of one very specific context in which gender wage gaps can matter: an export-oriented, developing country in which, at least in the early stages of industrialization, the export industries (for example, textiles and apparel or electronics assembly) use predominantly low-paid female labour. The model highlights the role of occupational segregation (the concentration of women workers into certain industries or occupations) and discrimination in pay (women workers receiving lower wages than comparably educated and skilled men). Our presentation will follow the modelling approach of Blecker and Seguino (2002, 2007),[8] but for reasons of space we will confine ourselves here to an outline of their short-run model with a mostly graphical and intuitive presentation.

Blecker and Seguino use a two-sector model for an open economy: one industry produces home goods $H$ that are only sold domestically, which are used for both consumption and investment, while the other industry produces exportable goods $X$ that can either be exported or consumed domestically. To model occupational segregation in an analytically tractable way, Blecker and Seguino assume that only female workers ($f$) are employed in the exportables sector while only male workers ($m$) are employed in the home goods industry. This is clearly a very special case, but it is also an important one from a policy perspective given the frequently found concentration of women workers in low-wage, labour-intensive export production in many developing nations, and the model demonstrates techniques that can be used to portray other forms of wage inequality and occupational segmentation.

Prices in the two industries are set in neo-Kaleckian fashion by markups over the unit costs of labour and imported raw materials

$$P_H = (1 + \tau_H)(W_m a_H + EP_n n_H) \tag{7.16}$$

$$P_X = (1 + \tau_X)(W_f a_X + EP_n n_X) \tag{7.17}$$

where $\tau_H > 0$ and $\tau_X > 0$ are the markup rates in each sector; $W_m$ and $W_f$ represent male and female nominal wage rates, respectively (assuming $W_f < W_m$);[9] $a_i$ is the labour coefficient in sector $i$ ($i = H, X$); $n_i$ is the intermediate input coefficient in sector $i$ ($i = H, X$); $P_n$ is the world price of intermediate inputs in foreign currency; and $E$ is the nominal exchange rate (domestic currency price of foreign exchange). All the input–output coefficients and

$P_n$ are exogenously fixed. The wage rates and exchange rate are taken as exogenously given in the short run, although they adjust dynamically in the medium run (due to space constraints, we will only cover highlights of the short-run model here).

The Blecker–Seguino model then incorporates a key difference in the market structures of the industries in which the men and women workers are employed. The home sector markup $\tau_H$ is assumed to be rigid as a result of protectionist barriers, government subsidies and a highly concentrated oligopolistic structure.[10] In contrast, the export-sector markup $\tau_X$ is assumed to be flexible in response to international competitive pressures. Similar to how we modelled a flexible markup in Chapter 4, $\tau_X$ is an increasing function of the real exchange rate for export goods, $EP_f/P_X$, where $P_f$ is the price (in foreign currency) of foreign goods that compete with domestic exports in global markets:[11]

$$1 + \tau_X = \mu\left(\frac{EP_f}{P_X}\right)^{\eta} \tag{7.18}$$

As in Chapter 4, $\mu$ is a positive parameter reflecting the degree of oligopolistic power (in this case, in the exportables sector) and $\eta$ is a positive elasticity (here, the subscript $f$ means foreign for prices but female for wages).

Assuming that domestic exports and foreign goods are imperfect substitutes, there is a finite price-elasticity of demand for the country's exports in response to changes in the real exchange rate (the relative price of foreign competing goods). Imports consist of the imported intermediate goods for export production (the export industries are labour-intensive assembly operations) plus a fixed proportion of the capital goods used for investment. Investment depends on the profits obtained in the two industries, while consumption depends on the incomes of the four groups of income-earners (workers and capitalists in each industry) and their respective marginal propensities to consume and the proportions in which they buy the two goods (recall that the exported goods may be consumed at home as well as sold abroad).

Without giving all the mathematical details here, we can summarize the short-run analysis using the diagram in Figure 7.1. The downward-sloping AD (aggregate demand) curve depicts macroeconomic equilibrium (saving equals investment plus net exports) in terms of the trade-off in the composition of output between the two goods ($X$ and $H$) for any given level of aggregate demand (as determined by exogenous parameters such as the saving propensities and the constant terms in the investment and consump-

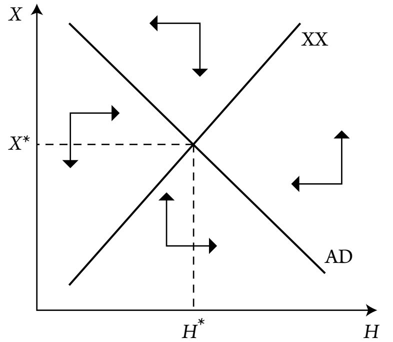



**Figure 7.1** Short-run equilibrium and dynamics in the Blecker–Seguino gender gap model

tion functions). If output of either good rises, this creates an excess aggregate supply of goods (excess of saving over investment), and to restore goods market equilibrium the output of the other good has to fall; hence the downward slope of AD. The XX curve represents market clearing (supply-equals-demand) equilibrium for the exported good $X$. XX slopes upward because, if output of (say) $X$ rises, this generates additional wage and profit income, part of which is spent on consumption of $H$, and to eliminate the resulting excess demand for $H$, more $H$ has to be produced (assuming that the $H$ sector does not hit a capacity constraint). There is excess supply in the market for $X$ above and to the left of the XX line, and excess demand below it; there is excess aggregate demand below and to the left of the AD line, and excess aggregate supply above it. Equilibrium occurs at the point $(H^*, X^*)$ where AD and XX intersect, so that the $X$ market clears while macroeconomic equilibrium is also fulfilled (by Walras' law, the $H$ market must also clear at the same point).[12] The short-run dynamics of the system are as shown in Figure 7.1, and, under standard assumptions, the equilibrium is locally stable in the neighbourhood of the equilibrium point.

In their short-run analysis, Blecker and Seguino (2002) consider the effect of a reduction in the gender gap, in the sense of an exogenous increase in the women's wage holding the men's wage constant.[13] A rise in $W_f$, holding all other factors constant, affects the equilibrium of the economy through three distinct channels: (1) a relative price effect, since $P_X/P_H$ increases; (2) a gender redistribution effect, since the women's real wage rises while the men's real wage falls as long as men consume some of the exported good; and (3) a functional or class redistribution effect, since the rise in the women's wage squeezes the profit markup and reduces the profit share in the $X$ sector. Similar to other models of neo-Kaleckian inspiration, the net effects ultimately boil down to whether the rise in consumption of women workers outweighs the negative effects of their higher wages on export competitiveness as well as on profits and investment in the export sector.

In an optimistic scenario, if working-class households have a much higher marginal propensity to consume than capitalist households (so that higher women's wages give a big boost to consumption), there is significant domestic demand for the exported good and exports are relatively price-inelastic, then AD and XX both shift to the right. As a result, the equilibrium output of *H* rises strongly while the change in *X* output is ambiguous, and the equilibrium levels of total employment and real national income would rise. In a pessimistic scenario, if exports are highly price-elastic, there is relatively little home consumption of the exported good and the marginal propensity to consume of workers is only slightly greater than that of capitalists, AD and XX both shift downward. In this case, the equilibrium output of *X* definitely falls, the change in *H* is ambiguous and equilibrium levels of total income and employment are reduced.[14]

This analysis thus illustrates the possible incentives for employers in export industries to discriminate against women workers by paying them lower wages than they would have to pay male workers. In the pessimistic scenario, a large gender gap can boost economic growth at the cost of impoverishing the women workers whose low wages make the exports competitive. This model also highlights the difficulties that could be faced in trying to close the gender gap in an economy of this type unless the country can escape from the trap of specializing in labour-intensive manufactured exports that sell in highly competitive global markets and is able to broaden the base for domestic consumption of exported as well as domestic goods.

Of course, the Blecker–Seguino model depicts only one, relatively special case of the impact of trying to close a gender wage gap. Nevertheless, it illustrates how gender differences and gender discrimination can be analysed using the tools of heterodox macro modelling covered in this book – tools that could easily be extended to cover gender wage gaps in other contexts or other aspects of gender relations. Moreover, the same kinds of analytical tools employed in this model of the gender wage gap could also be applied to other types of wage gaps such as those between workers of different races, ethnicities or national origins. However, the focus on gender relations results in one unique feature of this model: workers are segmented by gender in terms of their productive roles in the market economy (different industries, different wages), but they are still linked together in the same working-class households.[15]

In spite of the obvious differences in their model construction, the Tavani–Vasudevan analysis of raising the wage share of production workers and the Blecker–Seguino analysis of closing the gender gap actually have some important common features. In both models – at least in the simple, comparative static versions presented here – the positive effects of an inequality-reducing redistribution of income on workers' consumption have to be weighed against possible negative effects on profitability, investment and/or net exports. In contrast, the Palley model presents a more optimistic outlook, in which narrowing wage inequality is always expansionary, but this conclusion rests on the strong assumption that the overall profit share is held constant. Palley's analysis separates the 'pure' effects of wage inequality from the overall wage (or profit) share, but it leaves open the question of whether in practice such inequality can be reduced without the overall functional shares changing. In any case, Palley's model highlights the potential for income redistribution among different strata of workers to be beneficial to the overall economy under appropriate structural conditions.

## 7.3 Rentier income and distributional effects in financial models

In this section, we will consider a set of models in which corporations finance their investment externally, either by borrowing through the issuance of bonds or by issuing new equity shares to shareholders. In each case, the financial investors (bondholders or shareholders) are portrayed as a distinct class of 'rentiers', that is, wealthy individuals who live off the income on their financial assets and do not need to perform labour.[16] In the first case, the rentiers receive interest on their stocks of bonds; in the second case, they receive dividends on their shares of equity (for simplicity, capital gains are not considered). To be clear, we have a limited objective in this section, which is to elucidate the distributional consequences of changes in rentier income (either interest or dividends) and how these changes in turn impact capacity utilization and growth in the economy. We will therefore restrict our attention here to static versions of these models. For the sake of both simplicity and brevity, as well as to facilitate comparisons across model specifications, we will analyse short-run situations in which either the stock of bonds or the equity–capital ratio is exogenously given; we will not analyse the dynamics of debt accumulation or equity issuance over time in this section (even when those were covered in the original articles about these models). However, two dynamics models of debt finance will be presented later, in section 7.4.

### 7.3.1 A Kalecki–Steindl model with bondholders

The model in this section is based on several models found in Hein and Stockhammer (2011b), Hein (2014) and earlier work by Lavoie (1995a), among others, but modified and simplified for our purposes (and expressed in our notation). In this model, the rentiers are bondholders who lend money to firms by buying corporate bonds and receive interest payments on those bonds. All debt consists of stocks of outstanding corporate bonds, which are liabilities to the firms and assets to the rentier households who own them. All bonds are owned by the rentier class; workers do not save, hold bonds or receive any interest income. Rentier households live entirely off their interest earnings, which they partly spend on consumption and partly save by acquiring more bonds (although we will not analyse the dynamics of bond accumulation here). Firms are corporations that cannot consume, so they save 100 per cent of their net (retained) profits after paying interest on the debt (for simplicity, we ignore repayment of principal).

If we let $B$ equal the total outstanding stock of bonds in nominal terms, which can be taken as given or predetermined at any point in time, and $i$ is the interest rate, then interest payments on this debt $iB$ represent a cost to the borrowing firms and income to the rentier households (bondholders). Assuming that firms receive profits at the rate $r$ on the value of their capital $PK$, firms' (corporate) savings in nominal terms are equal to their net retained profits, $rPK - iB$, while rentier households have a marginal propensity to save $s_R$ $(0 < s_R < 1)$ for the interest income they receive (equal to $iB$). Since there are no savings out of wages, total nominal savings consist entirely of the net profits of the firms plus the personal savings of the rentiers

$$S = rPK - iB + s_R iB \tag{7.19}$$

Dividing both sides by $PK$ and rearranging, we obtain the following expression for the saving–capital ratio $\sigma$

$$\sigma = S/PK = r - (1 - s_R) i d_B \tag{7.20}$$

where $d_B = B/PK$ is the debt–capital ratio for corporate bonds and $(1 - s_R)$ is the rentiers' marginal propensity to consume.

Next, we use a Kalecki–Steindl investment function,[17] similar to equation (4.19) in Chapter 4 but subtracting the cost of debt service $(id_B)$ from the gross profit rate $(r)$ to make investment a function of net profits and the utilization rate:

$$g = g_0 + g_1(r - id_B) + g_2 u \tag{7.21}$$

where $g_1, g_2 > 0$ (the necessary assumption about $g_0$ to obtain an economically meaningful solution will be specified below). Assuming no government or foreign trade for simplicity, the saving-equals-investment equilibrium condition is of course $\sigma = g$, which upon substitution of (7.20) and (7.21) and using $r = \pi u/a_1$ yields the solution

$$u^* = \frac{g_0 + (1 - s_R - g_1)\,id_B}{(1 - g_1)(\pi/a_1) - g_2} \tag{7.22}$$

Following the same methodology as in Chapter 4, the denominator of this solution must be positive assuming the Keynesian stability condition.[18] Note that the stability condition also implies that we must assume $g_1 < 1$, because the denominator is positive if and only if $1 - g_1 > (a_1/\pi)g_2 > 0$. This is essentially the same as assuming that the propensity to invest out of firms' retained profits $(g_1)$ must be less than the propensity to save out of those profits, where the latter propensity is unity. Given a positive denominator, the numerator must also be positive for an economically meaningful (positive) solution, which requires that $g_0 > -(1 - s_R - g_1)\,id_B$.

Now, consider the impact of an increase in the debt service of firms $id_B$ (which, as noted earlier, equals the interest income of rentiers), measured as a ratio to the capital stock. Taking the profit share as exogenously given (we will consider the case where it is endogenous below), the effect is simply

$$\frac{\partial u^*}{\partial(id_B)} = \frac{1 - s_R - g_1}{(1 - g_1)(\pi/a_1) - g_2} \tag{7.23}$$

which is positive if $1 - s_R > g_1$ and negative if $1 - s_R < g_1$. In other words, a rise in interest costs (due to either a higher interest rate or an increased debt–capital ratio) is expansionary if the rentier's marginal propensity to consume exceeds the responsiveness of the firms' investment to their net profit rate, and is contractionary in the converse case.[19] Of course, most economic models, both mainstream and heterodox, assume the 'normal' case of a negative net effect – largely because most macro models (of any persuasion) tend to ignore rentiers' income and consumption, in which case only the negative result is possible. But in this model, there is also the possibility of what Lavoie (1995a) called the 'puzzling' case, in which higher interest costs have a positive net effect on output and utilization, which can occur if the response of rentiers' consumption is sufficiently large (or, to put it another way, if the rentiers' saving propensity is sufficiently low).[20] Interestingly, if we think of a rise in $id_B$ as a redistribution of income towards rentiers, the condition for this to be expansionary or contractionary is analogous to what we found for the case of a redistribution of wages towards managers in the Tavani–Vasudevan model (see the discussion of equation 7.14 in section 7.2.2, above).

Thus, the possibility of the puzzling case arises from an aspect of interest payments that is usually neglected in most macro models: the fact that they imply income for the owners of the assets (in this case, the rentiers who hold the corporate bonds). This analysis also makes it clear that monetary policy (which operates by setting the interest rate $i$[21]) has a distributional effect: holding the profit share $\pi$ and debt–capital ratio $d_B$ constant, a rise in $i$ redistributes income from corporations (which cannot consume, but do spend on investment) to rentier households (which consume with the propensity $0 < 1 - s_R < 1$). Hence, whether an interest rate hike is expansionary or contractionary depends on whether the induced increase in rentiers' consumption is greater or less than the induced decrease in firms' investment. Of course, this would depend very much on the social composition of the rentiers: if they are primarily retired workers, for example, living off the interest income on bonds held in their pension funds, then the puzzling case would be plausible; but if the rentiers are mainly wealthy individuals whose income and wealth far exceeds even their own lavish consumption expenditures, then the normal case would be more likely to result. The net effect also depends on the strength of the interest rate effect on investment, as measured here by $g_1$.

Next, consider whether demand is wage-led or profit-led. This depends on the sign of the partial derivative

$$\frac{\partial u^*}{\partial \pi} = \frac{-[g_0 + (1 - s_R - g_1) id_B](1 - g_1)(1/a_1)}{[(1 - g_1)(\pi/a_1) - g_2]^2} = \frac{-u^*(1 - g_1)(1/a_1)}{(1 - g_1)(\pi/a_1) - g_2} < 0 \tag{7.24}$$

Recalling that $g_1 < 1$ is necessary for stability and that $g_0 > -(1 - s_R - g_1) id_B$ is required for a positive solution, this derivative must be negative, indicating wage-led demand, which is not surprising given that it's based on a Kalecki–Steindl investment function in a model with no saving out of wages and no international trade.[22] But, what is most striking here is how the distinction between the normal and puzzling cases affects the *degree* to which the economy is wage-led. For any given debt service burden (interest payout ratio) $id_B$, the numerator will be less negative in the normal case where $1 - s_R - g_1 < 0$, indicating more *weakly* wage-led demand, and more negative in the puzzling case where $1 - s_R - g_1 > 0$, indicating more *strongly* wage-led demand. In addition, we can see that the level of the debt service burden $id_B$ affects the degree to which demand is wage-led differently in the normal versus the puzzling case. In the normal case, where $1 - s_R - g_1 < 0$, a higher debt service burden $id_B$ makes the numerator smaller in absolute value (less negative) and hence makes demand relatively *less* wage-led (although profit-led demand is not possible unless we modify the model, for example by using a different investment function, allowing for positive saving out of wages or introducing foreign trade). In contrast, in the puzzling case where $1 - s_R - g_1 > 0$, a rise in $id_B$ makes the denominator greater in absolute value (more negative), resulting in more strongly wage-led demand.

The results obtained so far have been based on the assumption that the profit share remains constant when the interest rate or debt service burden increases. However, as discussed in Chapter 4, neo-Kaleckian theory suggests that oligopolistic firms will try to raise their markup rates in response to higher fixed costs – which include interest payments to bondholders – thereby passing through part of the increased fixed costs to consumers (and reducing real wages in the process). To model this in a particularly simple way, suppose that the profit share (which is positively related to the markup, as we've seen previously) is an increasing function of the interest–capital ratio: $\pi = \pi(id_B), \pi' > 0$. In this case, the derivative (7.23) becomes

$$\frac{\partial u^*}{\partial(id_B)} = \frac{(1 - s_R - g_1) - u^*(1 - g_1)(1/a_1)\pi'}{(1 - g_1)(\pi/a_1) - g_2} \tag{7.23$'$}$$

where the endogenous adjustment of the profit share (the second term in the numerator) makes the impact of a rise in $id_B$ either more negative (in the normal case) or less positive (in the puzzling case). Indeed, if the adjustment of the profit share is sufficiently large, the derivative in (7.23′) could turn negative, even in the puzzling case. Intuitively, a rise in the interest rate (or debt burden) has a double redistributive effect in this extended version of the model: it redistributes part of profits from firms to rentiers, as before, but it also redistributes wages to profits by raising the profit share – which implies a reduction in the real wage, which is $w = (1 - \pi)/a_0$ in a neo-Kaleckian model as in Chapter 4. The latter distributive shift is definitely contractionary in a Kalecki–Steindl model, while the former can be either expansionary or contractionary depending on whether the economy exhibits normal or puzzling responses to increased interest payments.

### 7.3.2 A Bhaduri–Marglin model with shareholders

This section covers a model from Vasudevan (2017), in which the rentiers are corporate shareholders who own equity in the firms and receive dividends in proportion to their shares.[23] As in the previous subsection, the payouts to the rentiers reduce the retained profits of the firms, while channelling this paid out portion of the profits to a class that saves only part of its income and consumes the rest. Following Vasudevan, we adopt a simple specification in which the equity–capital ratio ($\chi > 0$) is exogenously given, and dividends on the equity shares are paid out at a fixed (exogenously given) rate $\zeta$ ($0 < \zeta < 1$). Rentiers (shareholders in this case) again save at the rate $s_R$ ($0 < s_R < 1$), while for simplicity workers do not save. Although Vasudevan assumes that the retained profits of the firms accrue to 'capitalists' who save at a higher rate than the rentiers ($s_r > s_R$), we prefer to think of the retained profits as accruing to the corporation as an institution, in which case by definition they are all saved (corporations cannot consume), so we will assume that $s_r = 1$.[24]

On these assumptions, the saving rate can be expressed as

$$\sigma = (r - \zeta\chi) + s_R\zeta\chi = r - (1 - s_R)\zeta\chi \tag{7.25}$$

where as usual $r = \pi u/a_1$ is the rate of profit. The first part of this equation says that saving is the sum of retained profits (profits net of dividend payouts) plus rentiers' saving out of dividend income, while the second part says that this is equivalent to total profits minus the consumption of the rentiers, all measured as ratios to the capital stock. Vasudevan (2017) then uses a modified version of a linearized Bhaduri–Marglin investment function,

$$g = h_0 + h_1\pi + h_2 u - h_3\zeta \tag{7.26}$$

where $h_1, h_2, h_3 > 0$ (a restriction on $h_0$ will be discussed below). The rationale for the $-h_3\zeta$ term is that payouts to shareholder households diminish the retained profits (which equal corporate saving) available for internal finance of investment.[25]

In this model, investment is financed by the issuance of new equity as well as by current saving, and since the equity–capital ratio $\chi$ is exogenously given, equity grows at the same rate as capital ($g$) and new equity issuance (measured as a ratio to the capital stock) equals $\chi g$. Therefore, macroeconomic equilibrium requires $g = \sigma + \chi g$, or, equivalently,

$$\sigma = (1 - \chi)g \tag{7.27}$$

Substituting equations (7.25) and (7.26) into (7.27) and using $r = \pi u/a_1$, we obtain the following solution for short-run equilibrium capacity utilization

$$u^* = \frac{(1 - s_R)\zeta\chi + (1 - \chi)(h_0 + h_1\pi - h_3\zeta)}{(\pi/a_1) - (1 - \chi)h_2} \tag{7.28}$$

where as usual the denominator has to be positive for Keynesian stability.[26] Therefore, the numerator must also be positive for an economically positive solution, which (assuming $\chi < 1$) means that $h_0$ cannot be too negative and $h_3\zeta$ cannot be too large.

Several interesting comparative statics follow from this solution. First, consider a rise in the proportion of capital financed by equity, $\chi$. The effect on equilibrium utilization is given by

$$\frac{\partial u^*}{\partial \chi} = \frac{(1 - s_R)\zeta - (h_0 + h_1\pi - h_3\zeta) - h_2 u^*}{(\pi/a_1) - (1-\chi)h_2} \tag{7.29}$$

Given that the denominator is positive, the sign of (7.29) depends on the sign of the numerator. Noting that the equilibrium growth rate is $g^* = h_0 + h_1\pi + h_2u^* - h_3\zeta$, the numerator will be positive if $(1 - s_R)\zeta > g^*$ and negative if $(1 - s_R)\zeta < g^*$. In other words, whether a greater degree of equity finance is expansionary or contractionary depends on whether the consumption out of increased shareholders' dividends is greater or less than the investment spending that is partially financed by the increase in equity (all measured as ratios to the capital stock). As in other models covered earlier in this chapter, if the rentiers are strong consumers (for example, retirees whose pensions are invested in equity shares), then more equity issuance is more likely to be expansionary, but if the rentiers are primarily savers (very wealthy households who are able to save most of their dividends), then more equity issuance is likely to be contractionary.

Next, consider an increase in the dividend payout rate $\zeta$. The effect on equilibrium utilization is given by

$$\frac{\partial u^*}{\partial \zeta} = \frac{(1 - s_R)\chi - (1-\chi)h_3}{(\pi/a_1) - (1-\chi)h_2} \tag{7.30}$$

where the sign again depends solely on the sign of the numerator, assuming Keynesian stability. This condition has an important economic interpretation. If $(1 - s_R)\chi > (1-\chi)h_3$, then the additional consumption out of increased rentiers' dividend income exceeds the reduction in investment due to increased dividend payouts by firms, and the increased dividend payouts are expansionary. Vasudevan calls this a 'shareholder-led regime', which again could be identified (for example) with a society in which equity is mainly held by retirees who use the dividends to pay for consumption, while corporations focus on investing in the future growth of their capital. However, if $(1 - s_R)\chi < (1-\chi)h_3$, then the reduction in investment exceeds the increase in consumption and the increase in the dividend payout rate is contractionary. Vasudevan calls this a 'shareholder-burdened regime', which she identifies with a 'financialized' economy in which rentier households pull funds out of the corporate sector and corporate executives are more concerned with creating 'shareholder value' than with the long-term growth of the firms' capital.

Finally, this model can be solved to find the condition for whether demand is wage-led or profit-led. For this purpose, Vasudevan adds the assumption that the dividend payout rate is an increasing function of the profit share, on the grounds that shareholders will expect to receive proportionately higher payouts when the firms are more profitable: $\zeta = \zeta(\pi), \zeta' > 0$. Using this function in (7.28), we can find the partial derivative

$$\frac{\partial u^*}{\partial \pi} = \frac{(1-\chi)h_1 - (u^*/a_1) + [(1-s_R)\chi - (1-\chi)h_3]\zeta'}{(\pi/a_1) - (1-\chi)h_2} \tag{7.31}$$

where once again the denominator is positive assuming the Keynesian stability condition. The first two terms in the numerator, $(1-\chi)h_1 - (u^*/a_1)$, represent the 'ordinary' determinants of whether demand is profit-led or wage-led in a Bhaduri–Marglin model, since this is equivalent to the difference between the response of firms' investment to a higher profit share and the increased saving out of those profits (recalling that we've assumed a saving rate of unity out of retained profits). If the former is greater, demand would be profit-led, while if the latter is greater, demand would be wage-led, in the absence of any change in the dividend payout rate. However, if the dividend payout rate is endogenous and $\zeta' > 0$, then there is an additional effect, the sign of which depends on whether the regime is shareholder-led $[(1-s_R)\chi > (1-\chi)h_3]$ or shareholder-burdened $[(1-s_R)\chi < (1-\chi)h_3]$. In the former case, the demand regime is more likely to be profit-led, since there is a high propensity of the rentiers to consume out of their increased dividends, while in the latter case the regime is more likely to be wage-led because of a weak response of rentiers' consumption to the induced rise in dividends.

The preceding discussion of the Hein–Lavoie and Vasudevan approaches gives some basic foundations for how financial variables and relationships can be incorporated into heterodox macro models. These models are just the tip of a very large and growing iceberg, with ever more work on the dynamics of finance and debt entering into heterodox growth theory (HGT). Moreover, as noted earlier, we have only covered their comparative static properties in this section while omitting the complex dynamics in the original presentations. In the next section, we turn to a few examples of recent work on macrodynamic models that are driven by financial relationships. To connect this discussion with earlier chapters, we will focus on two branches of this genre that serve as alternatives to the models of cyclical growth covered in earlier chapters, especially the Goodwin-type models (covered in Chapters 2 and 5) and the neo-Harrodian models (covered in Chapter 6). One of these branches draws on the literature on Minskyan financial fragility, while the other is part of the literature on household debt. In addition to their focus on financial dynamics, both models demonstrate that real–financial interactions can complicate our interpretation of the relationship between distribution and growth.

## 7.4 Wolves in sheep's clothing? Models that mimic profit-led dynamics

As discussed in Chapter 5, there is much controversy over the findings of profit-led demand and neo-Goodwinian dynamics in some recent empirical studies, especially those employing an aggregative method. But since capacity utilization and the wage share do often exhibit counterclockwise rotations, the question naturally arises as to what kinds of causal mechanisms could account for this behaviour if it is not in fact driven by the combination of profit-led demand and a profit-squeeze in distribution. Or, to put the point another way, what sort of dynamics could cause the appearance of neo-Goodwin cycles even if an economy does not exhibit such features? Inspired by the process of financialization and, in particular, the sharp increase in *household* debt in contemporary capitalist economies, various models have been proposed to demonstrate that seemingly profit-led macrodynamics can arise because of features of the financial system rather than properties of the real sector of the economy. In these models, the real sector may even be wage-led, but real–financial interactions nevertheless produce seemingly profit-led dynamics. One possibility associated with these models is that seemingly profit-led macrodynamics may be a result of *particular financial institutions* rather than any necessary relationship between distribution and growth.[27] This, in turn, raises the possibility that the relationship between distribution and growth is a historically-specific social construct, that may even be amenable to change by policy interventions (Palley, 2014, 2017; Carvalho and Rezai, 2016; Setterfield and Kim, 2017, pp. 54–5). In the remainder of this section, we will focus on two classes of models that can be associated with this broad result.

### 7.4.1 Pseudo-Goodwin cycles produced by financial fragility

It was demonstrated in Chapter 5 that a profit-led demand regime is a necessary condition for neo-Goodwin cycles – that is, interactions between the wage share and capacity utilization rate that produce a pattern of counterclockwise rotation in capacity utilization–wage share space (see Figure 5.6, panel (d)). As such, the appearance of this pattern in actual economic data is often interpreted as evidence that real-world aggregate demand regimes are, in fact, profit-led. It was also noted in Chapter 5, however, that this does not necessarily mean that the entire economic system is profit-led – that is, that boosting profitability will necessarily raise growth and capacity utilization as the economy transitions from one steady-state equilibrium position to another. It turns out, in fact, that the observation of apparent neo-Goodwin cycles does not even mean that the demand regime is profit-led, at least once we begin to take into account financial factors that may interact with real-sector developments in the course of growth.

This result has been demonstrated in two companion papers (Stockhammer and Michell, 2017; Stockhammer, 2017) that combine a demand regime with a profit-squeeze type distribution function *and* a financial fragility equation, which is designed to capture the essence of real–financial interactions in Minsky's financial instability hypothesis (to which we referred briefly in Chapter 6).[28] The complete model can be stated as follows

$$\dot{f} = f(-1 + \gamma_1 Y) \tag{7.32}$$

$$\dot{Y} = Y(1 - f) \tag{7.33}$$

$$\dot{\psi} = \psi(-\gamma_2 + \gamma_3 Y - \gamma_4 \psi) \tag{7.34}$$

where $Y$ is real output, $\psi$ is the wage share, $f$ is a measure of financial fragility and various parameters have been normalized to values of one for simplicity. Equation (7.32) describes the rate of increase of financial fragility as increasing in the level of real output. This equation is intended to capture, in simplified form, the relationship between economic activity and financial fragility in Minsky's (1982, 1986) financial instability hypothesis, according to which as the economy grows, firms take on more debt (and of increasingly precarious forms) so that the financial fragility of the economy (its vulnerability to, for example, a sudden increase in interest rates and/or a shock to the level of income) increases. Equation (7.33), meanwhile, describes goods market dynamics as being related inversely to the degree of financial fragility. This is a substitute for equation (5.22) or (5.23) in Chapter 5 (relating goods market dynamics to the state of the goods market itself and the wage share). Equation (7.33) states that financial fragility has an adverse effect on the rate of increase of real output, and is intended to capture the adverse consequences for the goods market of the build-up of financial fragility noted above. Unlike equations (5.22)–(5.23), equation

(7.33) involves no feedback from distribution to demand. As a result, the dynamics of the goods market are not profit-led (because there is *no effect* of distribution on demand), so that whatever the results of the system outlined above, these results *cannot* be the by-product of a profit-led demand regime.

Finally, equation (7.34) is analogous to equation (5.19) or (5.24) from Chapter 5, with $\omega_2 = -\gamma_4 < 0$ (so that the wage share feeds back negatively on its own rate of increase which is, therefore, self-stabilizing, *ceteris paribus*) and $\omega_1 \approx \gamma_3 > 0$ (so that the rate of increase of the wage share increases in the level of economic activity). As was demonstrated in Chapter 5, the sign of this last coefficient will create a profit-squeeze effect, whereby increases in the level of real economic activity will boost the wage share and so reduce the profit share.

Consider first the interaction of equations (7.32) and (7.33) in isolation. Setting $\dot{f} = \dot{Y} = 0$ in order to solve for steady-state equilibrium values of $f$ and $Y$ (and ignoring the solution $f = Y = 0$), we obtain

$$f^* = 1, Y^* = \frac{1}{\gamma_1} \tag{7.35}$$

Meanwhile, the Jacobian of equations (7.32) and (7.33) evaluated at this steady state is

$$\mathbf{J} = \begin{bmatrix} 0 & \gamma_1 \\ -\dfrac{1}{\gamma_1} & 0 \end{bmatrix} \tag{7.36}$$

from which it follows that $Det(\mathbf{J}) = 1 > 0$ and $Tr(\mathbf{J}) = 0$, which outcomes yield a limit cycle. More specifically, the resulting system produces counter-clockwise movement in $f \times Y$ space. The important thing to note about the sub-system in equations (7.32) and (7.33) is that its dynamics result entirely from the interaction of the goods market and financial conditions: there is no causal role whatsoever for distribution.

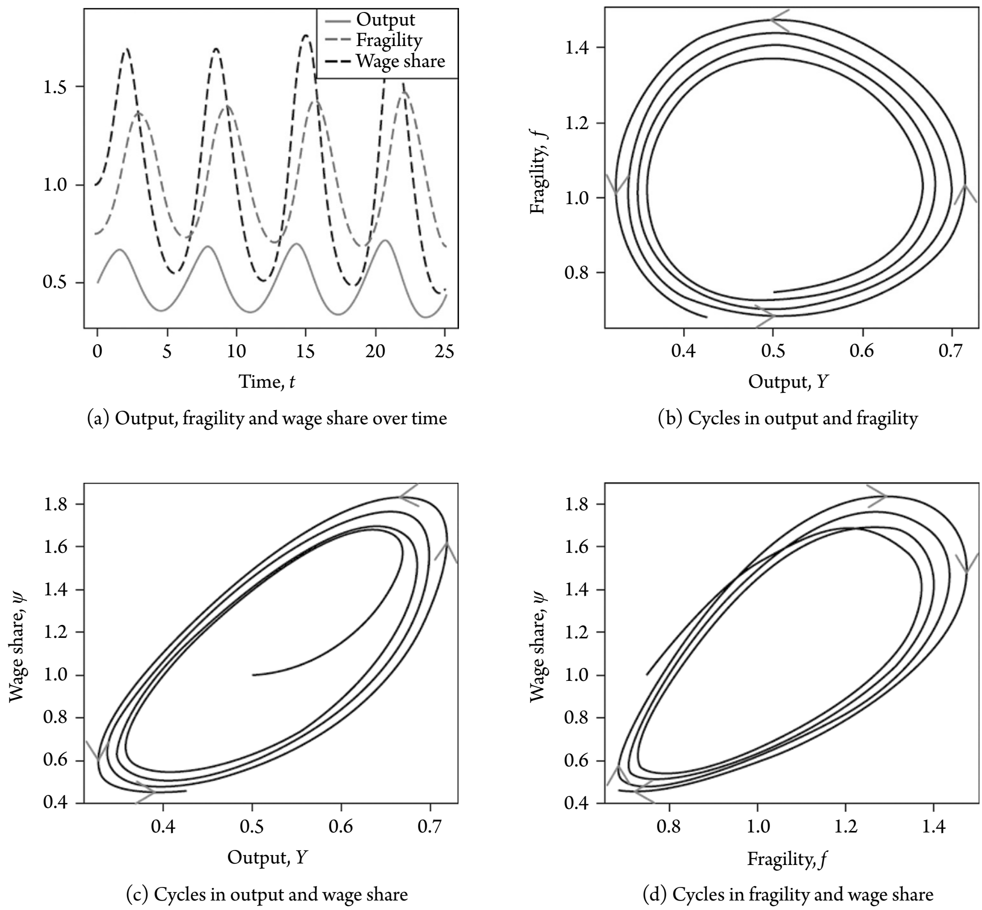
The importance of this last observation can be seen if we now extend our analysis to include equation (7.34). Note what is achieved by this extension: we now introduce distributional dynamics into the system by describing the rate of change of the wage share as a function of itself and the state of the goods market. But the distribution of income is no more than an adjusting residual: since equations (7.32) and (7.33) are unaltered, there is no causal role for the wage share in the dynamics of the system, the drivers of which are still $f$ and $Y$ (as a result of the self-contained interaction of equations 7.32 and 7.33 described above).

If we once again set $\dot{f} = \dot{Y} = \dot{\psi} = 0$ in order to solve for the steady-state equilibrium values of $f$, $Y$ and $\psi$ (again ignoring the solution $f = Y = \psi = 0$), we obtain

$$f^* = 1, Y^* = \frac{1}{\gamma_1}, \psi^* = -\frac{1}{\gamma_4}\left(\gamma_2 - \frac{\gamma_3}{\gamma_1}\right) \tag{7.37}$$

where we assume that $(\gamma_2 - \frac{\gamma_3}{\gamma_1}) < 0$ in order to satisfy $\psi^* > 0$. As can be seen from (7.37) – and not surprisingly in light of what has been said above – the equilibrium solutions for $f$ and $Y$ are unaffected by explicit consideration of equation (7.34). The stability of the system is affected, however, and with interesting consequences for the co-movements of the endogenous variables $f$, $Y$ and $\psi$ in the vicinity of the equilibrium in (7.37). Specifically, we must now consider the stability of the three-dimensional system of differential equations in (7.32)–(7.34). This analysis is conducted formally in Appendix 7.1.

What emerges from the stability analysis is a limit cycle that produces counterclockwise movements in $Y \times \psi$ space (Stockhammer and Michell, 2017, p. 114). This counterclockwise rotation of output and the wage share is illustrated in panel (c) of Figure 7.2. Panels (b) and (d) in the same figure illustrate the cyclical motion of other two-dimensional relationships, between output and financial fragility, and financial fragility and the wage share, that arises from the three-dimensional dynamics of the Stockhammer–Michell model, while panel (a) shows the cyclical trajectories of the three variables over time. Note, however, that despite its superficial resemblance to a neo-Goodwin cycle, the counterclockwise rotation of the wage share and output observed in Figure 7.2(c) is merely a *side effect* of dynamics that result entirely from the interaction of the goods market and financial conditions. As remarked above, despite the inclusion of equation (7.34) in our analysis, the distributional dynamics it describes are a passive residual response to the interaction of equations (7.32) and (7.33) in a system in which there is no causal role for distribution. In view of all this, Stockhammer and Michell (2017, p. 114) call the pattern of counterclockwise movement in $Y \times \psi$ space seen in Figure 7.2(c) a *pseudo-Goodwin cycle*, because it does *not* derive from the sort of profit-led/profit-squeeze dynamics associated with the original Goodwin




Source: Stockhammer (2017, p. 38). Reproduced with permission: output and wage share notation changed to match this book.

**Figure 7.2** Cyclical outcomes in the Stockhammer–Michell model

model or its neo-Goodwinian counterpart.[29] As Stockhammer (2017, p. 39) remarks, 'the pseudo-Goodwin cycle is generated as a side-effect as distribution is dragged along by fluctuations in output that are driven by financial factors'.

For good measure, Stockhammer and Michell (2017, pp. 115–17) extend the model outlined above by including a positive influence of the wage share on goods market dynamics in equation (7.33), thereby making the demand regime formally *wage*-led. They show that the same broad results outlined above – the emergence of pseudo-Goodwin cycles – obtain once again, with the exception that the cycles in $Y \times \psi$ space are explosive rather than conforming to a limit cycle. Hence seemingly Goodwinian patterns in the relationship between the level of economic activity and the wage share can emerge as a result of financial dynamics in a model in which the demand regime either *lacks* any causal influence of income distribution on the goods market or exhibits *explicitly wage-led* features. According to Stockhammer and Michell (2017), this demonstrates that the observation of what appear to be neo-Goodwin cycles in aggregate economic data does *not* necessarily provide evidence that aggregate demand is truly profit-led. More generally – and in keeping with the theme of this section as a whole – it demonstrates that real–financial interactions that are not, themselves, provoked by changes in the distribution of income can produce seemingly profit-led macrodynamics. In other words, the latter are a result of *particular financial institutions*, rather than any fundamental (causal) relationship between distribution and growth of the sort that the very notion of profit-led macrodynamics would appear to suggest.

### 7.4.2 Consumption-driven, profit-led growth

A commonplace argument in heterodox macroeconomics is that the relatively rapid growth of the US economy in the 1990–2007 period owed, in large part, to the willingness and ability of less affluent households to borrow in order to offset the otherwise negative impact on their consumption spending of increased income inequality (Palley, 2002a; Pollin, 2005; Brown, 2008; Cynamon and Fazzari, 2008; Barba and Pivetti, 2009; Setterfield, 2013b; Wisman, 2013). The counterpart to this argument in HGT has been the development of models that incorporate household debt accumulation and its consequences into various of the frameworks of analysis that have been outlined previously in this book.[30] The primary purpose of these models is to respond to the contemporary developments in capitalism called to attention by the literature just cited. They also, however, have important consequences for the relationship between distribution and growth and, in particular, what it means for economic activity to be profit-led.

In order to illustrate what is at stake in these models, it is useful to focus on the model developed by Setterfield and Kim (2017). This model blends several elements that will be familiar to the reader from previous chapters and even from earlier sections of this chapter, including wage inequality, rentier income and the debate over whether the economy is wage-led or profit-led. Indeed, the basic 'chassis' of the model is essentially a Kalecki-Robinson growth model of the sort presented in section 6.4 of Chapter 6, which combines features of the neo-Robinsonian model from Chapter 3 and the neo-Kaleckian model from Chapter 4. Recall from Chapter 6 that the equilibrium solution of this model can be found by equating equations (6.21) and (6.22), using the equilibrium condition $g = \sigma$, and then solving for $u$. As shown in section 6.4.2, this series of operations yields the solution

$$u^* = \frac{a_1 f_0}{(s_r - f_1)\pi} \tag{7.38}$$

Substituting this last expression into equation (6.21), we can write

$$g^* = \frac{s_r f_0}{s_r - f_1} \tag{7.39}$$

The operative point that emerges from equation (7.39) is that $\partial g^*/\partial\pi = 0$. In other words, the rate of growth (unlike the utilization rate) is unresponsive to changes in the profit share of income: growth is neither wage- nor profit-led.

Setterfield and Kim show that this distributional neutrality of the growth regime is transformed by the introduction of working households who borrow to finance consumption spending in an effort to 'keep up with the Joneses' – that is, pursue a consumption target based on the consumption spending of more affluent households. Specifically, the distributional neutrality of the growth regime is no longer assured: growth is, instead, more likely to be profit-led.

The essence of this result can be captured by examining the way in which consumption spending is transformed by a combination of borrowing and emulation-based consumption targeting on the part of working households. In Setterfield and Kim (2017, p. 50), consumption spending is described as follows:

$$C = C_w + C_R + \dot{D} \tag{7.40}$$

$$C_w = c_w W_L L \tag{7.41}$$

$$C_R = c_R\left(\frac{\omega}{\theta} W_L L + \Pi + iD_R\right) \tag{7.42}$$

$$\dot{D} = \beta(C^T - C_w) \tag{7.43}$$

$$C^T = \lambda C_R \tag{7.44}$$

where $C$, $C_w$ and $C_R$ denote aggregate consumption, consumption by workers and consumption by rentiers (managers and capitalists), respectively, $\dot{D}$ is borrowing by workers,[31] $iD_R$ is debt servicing by workers (the product of the real interest rate, $i$, and the debt owed to rentier households, $D_R$),[32] and $C^T$ is workers' target level of consumption spending. As previously defined,

$W_L$, $L$ and $\Pi$ are the real wage earned by workers, the number of workers employed and total profits, respectively. Equation (7.40) describes the composition of total consumption spending. Equation (7.41) states that workers consume a conventional fraction of their total wage income, based on their marginal propensity to consume $c_w$. Note that a corollary of this behaviour is that saving becomes a 'residual of a residual' – what remains after income has been consumed *and* debt servicing obligations have been met. In other words, saving by workers ($S_w$) can be written as

$$S_w = (1 - c_w) W_L L - iD_R \tag{7.45}$$

Equation (7.42) describes rentiers' consumption as a conventional fraction, represented by the marginal propensity to consume $c_R$ of their total wage, profit and debt-servicing income. As in the Tavani and Vasudevan (2014) model discussed in section 7.2.2, the managers' wage income depends on the ratio of managerial workers to production workers, $1/\theta$, and the ratio of managerial salaries to production workers' wages, $\omega$. Finally, equations (7.43) and (7.44) describe borrowing by workers as depending on the difference between their consumption target and the consumption they can fund from wage income, where the workers' consumption target is defined as a fraction ($\lambda$) of rentiers' consumption, to reflect working-class households' emulation of rentiers' 'conspicuous consumption' (in the sense of Veblen, 1912).

Combining equations (7.40)–(7.44) and collecting like terms, we get

$$C = \left([1 - \beta]c_w + [1 + \beta\lambda]c_R\frac{\omega}{\theta}\right)W_L L + (1 + \beta\lambda)c_R(\Pi + iD_R)$$

$$\Rightarrow C = \left([1 - \beta]c_w + [1 + \beta\lambda]c_R\frac{\omega}{\theta}\right)\frac{1 - \pi}{1 + \omega/\theta}Y + (1 + \beta\lambda)c_R(\pi Y + iD_R) \tag{7.46}$$

where $\frac{1-\pi}{1+\omega/\theta} = \psi_L$ is the workers' share of total income.[33] Equation (7.46) can be rewritten as

$$C = \Omega + \left(\frac{[1 + \beta\lambda]c_R - [1 - \beta]c_w}{1 + \omega/\theta}\right)\pi Y \tag{7.47}$$

where

$$\Omega = \frac{1}{1 + \omega/\theta}\left([1 - \beta]c_w + [1 + \beta\lambda]c_R\frac{\omega}{\theta}\right)Y + (1 + \beta\lambda)c_R iD_R$$

Consistent with equation (7.47), we can now write

$$\left.\frac{\partial C}{\partial \pi}\right|_{Y=\bar{Y}} = \left(\frac{[1+\beta\lambda]c_R - [1-\beta]c_w}{1+\omega/\theta}\right)Y > 0$$

if[34]

$$[1+\beta\lambda]c_R - [1-\beta]c_w > 0 \qquad (7.48)$$

$$\Rightarrow (c_R - c_w) + \beta(\lambda c_R + c_w) > 0$$

The first term in parentheses on the left-hand side of (7.48) can be considered negative by hypothesis (the propensity to spend of workers exceeds that of rentiers). But with borrowing and emulation ($\beta, \lambda > 0$) the second term on the left-hand side will always be positive, and if this second term is large enough, the response of aggregate consumption spending to an increase in the profit share of income will be positive. What this demonstrates is that borrowing and emulation behaviour can result in a redistribution of income towards profits, boosting demand formation and hence growth through the consumption channel. Setterfield and Kim (2017, p. 52) call this the 'paradox of inequality' whereby, contrary to conventional Keynesian wisdom, transferring income from high propensity to consume workers to low propensity to consume rentiers raises total consumption spending. This paradox is also the essence of what Kapeller and Schütz (2015) call 'consumption-driven, profit-led growth'. The growth regime is profit-led in the sense that a redistribution towards profits raises growth. But this is *not* because the impact of an increase in the profit share on growth operating through the investment channel outweighs its impact operating through the consumption channel, as in the Bhaduri–Marglin model introduced in Chapter 4 and some of the extensions of that model covered earlier in this chapter. Instead, motivated by a level of inequality that impairs their consumption spending, and facilitated by financial institutions, borrowing and emulation induce working households to more than offset the drop in consumption out of current income that results from a redistribution towards profit (because $c_R - c_w < 0$) by increasing their debt-financed consumption spending in an effort to keep up with the Joneses.

As demonstrated by Kapeller and Schütz (2015), consumption-driven, profit-led growth does not depend on introducing borrowing and emulation effects by less affluent households into a growth regime in which the rate of growth is initially invariant to the distribution of income (as in Setterfield and Kim, 2017). Kapeller and Schütz (2015, p. 59) posit a Bhaduri–Marglin type investment function, in which the rate of investment depends on the profit share and the rate of utilization and that, as demonstrated in Chapter 4, is capable (in the absence of borrowing and emulation effects) of producing wage-led growth outcomes. They show that results similar to those reported above – that is, an economy that is more inclined to produce profit-led growth outcomes – emerge even in this environment.

One final point of interest that has emerged from models of this genus is revealed by closer inspection of the term $\Omega$ that appears in equation (7.47) above. Hence note that

$$\left.\frac{\partial C}{\partial (iD_R)}\right|_{Y=\bar{Y}} = \left.\frac{\partial A}{\partial (iD_R)}\right|_{Y=\bar{Y}} = (1+\beta\lambda)c_R > 0 \qquad (7.49)$$

This derivative reveals that an increase in debt-servicing payments (by workers to rentiers) will *boost* consumption spending (and hence aggregate demand formation and growth). The result so obtained once again contradicts ordinary Keynesian logic, according to which even if borrowing boosts aggregate demand formation, the effects of borrowing – that is, debt accumulation – will eventually create a drag on the economy, by setting up ever-larger transfers of income away from high-spending, less affluent, debtor households towards lower-spending, more affluent, creditor households. Why does this ordinary Keynesian logic fail in the Setterfield and Kim (2017) model? The answer is surprisingly straightforward (but seemingly mundane): it is brought about by *the way in which debtor households service their debts*. Note, then, that according to (7.41) and (7.45), workers treat debt-servicing obligations as an expense (Cynamon and Fazzari, 2017a), and in so doing choose to privilege consumption spending out of their current income (equation 7.41) thus relegating saving to the status of residual of a residual: what is left after consumption spending and debt-servicing expenses have been funded (equation 7.45). This behaviour means that in the event of an increase in debt-servicing payments, some part of their wage income that workers would otherwise have saved (equation 7.45) is transferred to rentier households *and partially spent* (note that debt-servicing income contributes positively to total rentier income and hence rentier consumption, as in equation 7.42). In other words, debt servicing transforms a leakage from the circular flow of income into an injection and, in the process, boosts aggregate demand! The importance of this result, as emphasized by Setterfield et al. (2016), is that in an economy in which household borrowing contributes to demand formation, even the way in which households choose to service their debts (a topic as seemingly mundane as they come) can have an important effect on macrodynamics.

Thus, the propensity of an economy to be profit-led may be the product of demand formation by households through the consumption channel, rather than the interaction of the consumption and investment channels as originally emphasized in the Bhaduri–Marglin model. The significance of this observation is that profit-led growth may be a result of how low the wage share has already sunk, rather than a fundamental property of redistribution towards profits per se, and the concomitant proclivity of less affluent households to borrow in order to make up for lost ground relative to more affluent households. Note that this interpretation is consistent with the underlying behavioural assumptions of the models developed in this subsection: that income inequality has progressed to a point where working households need to borrow to finance consumption spending that they cannot fund out of income, given their (socially informed) consumption aspirations. As Cynamon and Fazzari (2008) argue, financial arrangements are required to make this work: households need to relax their borrowing norms, and creditors must relax their lending norms. Historically-specific financial institutions are once again important in shaping the relationship between distribution and growth, then. But unlike the Stockhammer and Michell (2017) model in which the financial sector is a key *driver* of macrodynamic outcomes, financial institutions play more of a *facilitating* role in the models of consumption-driven, profit-led growth due to Kapeller and Schütz (2015) and Setterfield and Kim (2017). The underlying cause of profit-led growth outcomes in these models is, in fact, the emergence of gross inequality in the distribution of income itself.

## 7.5 Supermultiplier models with exogenous components of demand growth

In Chapter 6, brief mention was made of a relatively recent class of Sraffian supermultiplier models that have, in turn, inspired neo-Kaleckian contributions to the Harrodian instability debate. The purpose of this section is to examine this class of models in its own right, starting with discussion of the supermultiplier concept itself. This is important because as we shall see, the supermultiplier provides the analytical basis not only for the Sraffian and neo-Kaleckian models discussed in what follows, but also for the modern Kaldorian approach to growth theory that is discussed in Part III of this book. One of the most important features of supermultiplier analysis is the weight that it places on the role of autonomous demand in the determination of equilibrium growth. So prominent is this role that supermultiplier analysis has ignited a debate among growth theorists about an apparent turn towards 'exogenous growth theory' in heterodox macrodynamics – this characterization having formerly been unique to the first-generation

(Solow) neoclassical growth model outlined in Chapter 1. In the remainder of this section we will describe the supermultiplier concept, outline recent Sraffian and neo-Kaleckian models that draw inspiration from this concept, and then discuss the alleged turn towards exogenous growth theory that these models represent. Chapter 8 will return to the supermultiplier concept (and the exogeneity of the growth rate) in the context of Kaldorian growth theory.

## 7.5.1 The supermultiplier concept

The term 'supermultiplier' was first introduced by Hicks (1950). Simply put – and as its name suggests – a supermultiplier is an expanded version of the familiar Keynesian expenditure multiplier, which makes the size of the latter larger. Consider, for example, the following system of equations, describing a simple static model of output determination

$$Y = C + I + A \tag{7.50}$$

$$C = cY \tag{7.51}$$

$$I = a_1 \Delta Y = a_1 yY \tag{7.52}$$

$$A = \overline{A} \tag{7.53}$$

where $Y$ is real output, $C$, $I$ and $A$ are real consumption, investment and autonomous demand (respectively), $y = \Delta Y/Y$ is the rate of growth, and $c$ and $a_1$ are (respectively) the propensity to consume and the ratio of capital to full-capacity output. Equation (7.51) is a simple proportional consumption function, and equation (7.52) describes investment spending in terms of the accelerator principle that was discussed extensively in Chapter 3. The exogenous component of demand, $A$, might represent exports, government spending or an exogenous component of either consumption spending by households or investment spending by firms. In different heterodox growth models that make use of the supermultiplier concept, it takes on different meanings – but for now we need not concern ourselves with exactly what $A$ can/should be thought to represent, focusing instead on its exogeneity.

Suppose that, to begin with, we focus on the solution of equations (7.50), (7.51) and (7.53), assuming for simplicity that $I = 0$.[35] This yields

$$Y = \frac{1}{1 - c}\overline{A} \tag{7.54}$$

The first term on the right-hand side of this expression is immediately recognizable as the simple Keynesian multiplier. In an expression of this sort, $A$ is very often assumed to represent total investment spending by firms, on the basis that the latter is determined with reference to expectational and monetary variables that are independent of current income. But suppose instead we introduce the accelerator theory of investment in (7.52) and now solve (7.50)–(7.53). This yields

$$Y = \frac{1}{1 - (c + a_1 y)} \bar{A} \tag{7.55}$$

The first term on the right-hand side of equation (7.55) is a *supermultiplier*. Note that since $a_1 y >> 0$, the supermultiplier just derived is strictly larger than the simple Keynesian expenditure multiplier identified previously.

As mentioned, different heterodox growth models appeal to the supermultiplier concept. In the remainder of this section, we will survey recent contributions from Sraffian and Kaleckian authors that appeal to this concept. We will return to the concept of the supermultiplier in Chapter 8, where it will be shown to be foundational to modern Kaldorian theories of growth that are based on the notion that growth in open economies (or particular global regions) is demand-driven and specifically export-led. Note, however, that supermultiplier models of growth share one important feature that is evident from even the simple static model described in equations (7.50)–(7.53). Hence note that by rearranging equation (7.55) to solve for the rate of growth $y$, we get

$$y = \frac{1}{a_1}\left(1 - c - \frac{\bar{A}}{Y}\right) \tag{7.56}$$

If the expression for $y$ in (7.56) is to be interpreted as a steady-state growth rate, then the ratio $\bar{A}/Y$ – the share of autonomous expenditures in total output – must be constant. In other words

$$\frac{\bar{A}}{Y} = \omega_A \tag{7.57}$$

where $\omega_A$ is some arbitrary constant,[36] so that

$$\hat{A} - \hat{Y} = 0$$

$$\Rightarrow y = \hat{A} \tag{7.58}$$

where $\hat{A}$ is the exogenously given rate of growth of autonomous demand.

Equation (7.58) bears an important resemblance to equation (1.32) from Chapter 1, which expresses the equilibrium rate of growth in the first-generation (Solow) neoclassical growth model as

$$y = \bar{q} + \bar{n} \tag{1.32}$$

In both cases, the long-run, steady-state rate of growth is simply exogenously given – in (7.58) by the (exogenously given) rate of growth of the autonomous component of aggregate demand. We have now created a new (heterodox) species of exogenous growth theory, comparable to the first-generation neoclassical growth theory (NGT) discussed in Chapter 1, wherein the rate of growth is imposed from without and determined independently of the equilibrium solution of the model. Of course, the two models remain substantively different insofar as the exogenous determinant of growth in equation (7.58) is found on the demand side of the economy (rather than the supply side, as in the Solow model). This difference is far from trivial. In fact, it encapsulates the Say's law versus 'Say's law in reverse' distinction between supply- and demand-led growth theories first raised in Chapter 1: the NGT exogenous growth model assumes that aggregate demand passively adjusts to accommodate a rate of growth exogenously determined on the supply side, whereas the HGT exogenous growth outcome in (7.58) assumes exactly the opposite – that supply conditions passively adjust to accommodate a rate of growth exogenously determined on the demand side. Nevertheless, the two models share a distinct methodological affinity, by effectively imposing the equilibrium rate of growth on the economy from without and offering no explanation of how this rate of growth is determined. We will return to discuss this feature of supermultiplier models towards the end of this section.

## 7.5.2 Sraffian supermultiplier models

As discussed previously in Chapters 1 and 3, part of the ambition of Sraffian or neo-Ricardian economics is to integrate classical surplus value theory with Keynes's principle of effective demand. Sraffian supermultiplier models of growth and distribution are therefore principally concerned with articulating a theory of demand-led growth in which distribution is exogenous (determined along Sraffian lines), and any steady-state equilibrium outcome is consistent with a classical 'fully adjusted' position (where market prices are equal to normal prices, yielding a uniform rate of profit, and capacity utilization is at its normal rate). The upshot is a post-Keynesian macrodynamic model that differs in certain crucial respects (to which we will draw attention below) from other (neo-Keynesian, neo-Kaleckian and neo-Harrodian) post-Keynesian models.[37]

Serrano (1995) is typically credited as the progenitor of the Sraffian supermultiplier model. Subsequent contributions to this tradition include Bortis (1997) and, more recently, Cesaratto (2015), Freitas and Serrano (2015), Pariboni (2016) and Serrano and Freitas (2017). Here, we will outline a model based on Serrano and Freitas (2017) which, in turn, draws on Freitas and Serrano (2015).

We begin with the familiar ratio of capital to full-capacity output, $a_1 = K/Y_K$. In other words, productive activity is never labour constrained; the dual economy assumption first noted in Chapter 1 is assumed to hold. It follows from the assumed fixity of $a_1$ (and absence of any depreciation of capital) that

$$y_K = g = \frac{I}{K} = \frac{I}{Y}\frac{Y}{Y_K}\frac{Y_K}{K} = \frac{(I/Y)}{a_1}u \tag{7.59}$$

where $u = Y/Y_K$ is the actual capacity utilization rate and $I/Y$ is the investment–output ratio. It follows, in turn, from the definition of $u$ and the equality of $y_K$ and $g$ reported in (7.59) that

$$\hat{u} = y - y_K$$

$$\Rightarrow \dot{u} = u(y - g) \tag{7.60}$$

Serrano and Freitas (2017, p. 72) imbue equation (7.60) with a behavioural interpretation according to which, in a competitive goods market, output adjusts faster than productive capacity in response to changes in aggregate demand (making the actual rate of capacity utilization variable in the short run). This, in turn, is the result of the economy operating in all periods at normal prices consistent with a uniform rate of profit calculated at the normal rate of capacity utilization, $u_n$ – all of which are distinctly classical (and hence Sraffian) features of the long period. At the same time, it follows from the behavioural interpretation of (7.60) that aggregate output at any point in time is demand-determined – which gives licence to describe determination of the level of output in terms of a *supermultiplier*.

The exact supermultiplier model presented in Serrano and Freitas (2017) is a variant of the generic supermultiplier model outlined at the beginning of this section, and can be written as

$$Y = C + I \tag{7.61}$$

$$C = wL + A = \psi Y + A \tag{7.62}$$

$$I = hY \tag{7.63}$$

$$A = \overline{A} \tag{7.64}$$

where the exogenously given component of aggregate demand, $A$, is now explicitly interpreted as exogenous *consumption* demand. Otherwise, consumption consists of wage income (that is, it is assumed that workers spend what they earn while all profit income is saved). Equation (7.63), meanwhile, is a variant of the accelerator relationship that makes investment spending endogenous to output on the basis of the premise that firms create capacity in order to meet demand. Note that in (7.63), $h$ is both the investment–output ratio *and* the 'marginal propensity to invest' $(dI/dY)$. It may appear from (7.63) that $h$ is parametric, but as will become clear this is not the case: $h$ is, instead, an endogenous variable that takes on a specific constant value only as a result of the steady-state solution of the model as a whole. Unlike our previous encounters with the accelerator mechanism, equation (7.63) thus embodies a flexible accelerator principle (see, for example, Freitas and Serrano, 2015, p. 270).

Finally, note that, according to equation (7.62), $\partial C/\partial Y = \psi$. In other words, the marginal propensity to consume is given by the wage share of income. It follows that the marginal propensity to save, $s$, is given by $s = 1 - \psi = \pi$, so that the propensity to save is identical to the profit share. As we will see, this observation becomes significant when we come to assess the responsiveness of growth to the saving propensity and the distribution of income.

Bearing in mind the result just obtained, solving equations (7.61)–(7.64) for $Y$ yields

$$Y = \left(\frac{1}{s - h}\right)\overline{A} \tag{7.65}$$

where the ratio in parentheses is the Sraffian supermultiplier.

We are now in a position to analyse the implications of the model for growth, both in the short run and in the long-run steady state. To begin with, note that in short-run equilibrium with $S = I$, and bearing in mind equation (7.63)

$$\frac{S}{Y} = \frac{I}{Y} = h \tag{7.66}$$

Meanwhile, Serrano and Freitas (2017, p. 74) write

$$\dot{h} = h\gamma(u - u_n), \gamma > 0 \tag{7.67}$$

In other words, firms increase (decrease) their marginal propensity to invest as capacity utilization rises above (below) its normal rate. Intuitively, this is consistent with firms modifying their investment in order to both keep pace with the expansion of output and ensure that they install sufficient capacity to keep the rate of capacity utilization at its normal rate in the long run. Since changes in $h$ are determined by firms in accordance with (7.67), and the saving ratio depends on $h$ as in (7.66), Serrano and Freitas (2017, pp. 70–71, 89) suggest that a key feature of their model is that the investment share of output determines the saving ratio. This is, indeed, true in the short run. Note, however, that it follows from equation (7.65) that

$$h = s - \frac{\bar{A}}{Y} \tag{7.68}$$

This points to the fact that the ultimate driver of $h$ is, therefore, the behaviour of the exogenously given component of consumption spending, $A$.[38] There is, then, a hierarchy in the model, according to which the investment–output ratio drives the saving ratio in the short run, whereas ultimately (in the long run) *both* variables are determined by the exogenous component of demand. Hence as Serrano and Freitas (2017, p. 74) note, 'the existence of a positive level of autonomous consumption is sufficient to make the saving ratio an endogenous variable . . . This endogenous determination of the saving ratio, with a given level of income distribution, is the distinctive feature of the . . . Sraffian supermultiplier growth model.' With respect to this last claim, note that endogenous adjustment of the saving ratio is also a feature of some of the early neo-Keynesian models surveyed in Chapter 3 – in which this is achieved by variation in the income distribution.[39] Meanwhile, most of the neo-Kaleckian models surveyed in Chapter 4 treat the distribution of income as exogenously given, but posit a *fixed* saving ratio.[40] While neo-Keynesian, neo-Kaleckian and Sraffian supermultiplier models are all part of the same broad post-Keynesian tradition, then, the Sraffian approach is rendered distinct by its capacity to combine demand-led growth with an exogenous distribution of income *and* variability in the saving ratio. As noted, this is ultimately achieved by the model's inclusion of an exogenously given, non-capacity creating component of aggregate demand.

Referring back to the structure of the supermultiplier model outlined above, by combining equations (7.61)–(7.63) we can write

$$Y = \psi Y + hY + A$$

$$\Rightarrow (1 - \psi)Y = sY = hY + A$$

$$\Rightarrow sy = \dot{h} + hy + g_A\frac{A}{Y}$$

$$\Rightarrow y = \frac{1}{s-h}\left(\dot{h} + g_A\frac{A}{Y}\right) \tag{7.69}$$

where $g_A$ denotes the (exogenously given) rate of growth of $A$. Finally, substituting equation (7.67) into this last expression, we arrive at

$$y = g_A + \frac{h\gamma(u - u_n)}{s-h} \tag{7.70}$$

Equation (7.70) describes the rate of growth of output in the Sraffian supermultiplier model in the short run. Substituting (7.59) and (7.70) into equation (7.60) and recalling the definition of $h$, we arrive at

$$\dot{u} = u\left(g_A + \frac{h\gamma(u-u_n)}{s-h} - \frac{h}{a_1}u\right) \tag{7.71}$$

Together, equations (7.67) and (7.71) form a system of two differential equations that completely describe the short-run dynamics of the Sraffian supermultiplier model. In order to describe the model's steady-state equilibrium, we must examine equations (7.67) and (7.71) under the conditions $\dot{h} = \dot{u} = 0$. Turning first to equation (7.67), note that the equilibrium conditions just stated imply that[41]

$$u^* = u_n \tag{7.72}$$

In other words, the steady-state equilibrium of the system is also a *fully adjusted position*. Meanwhile, imposing equilibrium conditions on equation (7.71), and bearing in mind both the result just derived (in equation 7.72) and equation (7.59), we arrive at

$$0 = u_n\left(g_A - \frac{h}{a_1}u_n\right) = u_n(g_A - g) \tag{7.73}$$

$$\Rightarrow g^* = g_A$$

Since we also know from (7.59) that $y = g$, we can therefore see that in the steady state, the rate of growth is given as

$$y^* = g^* = g_A \tag{7.74}$$

According to Serrano and Freitas (2017, p. 76), what (7.74) demonstrates is that 'the model generates an equilibrium path where economic growth is consumption-led (or, more generally, growth is led by autonomous non-capacity-creating expenditures)'. This is a somewhat roundabout way of saying that the equilibrium rate of growth (like the distribution of income) in their model is an exogenous given. In other words, recalling the nomenclature used to distinguish first- and second-generation neoclassical growth models in Chapter 1, the Sraffian supermultiplier model is an *exogenous growth model*.

Imposing equilibrium conditions on equation (7.71), and taking into account both equation (7.72) and equation (7.59), also yields the result

$$g_A = \frac{h}{a_1} u_n \tag{7.75}$$

$$\Rightarrow h^* = \frac{a_1}{u_n} g_A$$

Equation (7.75) is the steady-state value of the investment–output ratio which, given the full-capacity capital–output ratio and normal rate of capacity utilization, clearly depends on the rate of growth of non-capacity-creating autonomous demand. Equation (7.75) gives proper expression to the claim, made earlier, that it is ultimately the behaviour of $A$ that is responsible for determining the values of *both* the investment–output ratio *and* (by extension) the saving ratio. Indeed, given the formal equality of $S/Y$ and $h$ in equation (7.66) and the steady-state value of this last variable in (7.75), we can write

$$\frac{S}{Y} = \frac{a_1}{u_n} g_A \tag{7.76}$$

$$\Rightarrow f_R^* = \frac{(S/Y)}{s} = \frac{(a_1/u_n)g_A}{s}$$

where $f_R^*$ is the steady-state value of what Serrano and Freitas (2017) (following Serrano, 1995) term 'the fraction'. The steady-state value of the fraction is, according to Serrano and Freitas (2017, p. 78), the essential *closure* that, amidst its otherwise generally recognizable Keynesian structure, delineates the Sraffian supermultiplier model from its neo-Keynesian and neo-Kaleckian competitors.

These results (and their interpretation) are all very well, but can the steady-state equilibrium outcomes just described be reached? Is the Sraffian supermultiplier model *stable*, and so inclined to gravitate towards a steady-state equilibrium that is also a fully adjusted position? It transpires that this theme was an early bone of contention, the supermultiplier model not meeting with universal approval among Sraffians themselves – see, for example, Trezzini

(1995, 1998), Barbosa-Filho (2000) and Park (2000). But this criticism was a response to the early formulation of the model by Serrano (1995) and its early adoption by Bortis (1997), in which there was no formal stability analysis.[42] This lacuna has, however, since been addressed by Freitas and Serrano (2015, pp. 270–71), who demonstrate that the Sraffian supermultiplier model is stable provided that the 'expanded marginal propensity to spend' is less than one or, more specifically, that the parameter $\gamma$ – which captures the speed of adjustment of the investment–output ratio to departures of the actual rate of capacity utilization from its normal rate (see equation 7.67) – is sufficiently small. To understand all this, we begin by writing the Jacobian of the system of equations (7.67) and (7.71), evaluated at the steady-state equilibrium values $u^* = u_n$ and $h^* = (a_1/u_n)g_A$, as

$$\mathbf{J} = \begin{bmatrix} 0 & \dfrac{a_1\gamma}{u_n} g_A \\ \dfrac{-u_n^2}{a_1} & g_A\left(\dfrac{a_1\gamma}{s - (a_1/u_n)g_A} - 1\right) \end{bmatrix} \tag{7.77}$$

It follows that

$$Det(\mathbf{J}) = \gamma u_n g_A > 0$$

and

$$Tr(\mathbf{J}) = g_A\left(\frac{a_1\gamma}{s - (a_1/u_n)g_A} - 1\right)$$

Notice that in order for the system to be stable $(Tr(\mathbf{J}) < 0)$, we must observe

$$\frac{a_1\gamma}{s - (a_1/u_n)g_A} - 1 < 0 \tag{7.78}$$

$$\Rightarrow 0 < s - \frac{a_1}{u_n} g_A - a_1\gamma$$

which, bearing in mind that $s = 1 - \psi$, can be written as

$$\psi + \frac{a_1}{u_n} g_A + a_1\gamma < 1 \tag{7.79}$$

The term on the left-hand side of the stability condition (7.79) is the expanded marginal propensity to spend – that is, the responsiveness of the endogenous components of aggregate demand (consumption and investment spending) to changes in income during the model's disequilibrium adjustment process. To see this, note that as previously demonstrated, $\psi$ (the wage share of income) is also the *marginal propensity to consume* while $h^* = (a_1/u_n)g_A$ is the steady-state equilibrium value of the *marginal propensity to invest, h*. Finally, the term $a_1\gamma$ captures induced investment spending outside the steady state (see equation 7.67 and its interpretation). Together, these effects constitute an expanded marginal propensity to spend – 'expanded' in the sense that they take account of additional induced changes in (investment) spending in states of disequilibrium. Freitas and Serrano (2015, p. 271) and Serrano and Freitas (2017, p. 76) emphasize that the stability condition in (7.79) will hold for sufficiently low values of the adjustment parameter $\gamma$, which determines the size of the flexible accelerator response of investment spending to departures from the normal rate of capacity utilization. The behavioural meaning of this result is straightforward: *the investment response of firms designed specifically to restore the capacity utilization rate to its normal level must be sufficiently weak*. Note that this weak investment response is, from a behavioural perspective, anti-Harrodian – which observation is not surprising since its analytical purpose is to impose *stability* on the steady-state equilibrium of the Sraffian system (rather than introduce the local *instability* characteristic of the neo-Harrodian macrodynamics surveyed in Chapter 6).

Stability of the Sraffian supermultiplier model thus demands that we make certain assumptions about the size of the marginal propensity to spend and its component parts. These assumptions are no doubt contestable – as just illustrated, we can certainly imagine neo-Harrodian authors wishing to contest the key assumption that investment responds only relatively weakly to short-run departures of the actual from the normal rate of capacity utilization. Nevertheless, we can now state unequivocally that early reservations notwithstanding, there *are* conditions under which the Sraffian supermultiplier model is demonstrably stable.

Finally, note that by substituting the value of $h^*$ in (7.75) into the Sraffian supermultiplier in (7.65), we find that the level of output at any given point along the (stable) steady-state growth path is given as

$$Y^* = \left(\frac{1}{s - (a_1/u_n)g_A}\right)\overline{A} \tag{7.80}$$

Equation (7.80) makes clear the impact of $s = 1 - \psi = \pi$ on the level of output in the Sraffian model. Specifically, it is clear by inspection of (7.80) that an increase in the profit share (decrease in the wage share), which implies an increase in the marginal propensity to save, will lower the value of the supermultiplier and hence reduce real output, *ceteris paribus*. Since both a lower wage share and a higher propensity to save are thus seen to lower output, this provides us with a simultaneous demonstration that the level of output is wage-led and subject to the paradox of thrift. Note, however, that neither of these results is evident in the steady-state rate of growth. Indeed, since $y = g_A$, the exogenous steady-state rate of growth implied by the Sraffian supermultiplier model is completely invariant to both saving behaviour and the distribution of income. This result puts the Sraffian model at variance with almost all of the other heterodox growth models surveyed thus far in this book (classical-Marxian, neo-Robinsonian, neo-Kaleckian and neo-Harrodian), according to which the (endogenous) growth rate is sensitive to either saving behaviour or the distribution of income, or both.

### 7.5.3 Neo-Kaleckian supermultiplier models

According to Serrano and Freitas (2017, p. 71), the Sraffian supermultiplier model 'allows us to reconcile demand-led growth, exogenous distribution, and a tendency towards normal capacity utilization, even across steady states'. The first two of these attributes are identifiable features of neo-Kaleckian growth models; the addition of the third can be considered something of a holy grail in the neo-Kaleckian literature, in light of the controversy surrounding the emergence of Harrodian instability in the neo-Kaleckian model in the event that the model does not achieve a fully adjusted position. It is perhaps not surprising, then, that the Sraffian supermultiplier approach has piqued the interest of neo-Kaleckians such as Allain (2015, 2019) and Lavoie (2016) – although the fact that long-run (steady-state) growth is invariant with respect to either the saving rate or the wage share in the Sraffian supermultiplier model, as demonstrated in the previous subsection, would appear to suggest that the Sraffian and neo-Kaleckian approaches to growth are odd bedfellows.

The incorporation of a non-capacity-creating source of autonomous demand – a signature feature of the Sraffian supermultiplier model – into the neo-Kaleckian framework was discussed at some length in Chapter 6. There is no need to repeat that analysis here. It is, however, worth reflecting further on the question, implicitly raised above, as to whether or not the resulting model is properly regarded as neo-Kaleckian, given that its steady-state growth rate will not reflect either the paradox of costs or the paradox of thrift. Recall that in section 6.4.6 of Chapter 6, it was noted that according to Lavoie (2016), even if the steady-state growth rate is exogenous and thus bears no lasting influence of changes in the propensity to save or distribution of income, long-run growth can nevertheless be considered to display neo-Kaleckian properties on the grounds that growth over the course of the traverse from one steady-state equilibrium configuration to another will be positively affected by a fall in the saving propensity or a rise in the wage share.[43] For example, as was remarked in the previous chapter, following an initial reduction in the propensity to save in the neo-Kaleckian model depicted in Figure 6.7, the rate of growth rises and remains elevated during the discrete period of time that it takes for the ratio of autonomous demand to capital capacity $a^K = A/K$ to adjust to its new equilibrium value. If the long-run growth is calculated as the time-series average of the instantaneous rates of growth within every period during this traverse from one steady state to another, then the long-run growth rate so defined will be higher following a drop in the saving rate than it otherwise would have been (measured in the same manner over the same discrete period of time). According to Lavoie (2016), this serves to demonstrate that the long-run growth rate responds positively to a decline in the saving propensity – or in other words, growth is subject to the paradox of thrift.

Is this argument compelling? On the one hand, the steady-state growth outcomes of supermultiplier models show no influence of either the saving rate or the wage share, which, as we have seen, affect only the *level* of output along the steady-state growth path, not the steady-state growth rate itself. Moreover, comparison of steady-state positions (through the method of comparative dynamics) has long been the accepted method of analysis in the wage- versus profit-led growth debate. From a critic's perspective, therefore, Lavoie's (2016) argument could be regarded as 'shifting the goal posts' in a deliberate effort to produce certain desired results (the paradoxes of thrift and costs).

On the other hand, it was remarked in Chapter 1 that steady-state analysis is not the be-all and end-all of macrodynamics: it can sometimes be a straitjacket, and as we saw in Chapter 6, it is explicitly rejected in favour of emphasis on (locally unstable) disequilibrium dynamics in the neo-Harrodian tradition. Even in the macrodynamics literature that is tolerant of equilibrium analysis (including the treatment of equilibria as stable), there are warnings that equilibrium positions may never usefully describe actual (real-world) economic outcomes. According to Cornwall (1991, p. 107), 'if . . . real world change[s] in tastes, technologies and other institutional features are very rapid relative to the rate at which the economy can adjust, the convergence properties of the model take on much less interest and importance than the institutional changes themselves'. In other words, relatively slow convergence towards equilibrium may mean that despite the existence of stable equilibria, 'life is a traverse' (Halevi and Kriesler, 1992, p. 229. See also Harcourt, 1981 [1982], p. 218 and Fisher, 1983, p. 3 for similar views).[44] If it is thus possible – perhaps even necessary – to adhere only loosely to the implications, in the limit, of stable equilibrium models, then Lavoie's (2016) claims make sense. It would seem, then, that there is no necessarily correct answer to the question with which we began, which remains instead a matter of interpretation.

However, it may not be necessary to resolve the question as to whether or not Lavoie's (2016) argument is compelling in order to maintain neo-Kaleckian properties of long-run growth within a supermultiplier framework. Brochier and Macedo e Silva (2019) argue that a common problem with supermultiplier models – whether Sraffian or neo-Kaleckian – is that they are essentially flow models and lack stock-flow consistency (SFC). They therefore set out to construct an SFC supermultiplier model and to investigate whether or not the properties commonly ascribed to such models – in particular, the claims that changes in the propensity to save and the distribution of income affect only the *level* but not the (steady state) *growth rate* of income – survive the imposition of stock-flow consistency. A key feature of their model is that it involves an *autonomous* but nevertheless *endogenous* (rather than exogenous) component of aggregate demand – that is, a component of demand that is not funded by current income flows but that can, nevertheless, be related to the current value of some stock variable (and is therefore not simply taken as exogenously given). Brochier and Macedo e Silva (2019) show that, in their SFC supermultiplier framework, changes in the propensity to save and the distribution of income have both level *and* growth rate effects.

Although research into SFC supermultiplier models is in its infancy, the work of Brochier and Macedo e Silva (2019) is certainly suggestive of a means by which the central pillars of Sraffian supermultiplier analysis can be incorporated into HGT in a manner that preserves (rather than refutes) results familiar from earlier neo-Kaleckian growth models.

### 7.5.4 The turn to exogenous growth theory in heterodox macrodynamics

As we have seen, in supermultiplier models of any genus (Sraffian or Kaleckian), the driver of growth is a non-capacity-creating source of autonomous demand. Lavoie (2016, pp. 194–5) cites various time-series econometric studies that find that various types of expenditure – including exports, government spending, residential construction and components of total consumption – temporally precede and/or 'cause' changes in output. This, he argues, provides evidence that there are, in fact, various sources of non-capacity-creating expenditures that can plausibly be conceived as exogenous drivers of economic activity, as required by supermultiplier models.

Other heterodox macroeconomists are, however, less convinced. Skott (2017b, pp. 3–8) disputes the exogeneity of *any* of the sources of autonomous expenditure characteristically emphasized by supermultiplier models. He argues that exports can only be exogenous to regions but not at a global level (since the world is a closed economy);[45] luxury consumption (financed, for example, by drawing down wealth) responds to the state of the economy, while basic consumption of workers is fundamentally endogenous (the definition of a basic level of consumption rising with the level of development); residential investment responds to interest rates which are adjusted, by monetary authorities, in response to the state of the economy; and finally, government spending *may* be genuinely exogenous, but is (like monetary policy) more often responsive to the state of the economy.[46] Nikiforos (2018, pp. 668–71), meanwhile, argues that real–financial interactions are likely to defeat the autonomy of 'autonomous' expenditures, at least when the latter are debt-financed. If autonomous spending is debt-financed, then the growth of autonomous spending is equivalent to the growth of debt liabilities in the sector engaged in autonomous spending. In the steady state in supermultiplier models, output and income grow at the same rate as autonomous spending (and hence debt), meaning that the debt-to-income ratio remains constant. But suppose the economy is dislodged from the steady state by, say, a negative shock to income. This sudden spike in the debt to income ratio may cause borrowers to react (to the implied deterioration of their balance sheets) – perhaps by reducing their borrowing or even by deleveraging. Alternatively, creditors may react by limiting the supply of credit. In either event, 'autonomous expenditure stops being autonomous' (Nikiforos, 2018, p. 669) as a result of real–financial interactions, the like of which (as witnessed during the Asian financial crisis and later the Great Recession) are relatively commonplace. Finally, Dutt (2018, pp. 11–12) is more forgiving of the notion that there may be genuinely exogenous components of aggregate demand, but insists that nonetheless, compelling theoretical and/or empirical reasons are required to justify the treatment of any particular component of demand as exogenous.

Finally, it is important to draw attention back to a point made earlier: the equilibrium rate of growth in supermultiplier models is an exogenous given, meaning that what this class of models has created is a new class of *heterodox exogenous growth models* to accompany the first-generation neoclassical (Solow) model (which is also an exogenous growth model). Indeed, in the model developed by Allain (2019), the source of autonomous demand is a component of consumption spending whose rate of growth depends on the rate of population growth. This means that the economy's exogenously given steady-state rate of growth is (given the absence of technical change) equivalent to the natural rate of growth. Allain (2019) quite rightly notes that his model thus solves both of the fabled Harrod problems: Harrodian instability is tamed by the inclusion of an autonomous component of consumption demand (as demonstrated in section 6.4.6 of Chapter 6); and the first Harrod problem, highlighted in Chapter 1 (section 1.4.2) as a general feature of HGT models, never materializes. This is because the equilibrium rate of growth is the natural rate of growth – which eliminates concern with reconciling supply and demand in HGT, as first discussed in section 1.5 of Chapter 1. On the other hand, an equilibrium rate of growth that is determined by an exogenously-given natural rate is *exactly* the result of the Solow model: the rate of growth is not just exogenous, but also supply-determined!

If nothing else, developments in the direction of exogenous growth theory in HGT spurred by supermultiplier analysis can be considered a supreme irony. The recent history of neoclassical growth theory (as outlined briefly in Chapter 1) can be read as a struggle to pull away from exogenous growth theory (on the basis that it furnishes no *explanation* of growth at all, since the rate of growth is simply taken as given) and develop instead an endogenous growth theory (that *does* explain the economic origins of growth). The heterodox reaction to these developments, meanwhile, frequently involves taking umbrage at the neoclassical capture of the term 'endogenous growth', and reminding the profession that all HGT is endogenous growth theory and always has been (see, for example, Roberts and Setterfield, 2007). From this perspective, it might be argued that Sraffian-inspired developments in supermultiplier analysis have prompted a sudden, late, and undesirable turn towards exogenous growth theory in heterodox macrodynamics.

Where do these considerations leave us? Recall that according to Serrano and Freitas (2017) the Sraffian supermultiplier approach is rendered distinct among Keynesian growth models by its capacity to combine demand-led growth with an exogenous distribution of income *and* variability in the saving–output ratio. The neo-Keynesian Kaldor–Robinson models describe variability in the saving–output ratio, but do so by treating the distribution of income as an endogenous adjusting residual. In closed-economy neo-Kaleckian models, distribution is relieved of this role but the saving–output ratio is then rendered constant. In the traditional spirit of heterodox macrodynamics, however, *both* the neo-Keynesian *and* the neo-Kaleckian approaches furnish *endogenous growth models*. The problem with the Sraffian multiplier analysis, it might thus be argued, is that its combination of demand-led growth, exogenous distribution and variability in the saving ratio is achieved by another distinctive feature: the fact that it renders growth exogenous. Perhaps the key lesson that is emerging, at this late stage of our survey of models of growth and distribution, is that within this genus of models, one can never have it all.

## 7.6 Conclusions

Having largely devoted Chapters 2–5 to an exploration of the relationship between distribution and growth in core HGT models, the purpose of this chapter has been twofold. First, it has broadened the conception of distribution beyond the simple 'wages versus profits' distinction that has hitherto dominated discussion. This has been motivated in part by recognition of the fact that much of the increase in income inequality experienced in advanced capitalist economies such as the US since the 1980s has been brought about by increases in *wage* inequality,[47] while at the same time, this 'wage' inequality has been driven partly by the rapidly rising compensation of managers and executives (some or all of which is arguably part of the residual earnings or profits of the firm[48]). Concerns with these stylized facts are reflected in some of the 'three-class' models outlined in section 7.2. A second motivation for broadening the conception of distribution arises from a long-standing recognition in heterodox economics that the economy is stratified not just by social class (from which the basic wage–profit distinction arises) but also by factors such as race, ethnicity, gender and sexual orientation or identity. Thus far, HGT has had less to say about the role of these sources of social stratification in the determination of economic growth, but thinking along these lines (and the need for more of the same) is evident in the discussion of the Blecker–Seguino model in section 7.2.3 of this chapter and other references given there.

The second purpose of this chapter has been to broaden inquiry into the drivers of growth beyond the traditional concern with the distribution of income between wages and profits. One such driver is the financial sector. Traditional models of growth – whether classical, neoclassical or post-Keynesian – are notorious for their focus on the real sector of the economy, to the neglect of monetary and financial relations.[49] The inadequacy of this focus has been addressed recently by a variety of HGT models. Some – such as those analysed in section 7.3 – consider the relationship between finance and industry and its effects on growth by taking explicit account of the corporate debt and/or equity that is intrinsic to the financing of the growth process (and that, in a manner analogous to the compensation of managers discussed previously, creates another distributional claim – this time, that of a rentier class – on the total income generated by firms in the real sector of the economy). Other models consider the way in which borrowing shapes macrodynamics.[50] According to the Stockhammer–Michell model discussed in section 7.4.1, financial relations modelled in accordance with Hyman Minsky's financial instability hypothesis can drive the growth cycles that, in previous literature, have been identified with the interaction between growth and distribution. Other models, meanwhile, identify household debt, accumulated in response to the growing inequality noted above in an effort to increase consumption at a faster rate than real income, as a new driver of growth and instability in contemporary capitalism. In both cases, these models generate growth outcomes that are seemingly profit-led but, in fact, originate from channels that differ from the profit-squeeze/profit-led mechanisms characteristic of the Goodwin models (old and new) discussed in Chapters 2 and 5. In addition to moving beyond concern with the traditional growth–distribution nexus, then, these models suggest that we need to be cautious when ascribing characteristics to the growth process that appear to arise from redistribution towards profits but, in reality, may not.

A second 'new driver' of growth in contemporary HGT models is autonomous demand. Models that focus on the role of autonomous demand in the determination of steady-state equilibrium growth outcomes derive inspiration from the supermultiplier concept first introduced by Hicks (1950). As demonstrated by the discussion in section 7.5, these models now span a variety of different traditions within HGT. Much of the recent focus on supermultiplier models is attributable to the Freitas–Serrano Sraffian model discussed in section 7.5.2. But as was made clear in section 7.5.3 (and as the reader may recall from the discussion in section 6.4.6 of Chapter 6), these Sraffian developments have inspired neo-Kaleckians such as Lavoie (2016) in their search for mechanisms designed to tame Harrodian instability. The role of autonomous demand in HGT is, however, controversial, and as section 7.5.4 makes clear, supermultiplier models and their exogenous growth properties have not met with universal approval among scholars associated with the HGT project.

One class of supermultiplier models that we have yet to discuss in detail are those associated with modern Kaldorian growth theory. Kaldor's early contributions to (neo-Keynesian) growth modelling were outlined in Chapter 3. In the mid-1960s, however, Kaldor himself eschewed this earlier approach in favour of a new class of export-led growth models focused on structural change, increasing returns and cumulative causation. We now turn to a discussion of this contemporary Kaldorian tradition in Part III of this book.

## STUDY QUESTIONS

1) Does greater wage equality between production workers and corporate managers, or between female and male labour, necessarily have an expansionary impact? Compare alternative models and discuss why they reach different conclusions or under what conditions different scenarios result.
2) In what sense does monetary policy, considered as the setting of interest rates by a country's central bank, have distributional effects? Can interest rates affect the distribution of income between wages and profits, as well as between firms and rentiers, and if so how?
3) Explain how the degree to which an economy has wage-led or profit-led demand becomes endogenous in models with a financial sector, debt service, dividend payouts to equity holders, and/or interest income of rentiers.
4) Show how the interaction of the goods and financial markets can generate a Goodwin-like pattern of co-movements in the wage share and level of output. What is the significance of this result?
5) How can the introduction of household debt into a model of growth alter the relationship between distribution and growth? Why is this significant?
6) What is a supermultiplier? What is the basic implication of the supermultiplier for analysis of long-run growth?
7) Show how the Sraffian supermultiplier model gives rise to an endogenous investment to output ratio. How and why does this property contrast with other HGT models?
8) Discuss the 'turn to exogenous growth theory' in HGT. How might this help neo-Kaleckians to reconcile results such as the paradox of thrift and paradox of costs with a constant long-run equilibrium rate of capacity utilization? Why is exogenous growth theory controversial?

## NOTES

1 See Godley and Lavoie (2007) and, for recent surveys of the now-extensive literature on stock-flow-consistent macroeconomics, Caverzasi and Godin (2015) and Nikiforos and Zezza (2017).

2 Palley (2017) refers to these shares of worker and capitalist-manager households as 'ownership shares' of capital. Shares of capital ownership will be equivalent to shares of capital income (profits) if and only if the two classes of households receive equal rates of return on their assets.

3 Palley includes another term in his investment function, which in our notation can be written as $+ h_3 r$ where $r = \pi u / a_1$ is the profit rate. This essentially adds a neo-Robinsonian element into an otherwise neo-Kaleckian model, but it is not crucial to any of the qualitative results derived here.

4 Note that what is stated here pertains only to the sign of the numerator, and not the magnitude of the ratio on the right-hand side of (7.8), that is, not to the strength of the impact on the derivative. To determine the latter, it would be necessary to partially differentiate $\partial u^* / \partial \pi$ with respect to $\varphi_L$ and $\delta_L$, which yields very complex expressions given that these parameters are found in the denominator ($\Sigma^2$) as well as in the numerator.

5 Here, we differ from the approach of Tavani and Vasudevan (2014), who model gradual adjustment of $u$ to its equilibrium level, similar to the way utilization is treated in the neo-Goodwinian model of Barbosa-Filho and Taylor (2006) covered in Chapter 5. For ease of exposition and comparison with Palley's model, we instead derive a solution for short-run equilibrium $u^*$ and analyse the comparative static properties of this equilibrium.

6 Although the model presented here assumes a closed economy, we may infer that the negative effect on investment would be more likely to dominate in a highly open economy or a small country, especially if investment in tradable goods industries depends strongly on profitability in that sector (Razmi, 2016b; Ros, 2016).

7 The fact that the managers do not receive any part of the profits in Tavani and Vasudevan's model raises questions about what assets they save in and what is the saving propensity out of the returns they receive from their accumulated assets. Of course, the same issue arises in all models that allow for positive saving out of wages. See the discussion of Pasinetti (1962) and Kaldor (1966b) in Chapter 3.

**8** An alternative approach to gender modelling for an open economy is found in Seguino and Setterfield (2010), where the particular focus is on adjustment mechanisms that might ultimately reconcile the actual and natural rates of growth – and so address the interaction of aggregate demand and aggregate supply in long run growth theory discussed in Chapter 1 (section 1.5) – in an open-economy context.

**9** The assumption that the female wage is lower is made for realism and because it is important in the authors' dynamic analysis, but it does not play any role in deriving the short-run results discussed here.

**10** An alternative approach would be to define the two goods as tradables and non-tradables, and to assume that the prices of the former are given by world prices on the small country assumption. This is the approach of Razmi (2016b) and Ros (2016), and although neither of them considers gender wage gaps explicitly, they do allow for wage differentials between the two sectors. Although the modelling approach is different, the implications are qualitatively similar in at least one respect: raising wages in the tradables sector has a contractionary effect, which in the small country case occurs because higher wages squeeze profits (for any given world price) and this diminishes investment in that sector.

**11** This equation is analogous to equation (4.40) in section 4.4.3, except equation (7.18) applies only to the exported-goods sector in a two-sector model.

**12** Implicitly, there is a third curve HH for market clearing in the *H* sector, which is not drawn to avoid cluttering the diagram, and it must pass through the same equilibrium point where AD and XX intersect.

**13** In the medium-run dynamic model in Blecker and Seguino (2002), the two gendered wage rates and the exchange rate adjust endogenously. These dynamics were called 'short run' in the published article at the insistence of a referee, but are really medium run as this term is used in Chapter 5.

**14** Other scenarios are also possible, but these are the two most clear-cut cases. See Blecker and Seguino (2007) for more details as well as for the medium-run dynamics of the model in which women's wages are endogenized.

**15** The model implicitly assumes that household income is pooled, so that the same consumption and saving propensities apply to all labour income, regardless of whether women or men earn it. An important extension of this sort of model would be to consider differences in saving and spending patterns based on the gender of the income recipient.

**16** 'Rentier' comes from a French word meaning 'person of independent means' or 'annuitant', according to WordReference.com (http://www.wordreference.com/fren/rentier, accessed 11 September 2018).

**17** A Bhaduri–Marglin investment function could be used instead, with consequences as discussed below. The model of Vasudevan (2017) covered in section 7.3.2 below uses the Bhaduri–Margin version.

**18** To see this, note that the condition for $\partial(g - \sigma)/\partial u < 0$ implies a positive denominator in (7.22).

**19** See also the discussion in Chapter 6, section 6.4.6, about whether a rise in the interest rate is always and everywhere contractionary.

**20** In Lavoie and Hein's models, the distinction between the normal and puzzling cases also has implications for the stability or instability of the long-run debt dynamics, but we concern ourselves here only with the short-run comparative statics.

**21** Of course, the monetary authority may set a short-term interest rate that does not correspond to the rate paid on long-term corporate bonds, due to the existence of a term premium and possible risk premium on the latter. In that case, what matters is the degree to which the bond rate rises in response to an increase in the short-term policy rate.

**22** See section 4.3. Alternatively, if we either allowed for positive saving out of wages or utilized a Bhaduri–Marglin investment function of the form $g = h_0 + h_1\pi + h_2u - h_3id_B$, we could easily transform this model into one that allows for the possibility of profit-led demand. We leave it to the reader to solve such models as an exercise.

**23** Vasudevan (2017) also presents a model of demand with consumer debt, but for reasons of space – and since we will cover an alternative model of consumer debt later in this chapter – we will not present that model here.

**24** Perhaps Vasudevan's capitalists are top executives of the firms, who receive compensation paid out of retained profits (in some way that is not explicitly modelled) and consume part of that compensation. But to make the model more analogous to the Hein–Lavoie type of model discussed in the previous subsection, it seems better to ignore the top executives altogether and assume that any profits not paid out to the rentiers (shareholders or bondholders) are retained by the corporations, in which case 100 per cent of the retained profits must be saved.

**25** Recall that in this model, the only two sources of finance for investment are current savings and the issue of new equity; there are no bonds or bank loans.

**26** In Vasudevan (2017), aside from differences in notation and the fact that she does not assume $s_\pi = 1$, she writes this equation (and the subsequent partial derivatives) such that the signs of both the numerator and denominator are reversed (both are written to be negative, as if the top and bottom of 7.29 were both multiplied by –1). Although that is perfectly correct, we prefer to write both the numerator and denominator as positive for comparison with the other models covered here.

**27** As we will see in section 7.4.2, the (historically specific) initial distribution of income – specifically, the extent to which inequality between wage and profit income has previously accumulated – may also play an important role.

**28** See also Nikiforos (2017), who arrives independently at the key Stockhammer–Michell result: that the interaction of real–financial dynamics (the latter associated with the financial instability hypothesis) *not* growth and (profit-squeeze) distributional dynamics is what drives growth cycles in contemporary capitalism. For Nikiforos (2017), this result is important because the connection between a tight labour market and a profit-squeeze has broken down in capitalist economies since the early 1990s.

**29** For much the same reason, Stockhammer and Michell's pseudo-Goodwin cycle differs also from the quasi-Goodwin cycle associated with Skott's (1989, 2010) neo-Harrodian model in section 6.3.4 of Chapter 6: unlike the Skott model, distributional dynamics play no causal role in the Stockhammer–Michell model, but instead are a purely residual consequence of real–financial interactions.

**30** See, for example, Dutt (2005, 2006b, 2008), Kapeller and Schütz (2015) and Setterfield and Kim (2016, 2017) for developments of this kind within the class of neo-Robinsonian and neo-Kaleckian models explored in Chapters 3 and 4. Fiebiger (2018) relates the prominent role of household debt in generating aggregate macroeconomic outcomes to the work of Rosa Luxemburg, which is associated with the classical-Marxian tradition discussed in Chapter 2. Isaac and Kim (2013) construct a model of dual debt dynamics incorporating both household and corporate borrowing.

**31** The variable $D$ is defined as total debt owed by workers, so $\dot{D}$ is literally the rate of change of total debt owed by workers. However, because working households are assumed to meet interest obligations on outstanding debt in each period, $\dot{D}$ is equivalent to new borrowing by workers. Note that $\dot{D} > 0$ throughout this analysis since workers are assumed to accumulate more debt (rather than engage in deleveraging).

**32** Note that rentiers own only some part of workers' total household total debt, $D_R < D$. This is because, in the Setterfield–Kim model, workers do not consume all of their income and must, therefore, accumulate assets as a result of their saving. Setterfield and Kim (2017) assume that workers do not acquire corporate equity. As a result, workers themselves own (as creditors) some part $(D - D_R)$ of the total debt obligations that working households accumulate as they borrow.

**33** This result follows directly from equation (7.10).

**34** The expression in equation (7.48) is identical to that found in equation (37) in Setterfield and Kim (2017, p. 52), confirming that our analysis here is not lacking in generality by virtue of its exclusive focus on the consumption channel of aggregate spending.

**35** Alternatively, we could assume a fixed but positive level of investment spending is now one component of the autonomous component of demand, $A$.

**36** Note that $\omega_A$, which can be thought of as the share of autonomous aggregate demand in total output, is not related to $\omega$ (with no subscript) as used elsewhere in this chapter to represent wage inequality (the ratio of wages of managers to workers).

**37** Serrano and Freitas (2017, Table 1, p. 89) provide a detailed comparison and contrast of neo-Keynesian, neo-Kaleckian and Sraffian supermultiplier models, to which the interested reader is referred.

**38** This will be borne out when we analyse the steady-state properties of the model below.

**39** The reader is referred back to equation (3.14) which, suitably rearranged, expresses the saving ratio $S/Y$ as a function of the profit share given differing marginal propensities to save between workers and capitalists – the profit share being the key adjusting variable through which saving and investment are brought into equilibrium in the neo-Keynesian models of Kaldor and Robinson.

**40** An exception to the exogenous distribution of income is found in the open economy model in section 4.4.3. To see the point about the fixed saving–output ratio for a closed economy, recall that the neo-Kaleckian model (Kalecki–Steindl version) contains the neo-Robinsonian saving relation,

$$\sigma = s_r r$$

Multiplying on both sides by the capital stock, *K*, yields

$$S = s_r \Pi$$

$$\Rightarrow \frac{S}{Y} = s_r \pi$$

The saving–output ratio in the closed economy neo-Kaleckian (Kalecki–Steindl) model is, therefore, the product of two constants.

**41** Note that in seeking equilibrium outcomes, we ignore the trivial case where $h^* = 0$ and/or $u^* = 0$.

**42** As recounted by Lavoie (2016, p. 192), the original analysis of Serrano and Bortis rested on the assertion that if the demand expectations of firms are not systematically biased, the economy *must* eventually move towards a position characterized by equality of the actual and normal rates of capacity utilization.

**43** See also Dutt (2018, p. 12).

**44** We will return to this theme in Chapter 8 when trying to reconcile formal models of cumulative causation with Kaldor's preference for doing 'economics without equilibrium'.

**45** This point has a bearing on the interpretation of Kaldorian growth theory, which, as demonstrated in Chapter 8, has its origins in supermultiplier analysis. We will return to discuss the implications for growth theory of treating growth as export-led in the next chapter.

**46** As Skott (2017b, p. 8) argues, the Keynesian principle of functional finance due to Lerner (1943) suggests government fiscal and monetary policy *should* be endogenous in this fashion.

**47** According to Mishel and Schieder (2018), in 2017 the ratio of CEO-to-worker compensation in the US rose to 312-to-1, as compared to a ratio of 20-to-1 in 1965 and 58-to-1 in 1989.

**48** See, for example, Mohun (2006).

**49** Over 30 years ago, Kregel (1985) likened the neglect of such factors in Cambridge (neo-Keynesian) growth theory to staging 'Hamlet without the prince'.

**50** For a different way of incorporating monetary and financial factors into growth theory, see Isaac (2009).

# Appendix 7.1 Stability analysis for the Stockhammer–Michell model

Evaluating the Jacobian matrix of equations (7.32)–(7.34) at the steady-state values in equation (7.37) yields

$$\mathbf{J} = \begin{bmatrix} J_{11} & J_{12} & J_{13} \\ J_{21} & J_{22} & J_{23} \\ J_{31} & J_{32} & J_{33} \end{bmatrix} = \begin{bmatrix} -1 + \gamma_1 Y^* & \gamma_1 f^* & 0 \\ -Y^* & 1 - f^* & 0 \\ 0 & \gamma_3 \psi^* & -\gamma_4 \psi^* \end{bmatrix}$$

$$\Rightarrow \mathbf{J} = \begin{bmatrix} 0 & \gamma_1 & 0 \\ -\dfrac{1}{\gamma_1} & 0 & 0 \\ 0 & -\dfrac{\gamma_3}{\gamma_4}\left(\gamma_2 - \dfrac{\gamma_3}{\gamma_1}\right) & \gamma_2 - \dfrac{\gamma_3}{\gamma_1} \end{bmatrix}$$

According to the Routh–Hurwitz conditions for the stability of a three-dimensional system of differential equations, the matrix **J** above must possess the following properties

$$a_{\mathbf{J}1} = -Tr(\mathbf{J}) > 0$$

$$a_{\mathbf{J}2} = \begin{vmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{vmatrix} + \begin{vmatrix} J_{11} & J_{13} \\ J_{31} & J_{33} \end{vmatrix} + \begin{vmatrix} J_{22} & J_{23} \\ J_{32} & J_{33} \end{vmatrix} > 0$$

$$a_{\mathbf{J}3} = -Det(\mathbf{J}) > 0$$

$$b_{\mathbf{J}} = a_1 a_2 - a_3 > 0$$

where $J_{ij}$ represents the element in the $i$th row and $j$th column of the Jacobian matrix, **J**. Evaluating these conditions, we see that

$$a_{\mathbf{J}1} = -Tr(\mathbf{J}) = -\left(\gamma_2 - \frac{\gamma_3}{\gamma_1}\right) > 0$$

$$a_{\mathbf{J}2} = \begin{vmatrix} J_{11} & J_{12} \\ J_{21} & J_{22} \end{vmatrix} + \begin{vmatrix} J_{11} & J_{13} \\ J_{31} & J_{33} \end{vmatrix} + \begin{vmatrix} J_{22} & J_{23} \\ J_{32} & J_{33} \end{vmatrix} = 1 > 0$$

$$a_{\mathbf{J}3} = -Det(\mathbf{J}) = -\left(\gamma_2 - \frac{\gamma_3}{\gamma_1}\right) > 0$$

$$b_{\mathbf{J}} = a_1 a_2 - a_3 = 0$$

In other words, whereas the first three of the Routh–Hurwitz conditions for stability are satisfied, the fourth is not. Instead, as Stockhammer and Michell (2017, p. 114) demonstrate by means of simulations (see Figure 7.2), the system produces limit cycles in the endogenous variables $f$, $Y$ and $\psi$.

Part III

# Kaldorian approaches: export-led growth and the balance-of-payments constraint

# 8

# Export-led growth and cumulative causation

## 8.1 Introduction

This is the first of three chapters that discuss modern Kaldorian growth theory, which emphasizes the role of international trade and the balance of payments (BP) in open economies. This chapter presents models of export-led cumulative causation (ELCC), in which economic growth, technical progress, international competitiveness and export success interact in self-reinforcing 'virtuous' (or 'vicious') circles (an idea that originated with Myrdal, 1957). These virtuous and vicious circles result in patterns of chronic fast or slow growth in different economies, that can be associated with the historically observed tendency for divergence between advanced capitalist economies and the rest of the world – or, the rapid catch-up of some emerging economies and relative decline of some advanced ones.[1] The next two chapters then discuss models of balance-of-payments-constrained growth (BPCG), which focus on the constraints on growth imposed by the need to finance imports by means of export sales and/or financial inflows.[2]

ELCC and BPCG theory represent the two strands of modern Kaldorian growth theory that have been developed into the most commonly used and influential models for post-Keynesian economists concerned with long-run growth in open economies. However, Kaldorian growth theory also includes a third important strand, which is the analysis of structural change – especially in relation to the development of the manufacturing sector and the transfer of labour from agriculture to industry (or from these sectors to services), as well as the impact of global trade in manufactures and primary commodities (see Ros, 2013).[3] Each of these strands developed in the course of Kaldor's reorientation of his own thinking about growth towards a focus on the links between international trade and growth, so shifting away from his earlier closed economy analyses of the relationships between growth, distribution and technical change (including his early contribution to neo-Keynesian growth theory covered in section 3.3 of Chapter 3). Although we will not have a separate chapter on structural change, we will discuss it more briefly in all three chapters in this part of the book.

Despite important differences in their construction (which will become clear later in this chapter and the next two), the ELCC and BPCG models do coincide in certain respects: both maintain the post-Keynesian belief that aggregate demand constraints are paramount in determining a nation's output, even in the long run; and both see those constraints as emanating primarily from the international domain (rather than the domestic economy), so that increasing the growth rate of exports is ultimately key to raising a country's long-run growth rate.[4] As such, Chapters 9 and 10 will address some of the ways in which ELCC and BPCG theories can be reconciled. The main focus of this chapter, meanwhile, is to develop a model of ELCC that emphasizes the two fundamental properties of growth according to this model: that it is demand-led (with international trade playing the key role in generating the growth of autonomous demand); and that it is path dependent. In ELCC theory, path dependence is found in both the actual and the (Harrodian) natural rates of growth, and ultimately involves the economy evolving through an historically-specific series of technologically and/or institutionally specific regimes or episodes of growth.

The chapter is organized as follows. Section 8.2 outlines the core elements of the basic vision of growth developed by Nicholas Kaldor following his inaugural lecture at the University of Cambridge (Kaldor, 1966a [1989]). These elements include Kaldor's growth laws, the emphasis on structural change and the notion that exports are the ultimate source of autonomous demand that drives the demand-led growth process. Section 8.3 introduces the concept of cumulative causation in the growth process and presents the canonical formal model of ELCC developed by Dixon and Thirlwall (1975) and Setterfield and Cornwall (2002). Section 8.4 then discusses path dependence in the actual rate of growth along with Cornwall and Cornwall's (2001) conception of 'evolutionary Keynesian' macrodynamics. Particular importance is attached in this discussion to the recursive interaction of institutions, demand conditions and growth outcomes. Section 8.5 discusses path dependence in the natural rate of growth, drawing attention back to the response of supply conditions to demand conditions that is a basic feature of the Kaldorian vision of growth. Section 8.6 discusses some policy implications of the ELCC model. Section 8.7 examines various critiques of the ELCC model and alternative formulations of export-led growth, including the debate over 'Kaldor's paradox', the approach of Beckerman (1962) and the importance (or lack of importance) of relative prices and costs in driving exports. This section also discusses the lack of attention to imports and the BP in the ELCC model, thus setting the stage for the presentation of the BPCG model in Chapter 9. Finally, section 8.8 offers some conclusions.

## 8.2 The Kaldorian vision of growth

### 8.2.1 Kaldor's growth laws

Modern Kaldorian growth theory builds on the growth schema found in Nicholas Kaldor's post-1966 writings on growth and the importance of history rather than equilibrium in the economic processes (see, for example, Kaldor, 1966a [1989], 1970, 1972, 1985, 1996). Kaldor came to believe that the analysis of economic growth should be founded on a series of empirical generalizations or 'stylized facts', which have come to be known as 'Kaldor's growth laws'. Five of these laws, as summarized by Thirlwall (1983, emphasis in original) but renumbered here, are most relevant to the present discussion:

> 1) The faster the rate of growth of the manufacturing sector, the faster will be the rate of growth of Gross Domestic Product (GDP) . . ..
> 2) The faster the rate of growth of manufacturing output, the faster will be the rate of growth of labor productivity in manufacturing owing to static and dynamic economies of scale, or increasing returns in the widest sense . . ..
> 3) The faster the growth of manufacturing output, the faster the rate of labor transference from nonmanufacturing to manufacturing, so that overall productivity growth is positively related to the growth of output and employment in manufacturing and negatively associated with the growth of employment outside manufacturing.
> 4) The growth of manufacturing output is *not* constrained by labor supply but is fundamentally determined by demand from agriculture in the early stage of development and exports in the later stages . . ..
> 5) A fast rate of growth of exports and output will tend to set up a cumulative process, or virtuous circle of growth, through the link between output growth and productivity growth.[5]

Of these propositions, law 2 is based on Verdoorn's (1949) empirical finding of a positive correlation between the growth rates of labour productivity and output in manufacturing across countries, and thus is often referred to as 'Verdoorn's law'.[6]

These growth laws make clear the importance of *economic structure* (not just aggregate demand and path dependence) – and in particular, the prominence of the manufacturing sector – in Kaldor's vision of the growth process (see also Thirlwall, 2013). The ELCC model is typically developed with reference to total output as an aggregate, one-sector model, however, and so does not explicitly reflect the notion of there being a sectoral 'engine' of growth in the manufacturing sector. The importance of manufacturing in the composition of final output and export sales is frequently alluded to in the construction of ELCC models, but in principle the model should be disaggregated to capture this aspect of capitalist growth.[7]

## 8.2.2 Structural change and productivity growth

Kaldor formalized the first three of his growth laws with linear equations that he and others used to test the hypotheses econometrically using international cross-sectional data (this was in the 1960s and 1970s, before panel data estimation was widely available). Since our purpose here is to exposit Kaldor's theory, rather than to examine the econometric pitfalls in how he originally specified his equations, we will not enter into the debate about how they should be estimated empirically (see Thirlwall, 1983); rather, we will simply state them as algebraic expressions without the numerically estimated coefficients. Suppressing country subscripts and ignoring random error terms for brevity, these equations can be written as follows (where the $b_{ij}$ are all positive coefficients, with $i = 1, 2, 3$ indicating the first, second or third law and $j$ indicating the coefficient):

1) The growth rate of total output ($y$) is an increasing function of the growth rate of manufacturing output ($y_m$)

$$y = b_{10} + b_{11}y_m \tag{8.1}$$

2) The growth rate of labour productivity in manufacturing ($q_m$) is an increasing function of the growth rate of manufacturing output (this is the original version of Verdoorn's law)

$$q_m = b_{20} + b_{21}y_m \tag{8.2}$$

3) Overall (economy-wide average) productivity growth is an increasing function of the growth of output and employment in manufacturing and negatively related to employment growth in non-manufacturing

$$q = b_{30} + b_{31}y_m \tag{8.3}$$

$$q = b_{32} + b_{33}l_m - b_{34}l_n \tag{8.4}$$

where $q$ is economy-wide productivity growth (as in previous chapters) and $l_m$ and $l_n$ are the growth rates of employment in manufacturing and non-manufacturing respectively. The difference between $l_m$ and $l_n$ reflects structural change in the form of the transfer of labour between non-manufacturing and manufacturing activities.

The emphasis on manufacturing in Kaldor's analysis of structural change and economic development is based on four major advantages of manufacturing compared to other economic activities:[8]

1) Labour productivity (value added per worker) is typically higher and often grows more rapidly in manufacturing than in agriculture or services. As long as the level of productivity is higher in manufacturing, increasing the percentage of the labour force in manufactures raises the average productivity of labour for the economy as a whole – a phenomenon that is known as the 'structural change bonus'. Moreover, manufacturing is usually more capital-intensive than other sectors, which means that it offers greater opportunities for capital accumulation (investment), which in turn raises productivity further. However, some other industries are also highly capital-intensive (mining, utilities, transportation), including parts of agriculture in some advanced economies such as the US. Szirmai (2012) finds that manufacturing productivity has not grown more rapidly than agricultural productivity in most countries since 1975, even though it did in earlier historical periods, but manufacturing output continues to grow more rapidly than agricultural output and the potential for a structural change bonus still exists (since manufacturing industries continue to offer relatively high levels of value added per worker).
2) Manufacturing industries tend to exhibit greater opportunities for economies of scale or increasing returns, due to the high fixed costs of machinery and equipment as well as R&D (research and development). Indeed, the manufacturing sector has long been the main locus for technological progress, as most innovations (new products and processes) tend to originate in manufacturing or require manufactured inputs (for example, new machinery, computer hardware, or agricultural and mining equipment used in other sectors). Developing countries can adopt existing advanced technologies (via 'technology transfer' or 'international diffusion') relatively easily in manufacturing, which may help to account for the fact that there is unconditional convergence of productivity to the global frontier in manufacturing industries even though unconditional convergence is not observed for total, overall productivity (Rodrik, 2013). However, this too is changing, as many

services now require innovative activity (for example, software design in information technology) and have more opportunities for technological catch-up.

3) Manufacturing industries are also thought to offer the potential for strong backward and forward linkages in the development process. Backward linkages are the demand created for 'upstream' or supplier industries (raw materials, parts and components, machinery and equipment), while forward linkages are the supplies of goods to 'downstream' or user industries (assembly, final manufacture, transportation, distribution, other services and so on). However, it should be noted that it is more difficult to contain these linkages within individual countries today than it was in the past, because production is now broken up into 'global supply [value] chains' in which different parts of the production process (R&D, production of parts and components, assembly and so on) are performed in different countries. On the positive side, manufactures still offer what might be called horizontal linkages or positive externalities in production, since the skills, knowledge and training (of labour, management and so on) acquired in a given manufacturing firm or industry are often transferrable to other firms, occupations or sectors. These positive spillover effects help to account for the frequently observed clustering of activities in particular locations (such as apparel production in Bangladesh or automotive production in north-central Mexico).

4) Last, but very importantly from a Kaldorian perspective, manufacturing also offers significant advantages on the demand side. Engel's law implies that, as per capita income rises, households spend a decreasing share of their income on food and a higher share on manufactured goods. As a result, the income elasticity of demand is higher for manufactures than for food, which implies greater potential for long-run growth in manufacturing production than in agriculture (as we will see in Chapter 9, this also implies that manufacturing exporters will have an easier time relaxing BP constraints on their growth). In addition, the terms of trade (relative prices) of primary commodities (energy goods and other mineral products as well as agricultural raw materials) are notoriously volatile and sensitive to global business cycles, even if it is debatable whether they have a long-term decreasing tendency. Of course, some types of modern services also have desirable demand characteristics, including high income elasticities (for example, information technology, computer software and entertainment). But the benefits of a manufacturing specialization relative to an agricultural one are still found in the fact that manufacturing output tends to grow faster than agricultural output, even if productivity no longer grows faster in manufacturing than in agriculture (Szirmai, 2012).

Although it is far from the first such effort,[9] Rodrik (2014) provides a convenient way of formalizing the productivity benefits of structural change oriented towards the expansion of the manufacturing sector that links up to contemporary empirical studies. Following earlier scholars of structural change such as Baumol et al. (1989), Rodrik allows that there are some 'modern' services that have relatively high productivity, while services in the 'traditional' sector have lower productivity (along with agriculture, which for simplicity is assumed to be entirely traditional). Thus, suppose an economy has three sectors: manufacturing ($m$), modern services ($s$) and traditional services and agriculture ($t$); the last one could also include the so-called informal sector. Average labour productivity for the entire economy ($Q = 1/a_0$, where $a_0$ is the aggregate labour–output coefficient from earlier chapters) is a weighted average of productivity in these three sectors (similarly defined for each sector)

$$\begin{aligned} Q &= \alpha_m Q_m + \alpha_s Q_s + \alpha_t Q_t \\ &= \alpha_m Q_m + \alpha_s Q_s + (1 - \alpha_m - \alpha_s) \end{aligned} \tag{8.5}$$

where $\alpha_i$ is the share of labour in sector $i$, $\alpha_m + \alpha_s + \alpha_t = 1$, and we assume $Q_t = 1$ (thus, output is normalized such that it equals the product of one $t$-sector worker, which is held constant).

Let the relative productivity of each sector ($i$) compared to average productivity be denoted by $\pi_i = Q_i/Q$, and recall that productivity growth is $q = \dot{Q}/Q$. Rodrik (2014) assumes that productivity in modern services grows through conditional convergence of average productivity to a steady-state level $Q^*$ that depends only on institutional and policy 'fundamentals' ($\Theta$)[10]

$$q_s = \xi[\ln Q^*(\Theta) - \ln Q] \tag{8.6}$$

where $\xi$ measures the speed of convergence conditional on the set of fundamentals $\Theta$. He also assumes, based on econometric evidence in Rodrik (2013), that productivity growth in manufacturing exhibits unconditional convergence: it is faster in countries or industries that are farther behind the global frontier, regardless of institutions, policies or other fundamentals (although the latter can also help). Thus manufacturing productivity grows according to

$$q_m = \beta(\ln Q_m^* - \ln Q_m) + \xi[\ln Q^*(\Theta) - \ln Q] \tag{8.7}$$

where $Q_m^*$ is manufacturing productivity in the leading advanced economies and $\beta > 0$ is a coefficient reflecting the speed of unconditional convergence in this sector. Productivity is assumed to remain constant in the traditional sector, where $Q_t = 1$, so $q_t = 0$.

On the basis of these assumptions, the average growth rate of productivity in the economy as a whole can be expressed as the sum of four parts

$$\begin{aligned} q = {} & (\alpha_m \pi_m + \alpha_s \pi_s) \cdot \xi [\ln Q^*(\Theta) - \ln Q)] && \text{(a)} \\ & + \alpha_m \pi_m \cdot \beta (\ln Q^*_m - \ln Q_m) && \text{(b)} \\ & + (\pi_m - \pi_t) d\alpha_m && \text{(c)} \\ & + (\pi_s - \pi_t) d\alpha_s && \text{(d)} \end{aligned} \tag{8.8}$$

where (a) = conventional conditional convergence based on fundamentals, (b) = unconditional convergence in manufactures, (c) = the impact of the transfer of labour from the traditional sector to manufacturing and (d) = the effect of the transfer of labour from the traditional sector to modern services.

Regarding (a), a neoclassical economist might consider that the fundamentals would include things like saving propensities, rule of law and 'sound' macro policies (low inflation, small budget deficits and so on), while a heterodox (Kaldorian/Schumpeterian) economist would be more likely to emphasize factors such as a country's capabilities for technological innovation (or diffusion), public investment in infrastructure, and the administrative effectiveness of the state (probably both could agree on the importance of education).[11] In fact, Roberts (2007) has shown that a Kaldorian ELCC model also has conditional convergence properties. But from a Kaldorian perspective, the two key elements in Rodrik's model are (b) and (c): (b) implies that having a *higher* share of labour in manufactures ($a_M$) increases the ability of a nation to take advantage of the unique opportunities afforded by manufacturing activities for promoting productivity growth, while (c) implies that *increasing* that share ($da_M > 0$) is beneficial to the extent that productivity in manufactures exceeds the level in the traditional sector. The benefit (c) is the structural change bonus, which implies that transferring workers from low-productivity traditional or informal activities to high-productivity manufactures raises average productivity, independently of the growth rates of productivity in each sector. And, to the extent that modern services also offer a positive relative productivity differential ($\pi_s - \pi_t$), then (d) implies that a rising share of labour in this sector can yield a similar benefit – provided that the labour comes out of the low-productivity traditional sector and not out of higher-productivity manufacturing.

Of course, just as a specialization in manufacturing industries can be beneficial, a move away from manufacturing towards services can be problematic unless the latter are also of a high-productivity, technologically innovative variety. Indeed, Kaldor's original formulation of his laws was motivated by the slow growth of the British economy compared with other countries (especially Japan and the former West Germany) in the 1960s and 1970s, and the deindustrialization of the UK was a key part of his explanation. More recently, it has become apparent that many developing countries are suffering 'premature deindustrialization', in the sense that their manufacturing shares of output and employment have started to decline at lower levels of real per capita income compared with the earlier experience of the advanced economies. Rodrik (2016) attributes this phenomenon to the way that globalization and trade liberalization have induced many developing countries to re-specialize in primary commodities. The developing regions that have experienced the most premature deindustrialization (especially South America) have grown more slowly relative to those that have not (especially the East Asian economies), which is exactly what we would expect from the Kaldorian approach to structural change. We will return to the importance of structural change in Chapters 9 and 10, but now we turn to our main themes in this chapter related to ELCC in an aggregative framework.

### 8.2.3 Export-driven, demand-led growth

Kaldor's basic vision of growth is based on a dynamic two-way interaction between supply and demand conditions. A central feature of this vision is Adam Smith's (1776 [1976], p. 21) famous dictum 'that the division of labour is limited by the extent of the market'. In other words, the expansion of demand induces changes in the potential supply of goods, by affecting the efficiency with which goods are produced. For Smith, the extent of the market had a decisive effect on specialization in the process of production (the division of labour). In modern Kaldorian growth theory, and in keeping with the second of Kaldor's growth laws as stated in section 8.2.1, Verdoorn's law – the empirical finding of a positive correlation between the growth rates of labour productivity and output in manufacturing – is understood as a dynamic analogue of Smith's original dictum, describing the workings of a demand-led economy that is subject to *dynamic increasing returns to scale*.[12] On this view, Verdoorn's law – which is otherwise something of a 'black box' – captures the influence of output growth not only on specialization in the production process à la Smith, but also on the extent of learning by doing at the point of production (Arrow, 1962); the propensity of firms to engage in R&D, the payoffs to which are highly uncertain (Schmookler, 1966; Brouwer and Kleinknecht, 1999); firms' willingness to invest in 'lumpy' physical capital that embodies technological improvements that would be chronically underutilized absent a sufficiently large market for final output (Scitovsky, 1956; Lamfalussy, 1961, 1963);[13] and, in dual economies, the evolution of the shares of employment in the traded goods and informal sectors that differ with respect to their levels and rates of growth of output per person (Dutt and Ros, 2007). It will immediately be recognized that all of these factors can be connected to productivity growth.

But if Verdoorn's law represents the influence of demand on supply conditions, how, if at all, do supply conditions influence demand? Here, Kaldor drew inspiration from Young's (1928) attempt to address this question by combining Smith's insights with Say's law. Young argues that, at a very aggregate level, economic activity essentially involves the exchange of goods for goods. Hence every increase in the supply of a commodity increases the market for other commodities (at least potentially – a vital qualification for Keynesians!). This leads Young to conclude that just as the division of labour depends on the extent of the market, so, too, the extent of the market depends on the division of labour, since rising output resulting from increased specialization will increase the market for other goods and services. Kaldor, however, regarded demand as being relatively autonomous of supply conditions – *influenced* but not *determined by* supply conditions, as in Keynes's principle of effective demand.[14] This Keynesian conception of demand formation, in which Say's law does not hold, privileges the causal role of demand in the two-way interaction between demand and supply conditions that otherwise emerges from the joint interaction of the division of labour and the extent of the market discussed so far. It suggests that despite their joint interaction, the supply side is more likely to play a passively adjusting and accommodating role, and the demand side an independently active role, in generating changes in growth outcomes. This makes demand formation the focus of growth analysis from a Kaldorian perspective, in which growth is thus conceived as an essentially demand-led process.

Kaldor placed particular emphasis on external demand (that is, exports) as the key source of the expansion of aggregate demand in the demand-led growth process.[15] Indeed, for Kaldor, the expansion of exports is the proximate source of growth, so that the basic 'equation of motion' in growth theory is

$$y = k_X x \tag{8.9}$$

where $y$ is the rate of growth of real output, $x$ is the rate of growth of real exports and $k_X$ is the dynamic foreign trade multiplier.

## 8.2.4 A fallacy of composition?

Note that if equation (8.9) were to imply that all growing economies must run BP surpluses, it would suffer a simple fallacy of composition. It would lack generality as a description of capitalist growth, because not all economies can simultaneously accumulate trade surpluses. However, equation (8.9) does *not* necessarily have this implication. To see this, consider the foundations of equation (8.9) based on the following simple static model of output determination (for an open economy, but with no government for simplicity)

$$Y = C + I + (X - M) \tag{8.10}$$

$$C = cY \tag{8.11}$$

$$I = a_1 \Delta Y = a_1 y Y \tag{8.12}$$

$$M = vY \tag{8.13}$$

where $Y$ is real output, $C$, $I$, $X$ and $M$ are (respectively) consumption, investment, exports and imports (all in real terms), and $c$, $a_1$ and $v$ are (respectively) the propensity to consume, the ratio of capital to full-capacity output, and the marginal propensity to import. The structure of this model is consistent with Kaldor's (1970) insistence that, ultimately, exports are the only truly autonomous source of demand: both consumption and investment are wholly endogenous to income.[16]

Solution of (8.10)–(8.13) yields

$$Y = \frac{1}{1 - (c + a_1 y) + v} X \tag{8.14}$$

Note immediately that equation (8.14) is a species of *supermultiplier* analysis, so that as remarked in Chapter 7, modern Kaldorian growth theory shares a methodological affinity with the approach to growth associated with the recent work of some Sraffians and Kaleckians. As will be discussed below, however, the unique exogenous driver of growth in Kaldorian models is *export* growth (or the fundamental determinants thereof) in a model of strictly *regional* growth. Exports (or their fundamental determinants) can, of course, satisfactorily be taken as exogenously given at the level of the individual region, without this presupposing anything about the nature of the growth process *globally*. Kaldorian growth models are not, therefore, identical to other supermultiplier-based growth models that seek to furnish explanations of global growth and, in so doing, seem to be creating a new species of exogenous growth theory comparable to first-generation neoclassical growth theory (NGT), as was discussed in Chapter 7. This property of Kaldorian models is in keeping with Kaldor's (1972, 1985) own vision of the long-run growth process, according to which it is ultimately unsatisfactory to regard anything other than the legacy of the past as exogenously given.[17]

Suppose we now assume that $1 - c = a_1 y \Rightarrow 1 - (c + a_1 y) = 0$, where in the present context of discrete changes the growth rate of output is $y = \Delta Y/Y$. Notice that, from equation (8.11), $1 - c = S/Y$ where $S = Y - C$ denotes aggregate saving, while, from equation (8.12), $a_1 y = I/Y$. In other words, we are now assuming that the savings–income and investment–income ratios are always equal, or in the language of the sectoral balances accounting framework of Wynne Godley (on which see, for example, Wray, 2012, pp. 15–20), the private sector runs neither a surplus nor a deficit (the savings of the household sector fully funds the investment spending undertaken by the corporate sector). This assumption is again consistent with Kaldor's thinking about the operation of capitalist economies.[18] Under these conditions, the solution to (8.10)–(8.13) above reduces to

$$Y = \frac{1}{v} X \tag{8.15}$$

where $1/v$ – the remains of the supermultiplier originally found in (8.14) – is the (original, static) Harrod foreign trade multiplier.[19] Finally, it follows from (8.13) and (8.15) that[20]

$$\dot{M} = v\dot{Y} \tag{8.16}$$

and

$$\dot{Y} = \frac{1}{v} \dot{X} \tag{8.17}$$

The latter expression (8.17) is a dynamic version of Harrod's foreign trade multiplier, similar to (8.15) but written in terms of changes (time derivatives) instead of levels of the variables.

If we now combine (8.16) and (8.17) we can see that

$$\dot{M} = v\frac{1}{v}\dot{X} = \dot{X} \tag{8.18}$$

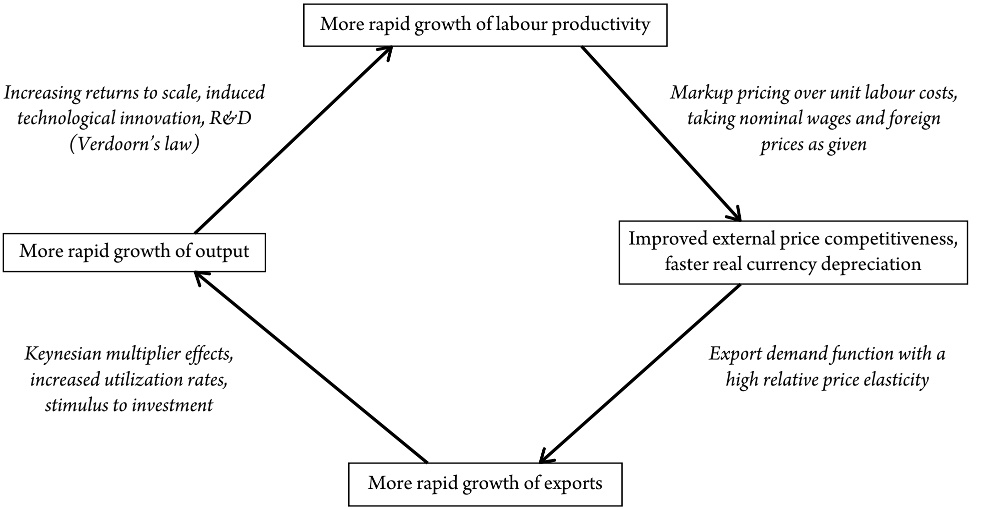
In other words, starting from a position of external balance ($X = M$), any expansion of output due to an expansion of exports ($\dot{X} > 0$) will automatically be consistent with the *maintenance of external balance*, since $\dot{M} = \dot{X}$. In short, the notion that export-led growth (as in equation 8.9) necessarily suffers a fallacy of composition – in the sense that not all countries can pursue export-led growth simultaneously – is false. This result is intuitive. It holds for the same reason that an increase in the size of Firm A does not necessarily come at the expense of Firm B in the context of the domestic economy: both firms can expand simultaneously as a result of a *general expansion of trade*.[21] At the same time, however, it should be noted that the simple logic of this result may not materialize in practice. In reality, countries that pursue export-led growth and do so by targeting the domestic markets of other (usually larger and more developed) economies are undertaking a strategy that *cannot* be pursued by all countries simultaneously, and that does, therefore, suffer a fallacy of composition (see, for example, Blecker and Razmi, 2010).

The analysis above draws attention to fact that, rather than suffering an inevitable fallacy of composition, Kaldorian models of export-driven, demand-led growth are nevertheless models of *regional* growth, designed to describe outcomes in a single region (which could be an individual nation state, a bloc of nations such as the European Union, or even a subnational unit) that engages in inter-regional trade. Hence in the ELCC and BPCG models discussed in this and the next two chapters, the rate of growth of world income is the ultimate (exogenously given) determinant of growth at the level of the region, usually interpreted as an individual nation state. This variable itself demands explanation by growth theory, but necessarily remains beyond the scope of models of regional (or national) growth. We can therefore think of the question of *global* growth as being beyond the purview of both ELCC and BPCG theory. This, in turn, means that the Kaldorian models discussed in these chapters are not ultimately antagonistic to some of the other models discussed earlier in this book, which by virtue of their closed economy structure *can* claim to offer explanations of global growth and hence the rate of growth of world income that Kaldorian models take as given in order to furnish explanations of regional (national) growth at a different (lower) level of aggregation.

## 8.3 Cumulative causation and the Dixon–Thirlwall model

### 8.3.1 Cumulative causation and the path dependence of the growth process

For Kaldor (as for Young and Myrdal before him), the two-way interaction between demand and supply conditions that has been discussed above is




Notes: The text boxes show the changes in variables; the arrows and italicized phrases show the causal mechanisms. Verdoorn's law operates mainly in the manufacturing sector, as discussed in the text. What are depicted here are positive changes reflecting a 'virtuous circle' of causality; negative changes would reflect a 'vicious cycle'.

**Figure 8.1** Export-led growth with cumulative causation: a schematic representation

properly interpreted as a process of *cumulative causation* – that is, a self-reinforcing, causal-recursive process as a result of which initial success in the growth process begets subsequent success (and failure begets failure). Specifically, in the Kaldorian schema, initially rapid output growth induces dynamic increasing returns (via Verdoorn's law), which enhances export competitiveness and hence export growth, which results in further rapid output growth (via equation 8.9), and so on. Figure 8.1 illustrates the basic logic of this schema, showing the 'circular and cumulative causation' between export growth, output growth, productivity growth and international competitiveness (measured by the rate of increase in the real exchange rate or relative price of foreign goods). Figure 8.1 represents a virtuous circle of self-reinforcing rapid growth; a vicious circle would involve cumulatively slow growth arising from slow initial output growth inducing weak productivity performance and declining international price competitiveness, resulting in slow export growth and hence a slow rate of export-driven, demand-led output growth, and so on. One note of caution is that the graphical representation in Figure 8.1 ignores the special role accorded to the manufacturing sector in Kaldor's growth laws, as discussed in the previous section. Instead, it focuses only on aggregate output, productivity, and exports, an implicit assumption being that manufactured goods are a significant component of aggregate output and exports.

In the Kaldorian growth schema outlined above, growth is certainly endogenous in the 'narrow' sense identified by Roberts and Setterfield (2007). First, technical change is explicitly modelled (in the form of Verdoorn's law). Second, instead of being imposed upon the system from without (as, for example, in the first-generation-NGT Solow model), the observed or actual rate of growth arises from causal interactions within the schema itself.[22] But Kaldor's growth schema is also consistent with Roberts and Setterfield's 'deeper' conception of endogenous growth, in which the growth rate today is sensitive to the pace of growth in the past. In other words, growth is *endogenous to its own past history*, or is *path dependent*. The importance of this theme to Kaldor is made forcefully evident in the following quotation:

> it is impossible to assume the constancy of anything *over* time, such as the supply of labour or capital, the psychological preferences for commodities, the nature and number of commodities, or technical knowledge. All these things are in a continuous process of change but the forces that make for change are endogenous not exogenous to the system. The only truly exogenous factor is *whatever exists at a given moment of time*, as a heritage of the past. (Kaldor, 1985, p. 61; emphasis in original)

Along with the importance of trade for aggregate demand formation, the notion of growth as an historical or path-dependent process has also informed much of the Kaldorian literature that has built on Kaldor's growth schema. This will become clear in the development and discussion of the ELCC model that follows.

## 8.3.2 The Dixon–Thirlwall model of export-led growth

Several economists have developed explicit models of export-led growth seeking to incorporate the idea of cumulative causation. Early contributions that foreshadow the emergence of Kaldorian models include Beckerman (1962) and Lamfalussy (1963).[23] An early model that builds on explicitly Kaldorian foundations was developed by Cornwall (1977). But the most widely known and accepted version, which we can therefore refer to as the canonical formal model of Kaldor's growth schema for a 'representative' capitalist economy, is due to Dixon and Thirlwall (1975). The Dixon–Thirlwall model can be stated as follows[24]

$$y = k_X x \tag{8.9}$$

$$x = \varepsilon_X(\hat{P}_f - \hat{P}) + \eta_X y_f \tag{8.19}$$

$$\hat{P} = \hat{W} - q \tag{8.20}$$

$$q = q_0 + \rho y \tag{8.21}$$

where $\hat{P}$ is the rate of price inflation, $\hat{W}$ is the rate of growth of nominal wages, $q$ is the rate of productivity growth, $\varepsilon_X > 0$ is the price elasticity of demand for exports,[25] $\eta_X > 0$ is the income elasticity of demand for exports, the subscript $f$ denotes a 'foreign' (rest-of-world) variable, and all other variables are as previously defined.

Equation (8.9) is already familiar from our discussion of the basic principles of export-driven, demand-led growth in section 8.2. Equation (8.19) describes the rate of growth of exports in terms of the inflation differential $(\hat{P}_f - \hat{P})$, which measures the rate of change of the relative price of foreign goods,[26] and the rate of growth of income in the rest of the world. It can be derived from an export demand function written in the Cobb–Douglas (constant elasticity) form

$$X = X_0 \left(\frac{EP_f}{P}\right)^{\varepsilon_X} Y_f^{\eta_X} \tag{8.22}$$

where $P$ denotes the domestic price level, $E$ is the nominal exchange rate (the domestic currency price of foreign currency, which is assumed fixed for simplicity), $X_0$ is a positive constant, $P_f$ and $Y_f$ are the foreign price level and national income, respectively, and the other variables are as previously defined. This expression – and hence its dynamic analogue in equation (8.19) – relates the demand for exports to relative prices and the level of income, and in this respect it represents a quite conventional description of the demand for goods.[27] Equation (8.20), which describes the rate of inflation, follows from a standard markup pricing equation where prices are set as a (fixed) markup over unit labour costs, the latter determined as the ratio of the nominal wage to the average product of labour (or level of labour productivity).[28] Finally, equation (8.21) represents an aggregative version of Verdoorn's law, which is (once again) already familiar from the discussion in section 8.2 (and ignoring the focus on the manufacturing sector noted there).[29] The parameter $q_0$ captures exogenous influences on productivity growth, while $\rho$ – the 'Verdoorn coefficient' – measures the elasticity of productivity with respect to real output, and hence the extent to which a change in the rate of growth of output induces a change in the rate of growth of labour productivity.

Combining equations (8.9), (8.19) and (8.20) yields

$$y = k_X(\varepsilon_X[\hat{P}_f - \hat{W} + q] + \eta_X y_f) \tag{8.23}$$

Suppose we now use the following equations to describe inflation and productivity growth in the rest of the world

$$\hat{P}_f = \hat{W}_f - q_f \tag{8.20$'$}$$

$$q_f = q_0 + \rho_f y_f \tag{8.21$'$}$$

where we assume, for simplicity, that the exogenous component of productivity growth, $q_0$, is identical in all regions. In other words, both inflation and productivity growth in the rest of the world are determined in exactly the same fashion as they are in our representative economy. Assume further that nominal wages increase at the same rate at home and abroad: $\hat{W} = \hat{W}_f$.[30] On the basis of these assumptions, equation (8.23) can be rewritten as

$$y = \Omega + k_X \varepsilon_X q \tag{8.24}$$

where $\Omega = k_X([\eta_X - \rho_f \varepsilon_X]y_f - \varepsilon_X q_0)$.

The term $\Omega$ in equation (8.24) is no more than a constant composite coefficient, but despite its intimidating appearance, it does have a straightforward interpretation in terms of the economics of demand formation and growth. To see this, first observe that the term in parentheses on the right-hand side of the expression for $\Omega$ above can be rewritten as $\eta_X y_f - \varepsilon_X(q_0 + \rho_f y_f)$. The first term in this last expression, $\eta_X y_f$, captures the positive effect on demand formation and growth in our representative economy of foreign income growth, $y_f$, working through the income elasticity of demand for exports $\eta_X$. The second term, meanwhile, captures the negative effect on demand formation and growth of $y_f$ working through the price elasticity of demand for the exports of our representative economy, $\varepsilon_X$, via the impact of $y_f$ on *foreign* productivity growth and hence foreign relative price competitiveness in international trade. Either of these effects can, in principle, be dominant, without affecting the unequivocally positive relationship between $q$ and $y$ captured by the second term in equation (8.24), which (at least in the first instance) is the primary focus of our attention.

Following Setterfield and Cornwall (2002), we can now identify Verdoorn's law (equation 8.21) as the *productivity regime* (PR) of the ELCC model, describing how productivity growth is determined through (among other things) growth-induced technical progress, and equation (8.24) as the *demand regime* (DR), which describes the dynamics of demand formation. Equation (8.24) summarizes a process of demand formation that includes the influence of productivity growth on domestic inflation (in equation 8.20) and hence export growth (in equation 8.19) and hence output growth (in equation 8.9) – thus capturing the feedback effect of productivity growth on demand-led output growth and so establishing the influence of supply conditions on aggregate demand as supposed by Young (1928).[31] But the dynamics of demand formation are not *limited* to this influence of supply on demand (thanks to the role of $\Omega$), thus establishing the relative autonomy of aggregate demand from aggregate supply presupposed by Kaldor (following Keynes's principle of effective demand).

Together, the productivity and demand regimes outlined above describe the recursive interaction of aggregate demand and aggregate supply in the determination of the growth rate, as envisaged by Kaldor in his discussions of the process of cumulative causation. This is illustrated in Figure 8.2, in which $y^*$ and $q^*$ denote the equilibrium rates of growth of output and productivity, respectively, and where it is assumed that $\Omega > 0 > -q_0/\rho$ and $1/\rho > k_X\varepsilon_X \Rightarrow k_X\rho\varepsilon_X < 1$. The significance of the first of these conditions is clear by inspection of Figure 8.2; the second implies that, as they are presented in Figure 8.2, PR is steeper than DR. Together, these conditions are sufficient to ensure the stability of the growth equilibrium depicted in Figure 8.2 at economically meaningful (positive) values of the growth rates of output and productivity, $y$ and $q$. This is captured in Figure 8.2 by the values of $y^*$, $q^* > 0$, coupled with the observation that if we begin with any value of $q$ that is lower (higher) than $q^*$, the resulting rate of growth (read off DR) will cause a subsequent increase (decrease) in $q$ due to movement along PR, which will induce a rise (fall) in $y$ due to movement along DR and so on,

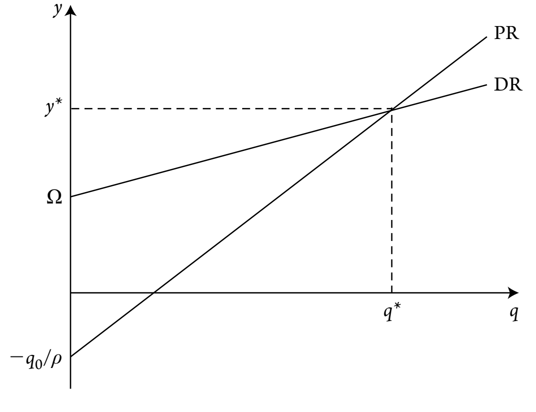



**Figure 8.2** The canonical Kaldorian (Dixon–Thirlwall) ELCC growth model

until the point $(q^*, y^*)$ is reached. Note that were the equilibrium in Figure 8.2 unstable, any initial $q > q^*$ would result in ever-increasing rates of output and productivity growth. The stability of the equilibrium depicted in Figure 8.2 thus ensures that in the first instance, the joint interaction of $y$ and $q$ in the ELCC model does not give rise to 'too much cumulation'.[32]

It is worth remembering that the ELCC model just presented was originally developed (for example, in Kaldor, 1966a [1989]) as a generalization of empirical regularities found in early cross-country regression analysis that became enshrined in Kaldor's growth laws. More recently, León-Ledesma (2002) has estimated an extended version of the Dixon–Thirlwall ELCC model that includes a fifth endogenous variable – a measure of R&D expenditures – together with various exogenous variables that are included to help identify the structural equations in a simultaneous equations framework.[33] León-Ledesma finds that most of the coefficients representing the key causal relationships in his extended ELCC model have the 'correct' (theoretically expected) signs and are statistically significant. He does not, however, test the predictive power of the ELCC model *relative to* any other particular growth model for explaining actual long-run growth rates.

## 8.3.3 Income divergence in the Dixon–Thirlwall model

The model developed so far serves to illustrate an important theme in Kaldorian growth theory: the possibility of income divergence, and hence growing inequality, between economies in the course of growth. To see this, consider two economies, *A* and *B*, that differ only with respect to their income elasticities of demand for exports, $\gamma$, such that

$$\eta_X^A > \eta_X^B \tag{8.25}$$

Then in terms of their respective DR curves (and as is revealed by inspection of equation 8.13 and the definition of $\Omega$) we have

$$\Omega^A > \Omega^B \tag{8.26}$$

and hence, as is illustrated in Figure 8.3

$$y^{A^*} > y^{B^*} \tag{8.27}$$

Now assume that, in levels, output is higher in country or region *A* initially ($Y^A > Y^B$). The consequences of this assumption, when coupled with the growth outcomes depicted in Figure 8.3,[34] are illustrated in Figure 8.4. Figure

**Figure 8.3** Growth outcomes in two different economies

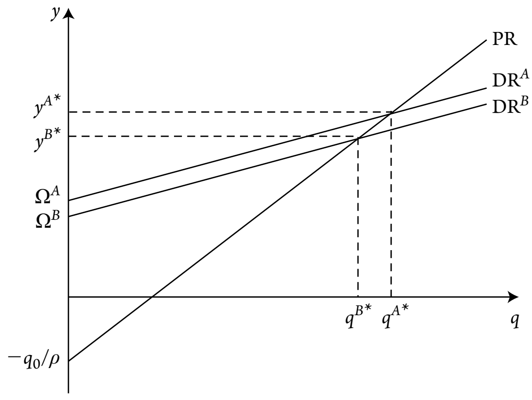
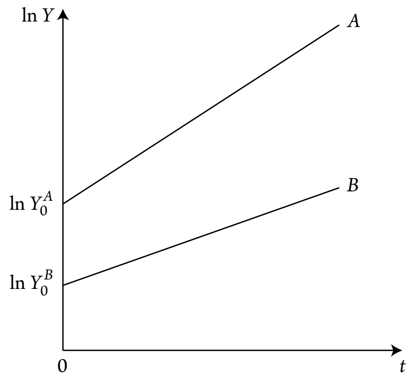



**Figure 8.4** Income divergence in the ELCC model




8.4 makes clear that, thanks to its initial advantage (at time $t = 0$) in the level of $Y$ and (from Figure 8.3) its self-perpetuating advantage in growth, economy $A$ will grow ever richer than economy $B$ over time in both absolute and relative terms (see Appendix 8.1 for a formal demonstration).[35] In other words, the inequality of income as between economies $A$ and $B$ will steadily increase, in both absolute and relative terms, in the course of growth. This pattern of divergence between 'rich' and 'poor' economies is consistent with the observed experience of advanced capitalist economies vis-à-vis the rest of the world (see, for example, Maddison, 1991, Table 1.5).

Even as the model illustrates the potential for divergence between rich and poor economies, however, it is important to note that it is also consistent with empirical findings of 'conditional convergence' – the tendency of poorer countries to grow faster than richer ones once a variety of influences on growth *other than the initial level of development* has been controlled for (see, for example, Mankiw et al., 1992). These findings are usually interpreted in terms of a neoclassical growth framework, from which the result of conditional convergence was first derived. But as mentioned previously in section 8.2.2, Roberts (2007) has shown that the same result can be derived from the canonical Kaldorian model outlined above. Essentially, this is because the transitional dynamics of the model above are qualitatively identical to those of the neoclassical growth model: the growth rate will tend to rise (fall) over time in any economy that initially grows slower (faster) than its equilibrium growth rate, as was illustrated in Figure 8.2 (see Roberts, 2007, pp. 624–6). Conditional convergence results that are usually interpreted in terms of NGT are therefore compatible with the canonical formal model of Kaldor's growth schema that has been outlined in this section.

## 8.4 Path dependence in the actual rate of growth

The model developed in the previous section is certainly faithful to the circular interaction between actual and potential output emphasized by Kaldor. Nevertheless, it seems to lack the requisite emphasis on history and path dependence in the growth process: it is, to all appearances, an ahistorical, traditional equilibrium model.[36] But contrary to appearances, the model in fact provides a good vehicle for exploring path dependence in the growth process, as will be demonstrated in this and the following section.

### 8.4.1 A disequilibrium approach to historical contingency

It was noted in the previous section that, providing certain existence and stability conditions are observed, the rates of growth of output and productivity will automatically gravitate towards their equilibrium values if they are above or below these equilibrium values initially. In other words, equilibrium outcomes such as $(q^*, y^*)$ in Figure 8.2 act as *point attractors*. Of course, if the rates of growth of output and productivity *are* different from their equilibrium values initially, then throughout the process of adjustment towards equilibrium, their values will depend on the rates of growth established initially. To see this, note that the choice of any arbitrary initial rate of growth in Figure 8.2 will result in a sequence of subsequent rates of growth (produced by the process of disequilibrium adjustment) that is uniquely determined by the choice of initial growth rate. Formally, if we rewrite the PR equation (8.21) in discrete time (with a '–1' subscript representing a one-period lag) as

$$q = q_0 + \rho y_{-1} \tag{8.21''}$$

and combine this expression with DR in equation (8.24), we get (recalling the definition of $\Omega$)

$$y = k_X(\eta_X - \rho_f \varepsilon_X)y_f + k_X \rho \varepsilon_X y_{-1} \tag{8.28}$$

This expression, which is a first-order difference equation in $y$, can be rewritten as

$$y = (k_X \rho \varepsilon_X)^t y_0 + k_X(\eta_X - \rho_f \varepsilon_X)y_f \sum_{i=1}^{t} (k_X \rho \varepsilon_X)^{i-1} \tag{8.29}$$

where $y_0$ denotes the initial rate of growth of output and $t$ is the number of time periods that has elapsed since these initial conditions were established. According to equation (8.29), the choice of $y_0$ determines the value of $y$ in all subsequent periods, holding all other factors constant.

Moreover, as previously discussed in Chapter 7 (section 7.5.3), it may not be possible to 'get into' equilibrium if the speed of adjustment towards equilibrium is slow relative to the rate at which the data defining the equilibrium are changing over time (Harcourt, 1981 [1982], p. 218; Fisher, 1983, p. 3; Cornwall, 1991, p. 107; Halevi and Kriesler, 1992, p. 229).[37] The upshot of these considerations is the following: the existence of a point attractor such as $(q^*, y^*)$ in Figure 8.2 notwithstanding, the rates of growth of output and productivity actually observed in the economy may always be a product of their initial rates in a system characterized by perpetual disequilibrium adjustment. We thus have a model of 'weak' path-dependent growth 'in which initial conditions, but no other feature of the economy's growth trajectory, influence subsequent growth outcomes in a purely self-reinforcing manner' (Setterfield, 2002, p. 220).[38] This is in keeping with Kaldor's emphasis on the lasting influence of initial conditions on growth outcomes in a system that never 'settles down' into a steady (equilibrium) rate of growth (see, for example, Kaldor, 1985, pp. 61–3).

## 8.4.2 A unit root in the growth process

An alternative to the disequilibrium approach is to postulate the existence of a 'unit root' in the growth process.[39] Recall that, as was demonstrated in the previous subsection, we can summarize the interaction of the DR and PR of our ELCC model in terms the following first-order difference equation, rewritten here for convenience

$$y = k_X(\eta_X - \rho_f \varepsilon_X)y_f + k_X \rho \varepsilon_X y_{-1} \tag{8.28}$$

The root of this equation is the constant coefficient $\Lambda = k_X \rho \varepsilon_X$, which determines the effect of $y_{-1}$ (the growth rate in the previous period) on the current rate of growth, $y$. A unit root exists when

$$\Lambda = k_X \rho \varepsilon_X = 1 \tag{8.30}$$

It will immediately be recognized that by postulating a unit root, we have changed one of the two conditions identified earlier as sufficient for the existence and stability of the equilibrium identified in Figure 8.2. The consequence of the unit root assumption is easiest to demonstrate if we also assume that[40]

$$\Omega = -q_0/\rho \tag{8.31}$$

Now note that $k_X \rho \varepsilon_X = 1 \Rightarrow \rho = 1/k_X \varepsilon_X$ and $\Omega = -q_0/\rho \Rightarrow q_0 = -\rho\Omega$. If we substitute these last two expressions into PR in equation (8.21), we get

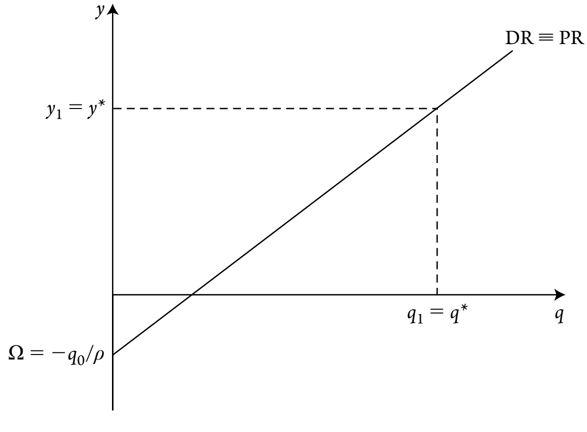
$$q = -\rho\Omega + \frac{1}{k_X \varepsilon_X} y \tag{8.32}$$

from which it follows that

$$y = \Omega + k_X \varepsilon_X q \tag{8.33}$$

(recalling that $k_X \rho \varepsilon_X = 1$ here). This is, of course, exactly the same as the expression for DR in equation (8.24). In other words, DR and PR are now identical, as depicted in Figure 8.5. And as is also illustrated in Figure 8.5, *any* initial choice of productivity growth rate (such as $q_1$) will generate a rate of growth of output ($y_1$), read off PR, that will, in turn, generate a rate of growth of productivity (read off DR) that is exactly equal to $q_1$. In other words, *ceteris paribus,* whatever growth rate is established initially will be indefinitely self-perpetuating. Put differently, all points along the DR ≡ PR curve depicted in Figure 8.5 are steady-state growth equilibria, so that $q_1 = q^*, y_1 = y^*$ for all $q_1, y_1$. The substance of this result is that, once again, the decisive influence

**Figure 8.5** The influence of initial conditions due to a unit root in the growth process



of initial conditions on subsequent growth outcomes (à la Kaldor) – that is, weak path dependence of growth – is established.

### 8.4.3 Strong path dependence I: technological lock-in and growth

The weak path dependence inherent in both the disequilibrium and unit root variants of the canonical Kaldorian model means that initial conditions affect long-run growth outcomes. But in these models, in the absence of unexplained, exogenous shocks, initial conditions are the *only* feature of the economy's prior growth trajectory that influence subsequent growth outcomes. However, a richer sense of historical contingency exists, which can be identified with strong path dependence. Strong path dependence involves structural change within an economy in response to its prior trajectory, where the latter may involve either a sequence of disequilibrium adjustments (as discussed in section 8.4.1 above), or cumulative experience of the same (equilibrium) outcome (such as that depicted in Figure 8.2). Specifically, strong path dependence exists when either the path towards or the cumulative experience of a particular equilibrium outcome affects the conditions of equilibrium (the data defining the equilibrium, such as the values of $\Omega$ and $\rho$ in DR and PR as depicted in Figure 8.2) and hence the position of equilibrium (that is, the precise equilibrium outcomes, such as $q^*$ and $y^*$ in Figure 8.2).[41] From this point of view, *all* positions of equilibrium (such as that depicted in Figure 8.2) are 'provisional' or 'conditional' (Chick and Caserta, 1997; Setterfield, 1997c). They exist only as long as the data defining them remain constant, and await subsequent redefinition resulting from discontinuous change in the structure of the economy that is induced by prior (equilibrium or disequilibrium) outcomes themselves. Hence, in the context of the model developed here, Figure 8.2 depicts no more than a transitory growth regime – a provisional or conditional characterization of the system that is adequate for the description of a particular episode of growth that may last for several consecutive business cycles, but which is ultimately susceptible to reconfiguration induced by the very outcomes that constitute the episode.[42]

There are various ways in which the structural change associated with strong path dependence may assert itself in the Kaldorian growth model. One of these concerns the pace of induced technological progress, as captured by the PR equation (8.21). Recall that $\rho$, the Verdoorn coefficient, captures the elasticity of productivity with respect to output – that is, the capacity of the economy to realize productivity gains on the basis of any given rate of growth of output. The value of this elasticity may be subject to discrete, growth-induced structural change due, for example, to technological interrelatedness and lock-in (Setterfield, 1997a, 1997b, 2002). Suppose, for instance, that rapid growth in the past causes an economy to get 'stuck' with certain industries and/or technologies inherited from the past. This might occur if rapid growth promotes specialization in production (as per Verdoorn's law), but at the same time, different components of the increasingly specialized production process (including plant, equipment and human capital both within and between firms, industries and the public sector) are *interrelated* – that is, subject to common technical standards that create interconnections between them. For example, certain types of computer software will work only on specific computer hardware, and require a specific skill set in order to be operated. Such interrelatedness makes it difficult to change one component of the production process without changing others. For example, an accounting firm may not be able to improve its software without simultaneously changing its computer hardware and retraining its employees.

The upshot is that technical change may become prohibitively costly and/or (in an environment of private ownership and decentralized decision making) difficult to coordinate in an economy that has grown extensively (that is, rapidly and/or over a protracted period of time) by accumulating certain interrelated types of human and physical capital, and in which the degree of interrelatedness between components of the production process has, as a result, surpassed a critical threshold level. Such an economy can be said to have become 'locked-in' to a particular technological base, inherited as a legacy of its past, from which it subsequently becomes difficult to deviate. And this, in turn, may impair the ability of the economy to realize induced technological progress in the future. Hence, if a technological improvement is incompatible with existing components of the production process, it may be foregone. The result is that the economy will experience a discrete drop in the size of its Verdoorn coefficient, $\rho$, which measures the ability of the economy to capture induced technological progress, as the threshold level of interrelatedness is surpassed and the economy experiences lock-in. The consequences of this are illustrated in Figure 8.6. Beginning with the same conditional growth equilibrium (at $q^*, y^*$) depicted in Figure 8.2, assume that cumulative experience of these growth outcomes creates lock-in to a particular technological base, as described above. This, in turn, will transform the economy's PR from

$$q = q_0 + \rho y \tag{8.21}$$

to

$$q = q_0 + \rho' y \tag{8.21'''}$$

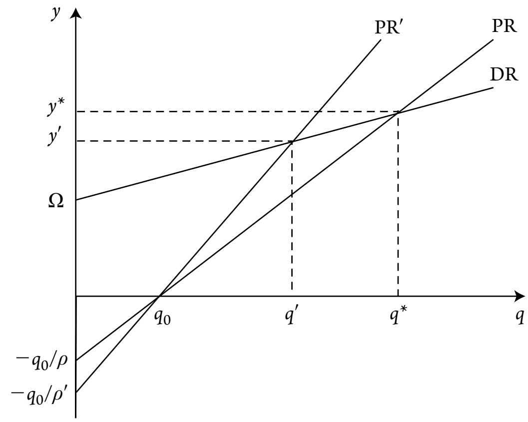



**Figure 8.6** The consequences of technological interrelatedness and lock-in

where $\rho' < \rho$. The upshot of this development is a reduction in the conditional equilibrium rates of output and productivity growth to $y'$ and $q'$ respectively, as illustrated in Figure 8.6. Clearly, Figure 8.6 exemplifies strong path dependence as defined earlier. In this case, the cumulative experience of a particular (conditional) equilibrium outcome affects the conditions of equilibrium (the Verdoorn coefficient, $\rho$) and hence the position of equilibrium itself.

## 8.4.4 Strong path dependence II: institutional change and growth

Technology is not the only source of discontinuous structural change that can be associated with strong path dependence. Another source is institutions, defined broadly to include conventions and norms as well as formal (for example, legal) rules. According to Setterfield and Cornwall (2002), institutions create a framework akin to a computer's operating system,[43] within which the income-generating process summarized in equations (8.9) and (8.19)–(8.21) is embedded. Hence the parameters (and even the precise functional forms) of DR and PR in equations (8.21) and (8.24) reflect the structure of the economy's institutional framework. For example, a 'value-sharing' norm that ensures that both workers and firms benefit from productivity gains may reduce conflict over technological change at the point of production, and thus increase the responsiveness of productivity growth to output growth (as captured by the Verdoorn coefficient, $\rho$). This, in turn, will affect the position of the PR curve in Figure 8.2 and hence the economy's rates of growth of output and productivity.

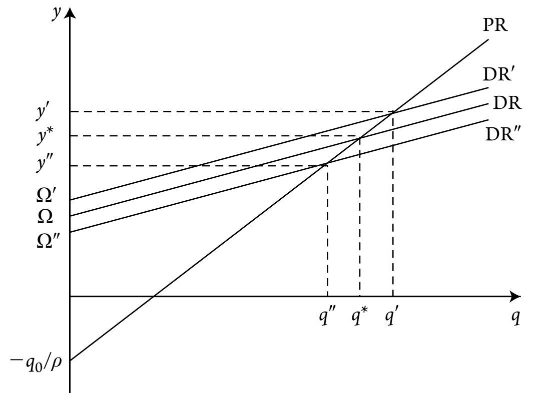
According to Setterfield and Cornwall (2002), the economy's institutional framework is relatively inert and hence enduring – sufficiently so as to give rise to precisely the sort of discrete episodes of growth, lasting for several consecutive business cycles, alluded to in the previous subsection. And as was suggested earlier, these growth episodes can be characterized by equilibrium growth outcomes of the sort depicted in Figure 8.2, as long as such equilibria are understood to be strictly conditional – in this case, conditional on the reproduction over time of the specific constellation of institutions within which DR and PR are embedded. This conditionality of the equilibrium draws our attention to the fact that, while relatively enduring, the institutional framework is not immutable. It can and does change over time, not least in response to the cumulative effects of the growth outcomes to which it gives rise.[44] For example, if sustained economic growth creates 'aspiration inflation' resulting in the breakdown of the value-sharing norm described earlier, then heightened distributional conflict at the point of production may impair the capacity of the economy to realize induced technological change, reducing the size of the Verdoorn coefficient, shifting PR and thus reducing the rates of growth of output and productivity in a manner similar to that depicted in Figure 8.8. In other words, the institutional framework shapes DR and PR in equations (8.21) and (8.24), thus creating a discrete episode of growth characterized by a conditional growth equilibrium (such as that depicted in Figure 8.2). But growth outcomes then have feedback effects on institutions, that eventually become manifest as institutional change. The upshot will be a new DR and/or PR, and hence a new episode of growth, and so on. Once again, we are describing a process whereby the cumulative experience of a particular (conditional) equilibrium outcome affects the conditions and hence the position of equilibrium – in other words, a system that displays strong path dependence.

Setterfield and Cornwall (2002) use the model described above to chart the rise and decline of the post-war Golden Age (1945–73) of macroeconomic performance in terms of discrete institutional changes interacting with the Kaldorian income-generating process summarized in equations (8.21) and (8.24). As a further example of their approach, consider the international transmission of the rise and decline of the financialized US growth process over the past 20 years. It is widely accepted that growth in the US economy over the last 20 years was consumption-led, and financed by unprecedented household debt accumulation (Palley, 2002a; Cynamon and Fazzari, 2008). According to Cynamon and Fazzari (2008), this financialized growth episode in the US was brought about by significant changes in the borrowing and lending norms of households and creditors, respectively. Moreover, the institutional change that Cynamon and Fazzari identify can be thought of as having been (in part) induced by the macroeconomic performance experienced in the US during what Setterfield and Cornwall (2002) identify as the low-growth 'Age of Decline' (1973–89).[45]

Hence one important macroeconomic outcome that was established during this low-growth episode was the tendency for real wages to grow slower than productivity for the majority of workers, thus depressing the wage share of income (see, for example, Palley, 2002a). This outcome can be traced directly to an important institutional feature of modern American capitalism that emerged during the Age of Decline – its 'incomes policy based on fear', associated with changes in corporate organization, labour law and macroeconomic policy designed to increase worker insecurity and reduce the relative power of workers in the wage bargain (Harcourt, 2007, pp. 63–4; Setterfield, 2006a, 2007). And as Cynamon and Fazzari (2008) argue, stagnant real wage growth has contributed to an increased acceptance among American households of debt accumulation as a mechanism for pursuing the 'American dream' of rising living standards. At the same time, the incomes policy based on fear alluded to above was designed to subdue inflationary pressures in the US economy – something it was successful in doing (Setterfield, 2006a, 2007). The resulting low (and stable) inflation environment that began to materialize towards the end of the Age of Decline helped to induce changes in creditors' lending norms, by reducing their macroeconomic risk and hence creating an incentive for them to pursue greater microeconomic risk, such as accepting greater household leverage and making subprime mortgage loans to financially vulnerable households (see, for example, Goodhart, 2005, p. 300).

The upshot of these developments was a debt-financed, consumption-led growth episode in the US after 1990, which has had beneficial effects for countries exporting to the US as a 'consumer of last resort' (see Blecker, 2013a; Hein and Mundt, 2013). The international transmission of this financialized US growth episode (and its recent demise) is captured in Figure 8.7.[46] Suppose, then, that we begin at the equilibrium denoted by $q^*$, $y^*$ as originally depicted in Figure 8.2. The emergence of the financialized growth process in the US can be reckoned to have had two effects on the DR of countries exporting to the US. The first, direct effect is an increase in $y_f = y_{US}$ and hence in $\Omega = k_X([\eta_X - \rho_f \varepsilon_X]y_f - \varepsilon_X q_0)$, where $y_{US}$ denotes the rate of growth of the US economy which is treated as a proxy for $y_f$ in economies exporting to the US as a consumer of last resort. The second, indirect effect operates via the income elasticity of demand for exports, $\eta_X$. The increased leverage of US households over the past two decades suggests that, for any given proportional increase in real income, the proportional increase in expenditures by US consumers on all goods and services (including imports) has increased (*ceteris paribus*), as income growth (which funds additional consumption) has been accompanied by debt accumulation (which finances additional consumption over and above what would be possible out of

**Figure 8.7**
International transmission of the rise and demise of the financialized US growth regime (for a country exporting to the US)




additional income).[47] This will manifest itself as an increase in $\eta_X$ and hence (again) in $\Omega = k_X([\eta_X - \rho_f\varepsilon_X]y_f - \varepsilon_X q_0)$. In other words, both the direct and indirect consequences for countries exporting to the US of the financialized US growth process involve an increase in $\Omega$ (to $\Omega'$ in Figure 8.7), which will shift DR upward (to DR′ in Figure 8.7) thus raising the equilibrium rates of output and productivity growth (to $y'$ and $q'$, respectively, in Figure 8.7).

As the events of 2007–09 demonstrated, however, the financialized US growth regime was unsustainable.[48] The financial crisis and ensuing Great Recession in the US economy had both direct and indirect effects on countries exporting to the US as the consumer of last resort, effects that are again captured in Figure 8.7. First, the direct effect of the Great Recession on such countries was to reduce $y_f = y_{US}$ and hence $\Omega = k_X([\eta_X - \rho_f\varepsilon_X]y_f - \varepsilon_X q_0)$. Second, the combination of the Great Recession and the financial crisis changed the proclivity of US households and creditors to borrow and lend respectively, with the result that the proportional expansion of expenditures accompanying any given proportional expansion of income – and hence the value of $\eta_X$ for countries that export to the US – may have decreased to some extent,[49] again lowering $\Omega = k_X([\eta_X - \rho_f\varepsilon_X]y_f - \varepsilon_X q_0)$. These developments are captured by the decrease in $\Omega$ (to $\Omega''$ in Figure 8.7), the resulting downward shift in DR (to DR″ in Figure 8.7), and the accompanying fall in the equilibrium rates of output and productivity growth (to $y''$ and $q''$, respectively, in Figure 8.7). The remaining question, of course, is whether these events prove to be temporary, or whether the financialized growth regime in the US is truly exhausted – in which case, *ceteris paribus*, lower growth outcomes similar to $y''$ and $q''$ in Figure 8.7 will persist as a new growth episode as the US leads the world into a period of secular stagnation.[50]

## 8.5 Reconciling the actual and natural rates of growth

In the Kaldorian model outlined in section 8.3, not only is the actual (equilibrium) rate of growth path dependent but so, too, is the Harrodian *natural* rate of growth – the maximum rate of growth that the economy can achieve in the long run. This is because the natural rate is sensitive to the actual rate of growth that the economy achieves, thanks to the operation of Verdoorn's law. This is illustrated in Figure 8.8. Figure 8.8 shows how the equilibrium rate of productivity growth, $q^*$, established by the intersection of DR and PR in the second quadrant of the diagram, determines the equilibrium natural rate of growth, $y_N^*$, in the third quadrant, given the rate of growth of the labour force, $n$.[51]

It is also evident from Figure 8.8 that, even though the natural rate of growth is endogenous, the first Harrod problem – inequality of the equilibrium and natural rates of growth – may persist (Cornwall, 1972). In fact, as in Harrod, $y^* = y_N^*$ will emerge only as a special case in the model developed thus far. The reasons for this can be made clear as follows. First, note that from the solution to equations (8.21) and (8.24), it follows that

$$y^* = \frac{k_X(\eta_X - \rho_f \varepsilon_X)y_f}{1 - k_X \rho \varepsilon_X} \tag{8.34}$$

Meanwhile, since by definition the natural rate of growth is

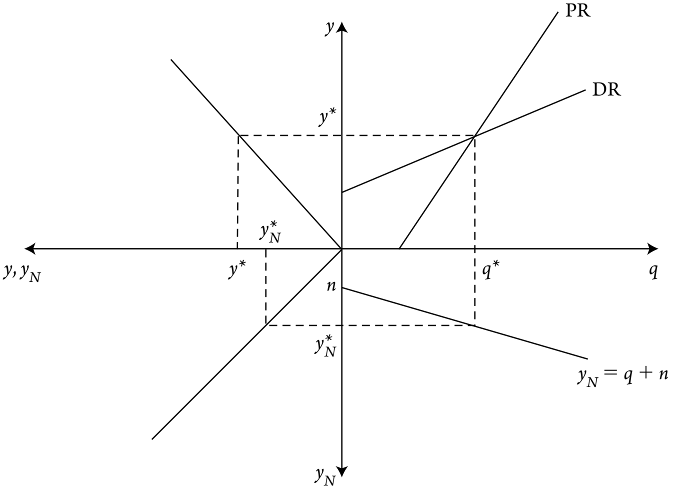



**Figure 8.8** The endogeneity of the natural rate of growth

$$y_N \equiv q + n \tag{8.35}$$

it follows from appeal to Verdoorn's law that

$$y_N^* = q_0 + n + \rho y^* \tag{8.36}$$

Finally, solving equations (8.34) and (8.36) under the equilibrium condition $y^* = y_N^*$ yields

$$\frac{q_0 + n}{1 - \rho} = \frac{k_X(\eta_X - \rho_f \varepsilon_X)y_f}{1 - k_X \rho \varepsilon_X} \tag{8.37}$$

It is clear by inspection that the equality in (8.37) is possible but not likely: it involves a constellation of independently determined parameters, and there is no obvious mechanism that will ensure these parameters take on values that exactly satisfy (8.37). Once again, we are confronted by the first Harrod problem.

The result derived above raises an important question about the *sustainability* of the equilibrium rate of growth depicted in Figure 8.8. Hence note that since by definition

$$y \equiv q + l \tag{8.38}$$

where $l$ denotes the rate of growth of employment, it follows from this definition and that of the natural rate of growth stated earlier that, if $y^* > y_N^*$ as in Figure 8.8, we will observe

$$q^* + l^* > q^* + n \tag{8.39}$$

$$\Rightarrow l^* > n$$

where $l^*$ is the equilibrium rate of growth of employment derived from the equilibrium rates of output and productivity growth determined in Figure 8.8, and the definition of the actual rate of growth stated above. Recalling our definition of the employment rate from earlier chapters as $e = L/N$, we can see that

$$\dot{e} = (l^* - n)e \tag{8.40}$$

Equation (8.40) tells us that, given the rate of growth of the labour force, the employment rate $e$ will keep increasing as long as $l^* > n$. But since the employment rate is bounded above (it cannot exceed one), this is impossible to maintain indefinitely.[52] The condition $y^* = y_N^*$ (or, equivalently, $l^* = n$) therefore constitutes a necessary condition for sustainable, long-run equilibrium growth – or at least, it does if $e$ is sufficiently high to begin with, and $n$ is sufficiently inelastic with respect to $l$ (as a result of impediments to migration, limits to increased labour force participation, and so forth). In other words, only if we are analysing a dual economy – that is, one with an abundant latent reserve army of labour in a subsistence or informal sector, that can be drawn (on demand) into the modern sector whose growth is described by the model we have developed so far – can the necessary condition $y^* = y_N^*$ be satisfactorily ignored.

But advanced capitalist economies are not dual economies, and it is clear from their post-war experience that they are capable of operating near to full employment – in which case any growth outcome similar to that depicted in Figure 8.8 must be regarded as ultimately unsustainable. Of course, it must be remembered that we are treating growth equilibria such as that depicted in Figure 8.8 as 'conditional' and that, as such, a growth regime or episode such as that in Figure 8.8 may come to an end before the logical bounds of the employment rate have been tested. Nevertheless, the possibility that a growth episode may become labour constrained (that is, unsustainable because $l^* \neq n$) should alert us to the potential importance of the necessary condition $y^* = y_N^*$ and hence to the importance of studying processes through which the equilibrium actual and natural rates of growth (and hence $l$ and $n$) might be brought into alignment, so that growth episodes can be made consistent with a constant employment rate and thus become (in principle) sustainable in the long run.

Because of the importance of this issue, it is not surprising to find that several such processes have been proposed and incorporated into Kaldorian models (see Chapter 1, section 1.5.2, and Cornwall, 1972). Models of this sort (for example, Palley, 2002c; Setterfield, 2006b) are structured so that the actual and natural rates of growth are equalized by means of processes that cause either the actual rate or the natural rate to adjust in the event of changes in $e$ – circumstances that, as demonstrated above, will always be observed if the actual and natural rates of growth are not equal to begin with. Much of this newer literature was written specifically in relation to the BPCG model, so we will just give a flavour of the analysis (adapted to an ELCC framework) here, while the original versions will be covered in greater detail in Chapter 10, section 10.4.

In Palley (2002c), the actual rate of growth bears the burden of adjustment, varying indirectly with the value of $e$ – so that if $y^* > y_N^*$ initially, $e$ will rise thereby reducing $y^*$ until $y^* = y_N^*$. An inverse relationship between $y^*$ and $e$ is justified by the notion that higher employment rates $e$ are associated with localized (industry- or even firm-specific) bottlenecks in the domestic economy, which divert demand abroad. As a result, further expansions of income of a given size will result in smaller realized expansions of demand for domestic output. In other words, the income elasticity of demand for domestic output will fall (and correspondingly, the income elasticity of demand for imports will rise), reducing the actual rate of domestic growth.[53]

In Setterfield (2006b), meanwhile, the natural rate of growth is the adjusting variable, varying directly with the value of $e$. If $y^* > y_N^*$ initially, the value of $e$ will rise, and this will now increase the value of $y_N^*$ until $y^* = y_N^*$. The direct relationship between $y_N^*$ and $e$ is explained by the notion that the size of the Verdoorn coefficient, $\rho$, is increasing in $e$. This is because the closer the economy operates to full capacity, the more likely are firms to engage in innovation and/or technical and organizational change in response to the *expansion* of the economy: when the economy is slack, similar expansion is likely to induce only increases in the utilization of existing (idle) capacity. In other words, it is not just the *rate of growth* but also the *level* of economic activity that is understood to influence induced technological progress via the Verdoorn law. The upshot of all this is that as $e$ and hence $\rho$ rise, so, too, does the natural rate of growth.[54]

As a result of either the Palley (2002c) or Setterfield (2006b) process, the necessary condition for sustainable, steady-state growth $(y^* = y_N^*)$ will be satisfied once equilibrium (with $\dot{e} = 0$) is achieved. The different adjustment processes postulated do, however, mean that the final conditional equilibria achieved will have different qualitative properties. The Setterfield (2006b) process gives rise to a conditional equilibrium consistent with what may be called 'fully demand-determined growth', because it involves adjustments on the supply side, which work through changes in the value of the natural rate of growth and accommodate the conditional equilibrium value of the demand-determined actual rate of growth. The Palley (2002c) process, in contrast, gives rise to 'semi-supply-determined growth', so called because the achievement of a final conditional equilibrium position now involves adjustment of the demand-determined actual rate of growth towards the natural rate – although the latter remains endogenous to the former (by virtue of Verdoorn's law), so that the demand-determined character of the final conditional equilibrium position is not lost altogether. Of course, the two adjustment processes are not mutually exclusive: they could operate simultaneously, with the result that adjustments on both the demand and the supply sides of the economy in response to the tightening (or slackening) of the goods and labour markets could play a role in determining the final conditional equilibrium rate of growth consistent with sustainable growth at the (endogenous) natural rate, $y^* = y_N^*$.

In order to formally demonstrate the possibility of reconciling the actual and natural rates of growth in the context of our ELCC model, let us focus our attention on the process proposed by Setterfield (2006b). This can be modelled as

$$\rho = \rho(e), \rho' > 0 \tag{8.41}$$

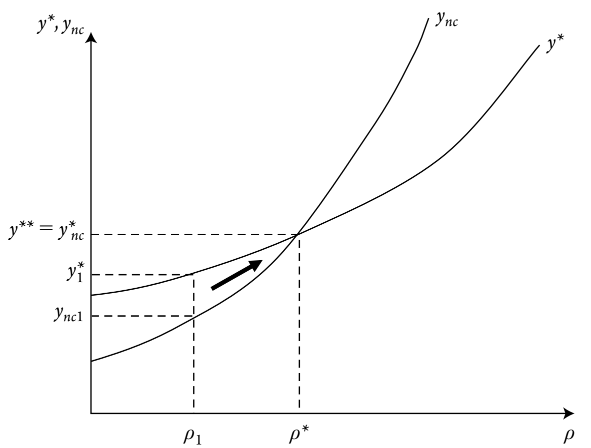
The consequences of equation (8.41), which captures the direct effect of $e$ on the Verdoorn coefficient discussed above, are illustrated in Figure 8.9 (see Appendix 8.2 for a formal derivation). This figure depicts both the ELCC equilibrium rate of growth from equation (8.34) (the curve denoted $y^*$) and the rate of growth that satisfies the necessary condition for sustainable long-run growth $y^* = y_N^*$ (the curve denoted $y_{nc}$, where '*nc*' denotes 'necessary condition') as functions of the Verdoorn coefficient, $\rho$. By referring to the left-hand side of equation (8.37), we can see that the latter relationship can be stated as

$$y_{nc} = \frac{q_0 + n}{1 - \rho} \tag{8.42}$$

Figure 8.9 depicts a situation where, with $\rho = \rho_1$, $y_1^* > y_{nc1}$ and hence, as demonstrated earlier, $l^* > n$. This will result in $\dot{e} > 0$ in equation (8.40), as a result of which $\rho$ will rise in equation (8.41), increasing the values of both $y^*$ and $y_{nc}$ in Figure 8.9. These adjustments will continue until $\rho = \rho^*$ in Figure

Note: The subscript '*nc*' refers to 'necessary condition' (see text for explanation).

**Figure 8.9**
Adjustment towards a sustainable equilibrium growth rate



8.9, at which point $y^{**} = y^{*}_{nc}$. At this point, the economy will have achieved a conditional equilibrium rate of growth that satisfies the necessary condition $y^{*} = y^{*}_{N}$ and is therefore sustainable in the long run.

## 8.6 Policy implications

Because it describes growth as resulting from a joint interaction of supply and demand conditions in a process of cumulative causation, the formal structure of the ELCC model suggests that both supply-side and demand-side policies can affect the rate of growth. Consider first the supply side of the economy. In the ELCC model, any policy that exogenously stimulates productivity growth (for example, an R&D subsidy or improved technical education) will increase the value of $q_0$, shift the PR curve in Figure 8.2 down and to the right, and so increase the equilibrium growth rate.

We must not forget, however, that the relative autonomy of demand conditions in Kaldorian growth theory (following from Keynes's principle of effective demand) privileges the position of demand formation in the process of growth, which is therefore (and despite the joint interaction of supply and demand conditions) understood to be demand-led. The focus of policy intervention in the ELCC model is therefore on the demand side, where numerous policy interventions may succeed in stimulating growth.[55] In the first place, any policy intervention that stimulates export growth will shift the DR curve upward in Figure 8.2, increasing the growth rate. Such policies might include an opening of foreign markets that raises the income elasticity of export demand $\eta_X$, or a policy-induced currency depreciation (a rise in the nominal exchange rate, $E$, so that $\hat{E} > 0$, where $\hat{E}$ denotes the rate of nominal currency depreciation).[56] Note that in the ELCC model, export and hence output growth can be increased by both cost reductions (resulting, for example, from currency depreciations) *and* any improvements to the quality of tradable goods, after sales service, product marketing and so on that affect the income elasticity of demand for exports. In other words, growth is sensitive to both price and non-price competitiveness in international trade.

Perhaps the most radical policy implication of the ELCC model, however, is that a stimulus to *domestic* demand can potentially spark a virtuous circle of export-led growth. To see this, suppose that, following Setterfield and Cornwall (2002), we rewrite equation (8.9) as

$$y = k_X(\omega_X x + \omega_A g_A) \qquad (8.9')$$

where $g_A$ denotes the rate of growth of the autonomous component of domestic demand and $\omega_j$ denotes the share of $j = A, X$ in total output.[57] Replacing equation (8.9) with equation (8.9′) transforms the DR of the ELCC model into

$$y = \Omega + k_X \omega_X \varepsilon_X q \qquad (8.24')$$

where $\Omega = k_X(\omega_A g_A + \omega_X[(\eta_X - \rho_f \varepsilon_X)y_f - \varepsilon_X q_0])$. It can now be seen that a stimulus to domestic demand (an increase in $g_A$) would have the same effect as a stimulus to export demand, resulting in an upward shift of the (modified) DR curve and a permanent increase in the equilibrium growth rate. In fact, given the logic of the ELCC model, part of the stimulus to growth would result from the fact that the domestic demand stimulus would increase export growth, by virtue of its causing productivity to grow faster and hence exports to become more competitive.

This may seem like a strong or even scarcely believable conclusion. It should be noted, however, that it stems in part from the (oversimplified) aggregate one-sector structure of the model. In a more refined (disaggregated) ELCC framework, in which the Verdoorn law was limited to manufacturing output, only that part of the increased domestic demand that contributed to the demand for domestically produced manufactures would stimulate the process of cumulative causation. On this view, the *structure* (not just the level) of domestic demand matters for the determination of the rate of growth.

## 8.7 Critiques, evaluations and extensions of the ELCC model

### 8.7.1 Omissions from the model

One obvious omission from the ELCC model, discussed earlier in this chapter, is the importance attached to structural change and the special role of the manufacturing sector in Kaldor's growth laws. This omission requires little further discussion here; as previously noted, disaggregation of the ELCC model is required to properly address the concern. Since this is something that has already been achieved in the BPCG strand of Kaldorian growth theory, we will defer further discussion of the impact of the sectoral structure of production on growth until Chapters 9 and 10.

As is commonly the case in HGT theory, the ELCC model makes no explicit reference to the labour market or monetary relations in its description of capitalist growth. The canonical Dixon–Thirlwall model is a purely real-sector model, in which there is no binding labour supply constraint on output growth,[58] nor any systematic effect of growth (and hence the relative tightness of the labour market) on nominal wage formation. In other words, both the labour market and the monetary sector are assumed to be passively accommodating of the growth outcomes it describes. The addition of labour market dynamics and an explicit monetary and financial sector would be required in order to relax these assumptions – although note that discussion of reconciling the actual and natural rates of growth (and thereby confronting issues associated with the first Harrod problem) in the ELCC model, as discussed in the previous section, goes some way towards addressing the absence of a labour supply constraint on the growth.

Another assumption commonly made in the Dixon–Thirlwall model is that the nominal exchange rate ($E$) remains constant, or else has an exogenously given rate of change.[59] This assumption can certainly be called into question. Much theory and intuition suggest that a country experiencing an export-led boom might expect to confront pressures towards currency appreciation to a degree that is sensitive to the extent of the boom itself (see Blecker, 2013b). This might be avoided through policy intervention in the currency market (as practised in China in the past), and so the assumption may not be without empirical merit in some specific cases. Nevertheless, explicit introduction of exchange rate dynamics is required in order to successfully relax the standard ELCC assumption of constancy or exogeneity of the nominal exchange rate (or rate of nominal exchange rate depreciation).

### 8.7.2 'Too much cumulation, too few contradictions?'

An important caveat in Kaldor's concept of cumulative causation and hence the canonical Dixon–Thirlwall ELCC model is that either fast or slow growth is strictly self-perpetuating: success breeds success or failure begets failure indefinitely. It would appear that once the initial relative success or failure of a region is known, so too is the rest of its history. Hence according to critics such as Gordon (1991), the ELCC model puts too much emphasis on self-reinforcing change, displaying (as a consequence) 'too much cumulation and too few contradictions'.

The extent of this problem should not be exaggerated, because of the relative autonomy of demand conditions in the principle of effective demand that is central to the Kaldorian vision. Consider the case of a virtuous circle of cumulative causation (as in Figure 8.1), for example. In any 'round' of the process of cumulative causation, the dynamics of the virtuous circle require sufficient demand formation in response to the realization of scale economies on the supply side to support the new potential level of output and so to propagate the next round of cumulative causation (Ricoy, 1987, p. 733). In terms of the formal structure of the Dixon–Thirlwall model, there exists a potential 'weak link' in the causal chain of a virtuous circle if the rate of growth of exports is subject to exogenous shocks.[60]

One possible objection to the preceding argument is that it relies on *unexplained* changes in demand to generate the breakdown of a virtuous circle: the weak link we have identified on the demand side of the ELCC model does not arise from *within* the process of cumulative causation itself. Note, however, that the discussions of strong path dependence in section 8.4 address precisely this problem, by conceiving cumulative causation as a growth dynamic that occurs, not within a vacuum, but instead within precise technological and institutional contexts. These technological and institutional regimes characterize the technical and social structure of the supply side of the economy. As the discussion in section 8.4 reveals, they may both influence and be influenced by the process of growth, with the result that, for example, a relatively fast-growing economy may, as a result of its relatively fast growth, induce technological and/or institutional change (or in the case of lock-in, a lack thereof) which is inimical to the maintenance of its high-growth dynamics. In this way, a virtuous circle of cumulative causation can endogenously break down.

The extensions to the ELCC model in section 8.4 allow the model to move away from the prediction that initially high relative growth is indefinitely self-perpetuating via the process of cumulative causation. They can therefore be thought of as bounding the process of cumulative causation so that there is not, to paraphrase Gordon (1991), *too much* cumulation. At the same time (and as we have already seen), these amendments are thoroughly in keeping with the spirit of Kaldorian growth theory: the limit or constraint they impose on the process of cumulative causation arises as a result of the prior growth process itself, lending emphasis to the theme that growth is an innately historical or path-dependent process. Of course, even the basic or canonical Dixon–Thirlwall model requires a restriction on the degree of cumulative causation in order to ensure that the ELCC equilibrium is stable and growth is not explosive, as discussed in section 8.3.2, but the key question is whether such an equilibrium can plausibly be maintained indefinitely or is likely to be subject to *endogenous* adjustments and reversals as well as possible exogenous shocks.

### 8.7.3 The Kaldor paradox and the Beckerman model

Early empirical tests of the ELCC model appeared to be unfavourable. In particular, the estimated elasticity of export growth with respect to a change in relative prices took the wrong sign in cross-sectional data ($\varepsilon_X < 0$ rather than $\varepsilon_X > 0$).[61] This phenomenon became known as the 'Kaldor paradox' (Kaldor, 1978), with various explanations offered for its observation. For example, the paradox may arise from reverse causality: rather than more expensive exports (real appreciation) raising export and output growth – which would seem perverse – it may be that faster export and output growth increases the demand for labour, which raises wages and so makes home country products more expensive. Nevertheless, Kaldor (1981) was convinced by the empirical evidence to abandon the ELCC model on the grounds that price or cost competitiveness – which is an important theoretical link between productivity improvements and export growth in the process of cumulative causation – didn't seem to matter in international trade.[62]

But was the evidence really convincing, or did Kaldor abandon his own theory too quickly? It is noteworthy, for example, that in the later empirical work of León-Ledesma (2002) discussed earlier, a result consistent with the predictions of ELCC theory emerges once investment rates and R&D expenditures are controlled for.[63] Moreover, as emphasized by Boggio and Barbieri (2017), the underlying concern of the Kaldor paradox – that there should be a causal relation running from the *rate of growth* of unit labour costs to the rate of growth of output (via the cost competitiveness of domestically produced goods and hence the rate of growth of exports) – overlooks Kaldor's (1971) original claim that:

> the main autonomous factor governing both the level and the rate of growth of effective demand of an industrial country with a large share of exports in its total production and of imports in its consumption is the external demand for its exports: and the main factor governing the latter is international competitiveness, which in turn depends on the *level* of its industrial cost relatively to other industrial exporters. (Kaldor, 1971, p. 7; emphasis added)

The operative phrase in the quotation above is '*level* of its industrial cost', rather than 'growth rate of its industrial costs'. In fact, in the earlier export-led growth model developed by Beckerman (1962), the growth rate of exports was assumed to be a function of the *level* (not the growth rate) of a country's prices (or unit labour costs) relative to those of other trading nations

$$x = z + \gamma(1 - \Gamma) \tag{8.43}$$

where $z$ is the growth rate of total world trade, $\gamma$ is a positive parameter and $\Gamma$ is some measure of relative competitiveness (such as the level of the real exchange rate or relative unit labour costs). Note that since equation (8.43) can be rewritten as

$$x - z = \gamma(1 - \Gamma) \tag{8.44}$$

it is effectively an equation expressing the rate of growth of a country's *share* of world trade as a function of the level of relative competitiveness, $\Gamma$. Note also the important contrast between the expression in equation (8.19) from the canonical ELCC model, in which the rate of growth of exports depends on the *rate of growth* of the real exchange rate $\hat{P}_f + \hat{E} - \hat{P} = \hat{P}_f - \hat{P}$ (assuming $\hat{E} = 0$), and equation (8.43), where with (for example) $\Gamma = P/EP_f$, the rate of growth of exports depends on the *level* of the real exchange rate (where, for consistency with the Beckerman specification, we have to use the inverse of the real exchange rate as usually defined, that is, the real value of the home currency).

Beckerman's approach has recently been revived by Boggio and Barbieri (2017),[64] who show that Beckerman's export share equation is mathematically equivalent to a 'replicator equation' in evolutionary biology. In replicator equations, the growth of a variety's share in the total population of a species is a function of the level of its fitness (not the rate of growth of its fitness) relative to that of other species. Boggio and Barbieri (2017) argue that the Beckerman/replicator formulation of the growth of export (equation 8.43) is more consistent with Kaldor's original quote from 1971 than is equation (8.19) in the standard (Dixon–Thirlwall) ELCC model. Moreover, they provide empirical evidence linking change in export shares to the level of competitiveness (measured by relative unit labour costs), consistent with equation (8.44). Indeed, they show that levels of this cost competitiveness variable are statistically significant in determining changes in export shares whereas as growth rates of the same variable are not. They conclude that dismissal of the importance of relative price (cost) effects in trade and growth based on the Kaldor paradox was premature, because relative costs – in levels – *do* significantly affect a country's comparative export growth performance. This finding is consistent with much recent empirical research showing the importance of levels of over- or undervaluation of the real exchange rate for growth (Rodrik, 2008; Berg et al., 2012; Rapetti et al., 2012).[65]

### 8.7.4 Cumulative causation and the balance-of-payments constraint

Almost as soon as it was developed, the ELCC growth model received an important challenge from Thirlwall (1979) and others, who, although sympathetic to the Kaldorian approach, believed that the ELCC models erred in ignoring the role of import demand and neglecting to incorporate a BP equilibrium condition. Thirlwall and Dixon (1979) criticized the ELCC model (including their own earlier version from which the canonical ELCC model derives) because

> No consideration is given to the possibility that the rate of growth of income determined by the model may generate a rate of growth of imports in excess of the rate of growth of exports, thereby imposing a constraint on the export-led growth rate if balance of payments equilibrium must be preserved. (Thirlwall and Dixon, 1979, p. 173)

The concern here is that in the course of its export-led growth, the ELCC model may describe a country running a chronic BP deficit or surplus on current account, thereby implicitly assuming that it can incur ever-growing foreign debt (running a capital account surplus by borrowing from abroad) or amass ever-increasing foreign assets *in perpetuity*. In fact this may not be sustainable, which compromises the claim of the ELCC to provide a description of long-run growth. Furthermore, if the equilibrium rate of growth consistent with the ELCC model for some countries implies continuous current account surpluses, the implicit assumption of the model is that *other countries* must be able to run perpetual deficits and hence borrow from abroad without limit, which raises the same questions of sustainability. Despite our earlier demonstration that export-led growth is *potentially consistent* with balanced trade, and hence does not necessarily assume that any individual economy is running a current account deficit or surplus, there is no mechanism in the ELCC model that *ensures* that this is the case.

It transpires that if import demand is incorporated into the ELCC model and a BP constraint is imposed, exports continue to play a key role in determining long-run growth. This is because faster growth of exports allows faster growth of imports without risking a chronic BP current account deficit.[66] Nevertheless, Thirlwall and Dixon (1979) show that, under certain circumstances, the cumulative causation mechanism characteristic of the ELCC model is thwarted, and the growth rate consistent with BP equilibrium is determined solely by the ratio of the growth rate of exports to the income elasticity of import demand *regardless* of whether Verdoorn's law (Kaldor's second law, which incorporates dynamic increasing returns) holds. This solution for the *BP-equilibrium growth rate*, originally found in Thirlwall (1979), is sometimes referred to (following Davidson, 1990–91) as 'Thirlwall's law'. Thirlwall's law and the broader BPCG approach that has arisen from it are the subject of investigation in the next two chapters.

## 8.8 Conclusions

This chapter has explored one of the main branches of modern Kaldorian growth theory, the ELCC model. The central principles of this approach are that growth is (1) demand-led, with exports playing a crucial role in aggregate demand formation; and (2) path dependent. In Kaldor's original vision, path dependence is associated specifically with the process of cumulative causation, in which initial conditions are self-reinforcing. In modern Kaldorian growth theory, the actual rate of growth may display either weak path dependence (sensitivity to initial conditions) or strong path dependence (when the growth path affects the conditions and hence the position of equilibrium). When growth is subject to strong path dependence, the experience of a particular (equilibrium or disequilibrium) growth trajectory can induce discrete structural change associated with the economy's technology and/or institutions, as a result of which the economy will evolve through a series of discrete regimes or episodes of growth – and a prior episode of relatively rapid growth may not be indefinitely self-reinforcing, so that the growth process so described need not suffer from 'too much cumulation'. The natural rate of growth is also path dependent in Kaldorian growth theory, although in and of itself this does not resolve important questions about the *sustainability* of any growth regime characterized by inequality of the equilibrium and natural rates of growth (the first Harrod problem). As has been shown, however, it is possibility to identify solutions to this sustainability issue. These solutions reconcile the basic Kaldorian vision of growth with precisely the type of balance in the growth process necessary to render growth outcomes sustainable in the long run.

The ELCC approach does not, however, reconcile the growth of the volume of exports with growth in the volume of imports, thus leaving open the possibility of chronic (even widening) current account surpluses or deficits in the long run. This raises new questions about the long-run sustainability of the equilibrium growth rate in the ELCC model, arising from a potential (or even likely) lack of BP equilibrium. As previously noted, reconciling Kaldor's vision of export-led growth with this concern is the subject of BPCG theory, to which we now turn.

### ? STUDY QUESTIONS

1) Why is the manufacturing sector so important in Kaldorian growth theory and how, as a result, might structural change affect the rate of growth?
2) Summarize the relationship between Kaldorian growth theory, the Hicks supermultiplier and the Harrod foreign trade multiplier. How does supermultiplier analysis in Kaldorian theory differ from its Sraffian and Kaleckian counterparts, and why might this be important?

3) Construct a basic model of export-led cumulative causation, and use this model to show how different forms of path dependence (weak and strong) can affect the growth process.
4) Outline the ways in which the actual and natural rates of growth can be reconciled in an ELCC model.
5) What are the policy implications of the ELCC model?
6) What criticisms have been made of the ELCC model? How might Kaldorians respond to the criticisms?
7) Is the Beckerman model, as revived by Boggio and Barbieri, an alternative to a Kaldorian approach, or a better way of representing Kaldor's original ideas about export-led growth? Discuss, making sure to refer to the debate over 'Kaldor's paradox'.

## NOTES

1 The advanced capitalist economies, which began the nineteenth century richer than the rest of the world, have subsequently grown faster than the rest of the world (Maddison, 1991, 2008). As a result, their incomes per capita have diverged (in both absolute and relative terms) from those of the rest of the world over the course of the past two centuries. This pattern of 'forging ahead' and 'falling behind' has contributed to a pattern of increasing global income inequality, with evidence of 'catching up' limited to a few Asian economies (most notably Japan and South Korea) that, since the middle of the twentieth century, have grown faster than the rest of the world, closed the per capita income gap, and so joined the elite club of advanced capitalist economies.

2 Our analysis and discussion in this chapter and the next draw partially on Blecker (2013b) and Setterfield (2013a).

3 Drawing on Kaldor's thinking as summarized in his Mattioli lectures (Kaldor, 1996), King (2010, pp. 165–9) identifies the third strand of modern Kaldorian growth theory as a two-sector North–South model in which the terms of trade between primary commodities and manufactured goods are central to the analysis. We agree that the Kaldorian framework encompasses important applications or extensions to North–South trade, as also recognized by Harcourt (2001, pp. 247–51), Skott (1999) and Cimoli and Porcile (2014), among others. Nevertheless, we think that Kaldor's recognition of the importance of international trade in manufactures and primary commodities is part of a broader emphasis he placed on structural change. A full treatment of North–South trade models (both Kaldorian and other) would be beyond the scope of this book, but some aspects of such work will be covered in Chapters 9 and 10. See also Dutt (1990, 2002) for North–South trade models that combine various types of heterodox modelling approaches.

4 In this respect, both views are at odds with other heterodox perspectives on growth discussed earlier in this book, which put more emphasis on domestic demand and in particular investment spending as the critical determinant of demand-led growth.

5 These laws have been renumbered from Thirlwall's list for our purposes, but are otherwise quoted verbatim from his summaries. Law 3 is taken from the shorter statement on p. 354 of Thirlwall (1983); the others are taken from pp. 345–7.

6 See section 8.2.3 and McCombie et al. (2003) for more extensive discussion of Verdoorn's law.

7 The multi-sector variant of BPCG theory presented in Chapter 9 is one way of achieving this disaggregation and restoring to Kaldorian growth theory the emphasis on economic structure and the manufacturing sector that is absent from the canonical one-sector model of ELCC. Chapter 10 will briefly discuss other efforts to incorporate structural change into BPCG models.

8 See Szirmai (2012) for a survey and additional sources on these advantages.

9 In a rather different theoretical genre, Pasinetti (1981) models structural change in the framework of a Ricardian analysis of 'vertically integrated' sectors. In a similar vein, Pasinetti (1993) connects structural change to the role of 'human learning' in the evolution of economic systems.

10 Here, Rodrik ignores Kaldor's idea (implicit in law 3, above) that productivity growth in services should be an increasing function of productivity growth in manufacturing, although some of his own empirical results in Rodrik (2013) could be taken to support that view (and this feature could be added into his model if desired). Nevertheless, his model supports a Kaldorian view in other respects, as discussed below.

**11** Rodrik is an example of what Lavoie (2014) calls a 'mainstream dissenter': an economist who uses orthodox analytical tools but often adopts unorthodox views that sometimes converge with heterodox perspectives. Rodrik et al. (2016) express scepticism about the importance usually attached to neoclassical fundamentals, especially for promoting the acceleration of growth in the least developed economies.

**12** See Magacho and McCombie (2017) on the notion that Verdoorn's law is consistent exclusively with the notion that the growth process is demand-led.

**13** See, for example, Setterfield (1997b, Chapter 3) for further discussion of these first three processes.

**14** See, for example, Toner (1999, Chapter 6) on the importance of the principle of effective demand in Kaldor's growth schema.

**15** It is noteworthy that, several centuries earlier, Adam Smith argued that one of the 'distinct benefits' of international trade was that

> By means of [foreign trade], the narrowness of the home market does not hinder the division of labour in any particular branch of art or manufacture from being carried to the highest perfection. By opening a more extensive market for whatever part of the produce of their labour may exceed the home consumption, it encourages them to improve its productive powers, and to augment its annual produce to the utmost, and thereby to increase the real wealth and revenue of the society. (Smith, 1776 [1976], pp. 468–9)

This emphasis on the dynamic effects of the expansion of trade was, of course, largely forgotten after Ricardo (1821 [1951]) shifted the focus of international trade theory to static efficiency gains based on comparative advantage. It was, however, revived by some of the early post-World War II development economists such as Myrdal (1957), whose work can also be seen as an important precursor of modern Kaldorian growth theory. See also Elmslie (1994) and Blecker (1997b) on the developmental implications of Smith's theory of international trade.

**16** See Palumbo (2009) for further discussion of Kaldor's treatment of consumption, investment and exports.

**17** We will explore the substance of this sentiment in more detail in sections 8.3 and 8.4 when discussing cumulative causation and the path dependence of the growth process.

**18** See Blecker (1997a) for a post-Keynesian perspective on the related Feldstein–Horioka puzzle (saving–investment correlation) in international finance and Palumbo (2009) for further discussion of Kaldor's thinking. The assumption of private sector balance ($1 - c = a_1 y$) is difficult to reconcile with the record of advanced capitalist economies such as the US over the last three or more decades, where sectoral balance analysis consistently reveals a private sector in deficit ($1 - c < a_1 y$). Recall, however, that the thrust of the analysis above is simply to show that export-led growth *may* be consistent with balanced trade – that it does not demand that all countries run BP surpluses, which is impossible. This, as we can now see, requires private sector balance in economies without an active public sector. The fact that the US has not witnessed private sector balance in recent decades suggests only (setting aside the fiscal position of the public sector) that it cannot have experienced balanced trade – which is, in fact, exactly what the historical record shows. It does not demonstrate fault with the analysis above, although it does call into question the veracity of Kaldor's own thinking that $1 - c = a_1 y$ can be regarded as a stylized fact of advanced capitalist economies.

**19** See also McCombie (1985) on the relationships between Hicks's supermultiplier, Harrod's foreign trade multiplier and Kaldorian models of growth.

**20** Note that it follows from (8.17) that, in this case, $k_X = 1$ in equation (8.9).

**21** The view that trade (specifically exports) can drive long-run growth *without* creating external imbalances is properly formalized in the BPCG model developed in Chapter 9.

**22** In general, the actual rate of growth so described is usually an *equilibrium* rate of growth, but processes of cumulative causation need not give rise to equilibrium outcomes.

**23** There are, in fact, important analytical differences between the ELCC model of Beckerman (1962) and what is identified here as the canonical Dixon–Thirlwall ELCC model. These differences and their implications for empirical research will be discussed in section 8.7.3 below.

**24** The formal structure of the Dixon–Thirlwall model is actually that of a traditional equilibrium model, in which the equilibrium rate of growth is defined and reached independently of the adjustment path taken towards it. It may thus appear to be at odds with the importance placed on path dependence in ELCC theory. But in fact, suitably extended, the Dixon–Thirlwall model provides a good vehicle for discussing growth as a path-dependent process, as will be demonstrated later in this chapter.

**25** Although price elasticities are, in principle, negative, we define them in this chapter and the next two chapters as positive, that is, they should be regarded as the absolute values of the 'true' negative elasticities.

**26** This is essentially the rate of real depreciation of the home currency, on the simplifying assumption of a fixed nominal exchange rate.

**27** However, this specification implicitly assumes that export supply is infinitely elastic. See Chapters 9 and 10 for critical discussion of this assumption about the structure of the export market.

**28** See Chapters 9 and 10 for versions of equation (8.20) in which the markup rate is not fixed.

**29** Implicitly, this means that the model presented in this section should be applied mainly to industrialized or semi-industrialized countries (or regions), and should be used with caution (or suitably modified, for example by incorporating structural change) in other contexts.

**30** Kaldor himself regarded constant long-run wage relativities between regions as a stylized fact, but for our purposes it need only be regarded as a simplifying assumption.

**31** Note, then, that consistent with the description of cumulative causation in Figure 8.2, the influence of supply on demand in the Dixon–Thirlwall model assumes that some importance attaches to cost competition in international trade. This is not a necessary feature of the model, however (nor of the ELCC model more generally). Its essential structure – the two-way interaction of supply and demand conditions – would remain unchanged if we were to assume constant relative prices ($\hat{P}_f = \hat{P}$) and that productivity growth enhances the *quality* of goods, and hence their *non-price competitiveness*, and hence the income elasticity of demand for exports ($\eta_X$). See, however, Carlin et al. (2001) and Boggio and Barbieri (2017) for evidence that unit labour costs are, in fact, an influence on export competitiveness, so that the latter does involve at least a component of price competitiveness.

**32** Recall from Chapter 1 that there is evidence in the historical record that growth rates fluctuate in the long run. But there is no evidence that either output or productivity growth rates rise continuously in the long run. The possibility that the ELCC model produces 'too much cumulation' even in the stable equilibrium case depicted in Figure 8.2 is addressed in section 8.7.2.

**33** León-Ledesma uses data for 17 countries averaged over four time periods between 1965 and 1994 and employs two- and three-stage least squares to solve identification problems (avoid simultaneity bias).

**34** Note that in Figure 8.4, $d \ln Y^A/dt = y^{A^*} > y^{B^*} = d \ln Y^B/dt$, where $t$ is time, which is consistent with the results in Figure 8.3.

**35** Harcourt's (1992, pp. 12–13) 'wolf-pack analogy' provides a useful metaphor for the tendency for income divergence that results from cumulative causation. As wolves break away from the pack, so forces are set in motion that allow them to get further and further ahead. This contrasts with a situation in which breakaway wolves are subject to forces that swiftly return them to the pack.

**36** Setterfield (1997b, p. 6) defines the traditional equilibrium approach to economic analysis 'as one in which the long-run or final outcomes of economic systems . . . are both defined and reached without reference to the (historical) adjustment path taken towards them'.

**37** The significance of this possibility is reinforced if the 'data' defining the equilibrium are understood to derive from relatively enduring but ultimately transmutable institutions, as in the model developed by Setterfield and Cornwall (2002). See section 8.4.3 below for further discussion.

**38** That the influence of initial conditions is strictly self-reinforcing can be demonstrated by differentiating the expression for $y$ in equation (8.29) with respect to $y_0$, from which we obtain

$$\frac{\partial y}{\partial y_0} = (k_X \rho \varepsilon_X)^t > 0$$

**39** For a related model of a unit root in a neo-Kaleckian model with cumulative causation features, see Dutt (2006a).

**40** The qualitative result reported below – that the existence of a unit root ensures that initial conditions always matter in the growth process – is unaffected by this second assumption, which is introduced only for purposes of simplicity. To see this, note that the assumption of a unit root transforms equation (8.29) into

$$y = y_0 + t[k_X(\eta_X - \rho_f \varepsilon_X) y_f]$$

where $t$ indicates the time period, from which it is evident by inspection that initial conditions *always* affect subsequent growth outcomes, regardless of the values of other parameters.

**41** Setterfield (2002, p. 227) identifies strong path dependence with hysteresis, on the basis that structural change is the *sine qua non* of hysteresis. The term hysteresis is, however, used in various different ways in economics – including that of a label for the unit root processes discussed earlier – and as such, is avoided altogether here. See Setterfield (2009) for fuller discussion of hysteresis.

**42** Figure 8.2 is thus analogous to what Robinson (1956, pp. 59, 66–7) describes as a 'state of tranquillity' – a special case where an innately historical process generates outcomes akin to those of a mechanical equilibrium process. See also Harris (1991, 2005).

**43** See Colander (1999) for the origins of this metaphor.

**44** In keeping with the durability of institutions (which, in turn, gives rise to the episodic nature of growth), such change will be discrete and discontinuous.

**45** The analysis that follows was inspired by, and is in part based upon, a conversation of Mark Setterfield with Wendy Cornwall that took place in August 2008.

**46** The domestic impact on the US economy itself can also be captured by the variant of the model developed in this chapter that is used by Setterfield and Cornwall (2002). For the sake of simplicity, this exercise is not pursued here.

**47** In addition, as discussed in Blecker (2013a), there was also a shift in the composition of US goods supply towards imports because of the globalization of production, that is, offshoring, tariff reductions, the creation of global value chains and so on, all of which was encouraged in part by a prolonged period of dollar overvaluation in the late 1990s and early 2000s. Thus, not only were US consumers spending more in those decades, but a greater proportion of what they were purchasing was imported (or had major imported components). However, this only amplifies the reasons to believe that $\eta_X$ increased for other countries exporting to the US market in that period.

**48** See, for example, Palley (2002a) and Godley and Izurieta (2002) for anticipations of this unsustainability that, in tandem with the discussion above, focus on the likely consequences for the aggregate-demand-generating process. Note that, in what follows, the shift in DR to DR″ in Figure 8.7 is hypothesized to have resulted from the *exhaustion* and subsequent collapse of a growth episode, rather than from institutional change induced by cumulative experience of the growth outcomes associated with the episode (and hence strong path dependence). In this sense, there is an important qualitative difference between the account provided above of the *rise* of the financialized US growth regime (which *does* involve appeal to strong path dependence based on institutional change induced by macroeconomic performance during the previous growth episode), and the account of the regime's subsequent *decline*.

**49** However, the extent of any such reduction in $\eta_X$ was surely attenuated by the deindustrialization of the US economy, which has hollowed out the US manufacturing industries that produce import-competing goods. As a result, although US households' expenditures are likely to be more constrained by their income in the post-crisis period, the manufactured consumer goods that they purchase still consist more of imports (or imported components) than they did in the past – partly as a hysteresis effect of past episodes of dollar overvaluation (on which see Setterfield and Ozcelik, 2018).

**50** For a variety of perspectives on the theory and reality of secular stagnation in the US economy and globally, see (among many others) Summers (2014), Backhouse and Boianovsky (2016), Blecker (2016b), Hein (2016), Bivens (2017) and Cynamon and Fazzari (2017b).

**51** A long-standing theme in heterodox growth theory (HGT), which goes back at least as far as Marx's alternative to Malthus's theory of the labour supply as discussed in Chapter 2, is that the growth rate of the labour force is also endogenous to the actual rate of growth. On this view, the rate of growth of the labour force adjusts endogenously to meet the needs of a growing capitalist economy through intersectoral and/or inter-regional migration of a global 'reserve army' of labour (see also Cornwall, 1972, 1977). Another possibility is that labour force participation is positively affected by the actual rate of growth, as has been observed in the aftermath of the financial crisis and Great Recession when the labour force participation rate has decreased. The possibility of endogenous labour force growth is overlooked here for the sake of simplicity. See also León-Ledesma and Thirlwall (2000, 2002) and León-Ledesma and Lanzafame (2010) for empirical evidence of the endogeneity of the natural rate of growth.

**52** Note that the employment rate is also bounded below – it cannot be less than zero – so an equilibrium growth outcome that involves $l^* < n$ will also raise a problem of unsustainability similar to that identified above.

**53** In contrast, the rise in the income elasticity of demand for imports reduces the BP-equilibrium growth rate (not the actual growth rate) in Palley's version of the BPCG model, as discussed more explicitly in section 10.4 in Chapter 10.

**54** Recall from equation (8.36) that the natural rate of growth varies positively with the Verdoorn coefficient. In the original version of Setterfield (2006b), the natural rate of growth adjusts to equal the equilibrium growth rate from the BPCG model, as explained in section 10.4 in Chapter 10.

**55** The reader should note, however, that the joint interaction of demand and supply conditions that is central to the ELCC model does *not* preclude supply-side policies – which could include industrial policies, education and training, and government support to R&D – from contributing to growth. The focus on demand-side policies is therefore relative rather than absolute.

**56** If we relax the assumption that $E$ is constant and assume instead that $\hat{E} > 0$, equation (8.19) is modified in a manner that alters the reduced form expression in equation (8.24) and, ultimately, raises the value of $\Omega$ in the DR in (8.24). Proof of this is left to the interested reader.

**57** Equation (8.9′) is consistent with a restatement of total output (from equation 8.10) as $Y = C + I + A + (X - M)$, where $A$ denotes the level of the exogenous component of domestic demand. Equation (8.14) – our expression for the Kaldorian supermultiplier – then becomes

$$Y = \frac{1}{1 - (c + a_1 y) + v}(A + X).$$

Note that, despite the inclusion of $A$ (in addition to $X$) as a component of autonomous demand, the multiplier in equation (8.9′) is still $k_X$. This is because we are still assuming that $1 - (c + a_1 y) = 0$, as a result of which the Kaldorian supermultiplier still reduces to $1/v = k_X$.

**58** In other words, the model rests on the dual economy assumption first introduced in Chapter 1.

**59** The assumption that $E$ is fixed can be relaxed and replaced with the assumption that the rate of nominal exchange rate depreciation ($\hat{E}$) remains constant, without having any fundamental effect on the properties of the ELCC model discussed up to this point. This, however, does not address the criticism raised here, that the ELCC lacks an explicit description of exchange rate dynamics.

**60** See Setterfield (1997b, Chapter 4) for an extension of the Dixon–Thirlwall model that accommodates this insight.

**61** Recall that we defined this elasticity to be positive, so that a real depreciation of the exchange rate or increase in the relative price of foreign goods leads to increased export demand, as shown in equation (8.19) or (8.22). In some of the empirical studies related to Kaldor's paradox, the relative prices or price elasticities are defined differently so that the 'right' sign may be negative.

**62** He instead endorsed the BPCG model of Thirlwall (1979) which, as will be demonstrated in the next chapter, allows no role for cost competitiveness in the determination of long-run, export-led growth, emphasizing instead the importance of non-price factors (advertising, product quality, after sales service and so on) in international competitiveness – at least, not in the standard or canonical formulation.

**63** Since León-Ledesma (2002) uses the rate of change in prices of home goods ($\hat{P} - \hat{P}_f$ in our notation), the elasticity he estimates is essentially $-\varepsilon_X$ and his estimated coefficient is negative and statistically significant.

**64** These authors also draw on empirical work by evolutionary/Schumpeterian economists including Verspagen (1993), Amendola et al. (1993), Amable and Verspagen (1995) and Verspagen and Wakelin (1997), who focus more on technology variables but who also test the importance of relative cost variables – in levels – as a determinant of international trade.

**65** For a sceptical view of the importance of relative prices or the real exchange rate and some contrary empirical evidence, see Ribeiro et al. (2018). However, as noted earlier, León-Ledesma (2002) found that rates of change in relative prices do affect export growth significantly in a more complex version of a Dixon–Thirlwall model, including many control variables and using simultaneous equations methods, which suggests that the evidence that led Kaldor to express his 'paradox' was not robust. The same issue about the importance of levels versus growth rates of relative prices (or the real exchange rate) arises in the BPCG model, as will be discussed in Chapter 10.

**66** Although this literature generally refers to 'balance of payments' equilibrium, it is clear from the context that what is really meant is balance on current account. Furthermore, the models usually ignore all other components of the current account besides trade in goods and services.

# Appendix 8.1 Absolute and relative income divergence due to cumulative causation

Assume, as in the text, that with $Y^A > Y^B$ initially, $y^{A^*} > y^{B^*}$. That the difference between $Y^A$ and $Y^B$ will grow over time in *absolute* terms becomes clear if we define the difference between these income levels at any point in time as

$$Gap = Y^A - Y^B = Y_0^A e^{y^{A^*}t} - Y_0^B e^{y^{B^*}t}$$

where $e$ is the base of the natural logarithm. It follows that

$$dGap/dt = Y_0^A e^{y^{A^*}t} y^{A^*} - Y_0^B e^{y^{B^*}t} y^{B^*} > 0$$

since both $Y_0^A > Y_0^B$ and $y^{A^*} > y^{B^*}$ by hypothesis.

That economy $A$ also becomes richer in *relative* terms can be demonstrated by first defining the difference between the log levels of $Y^A$ and $Y^B$ as the relative gap

$$RelGap = \ln Y^A - \ln Y^B = \ln(Y^A/Y^B)$$

It then follows, by inspection of Figure 8.4, that *RelGap* and hence the (log) level of income in economy $A$ relative to economy $B$ is increasing over time.

## Appendix 8.2 Formal analysis of Figure 8.9

The curves depicted in Figure 8.9 are based on the facts that, from equation (8.42)

$$\frac{dy_{nc}}{d\rho} = \frac{q_0 + n}{(1 - \rho)^2} > 0$$

and

$$\frac{d^2y_{nc}}{d\rho^2} = \frac{2(q_0 + n)}{(1 - \rho)^3} > 0$$

while, from (8.34)

$$\frac{dy^*}{d\rho} = \frac{\varepsilon_X k_X^2(\eta_x - \rho_f\varepsilon_X)y_f}{(1 - k_X\rho\varepsilon_X)^2} > 0$$

and

$$\frac{d^2y^*}{d\rho^2} = \frac{2\varepsilon_X^2 k_X^3(\eta_x - \rho_f\varepsilon_X)y_f}{(1 - k_X\rho\varepsilon_X)^3} > 0$$

Note also that $\lim_{\varepsilon_X \to 0}(dy^*/d\rho) = 0$, so a small enough value of $\varepsilon_X$ (the price elasticity of demand for exports) is sufficient to ensure that $dy^*/d\rho < dy_{nc}/d\rho$ (as depicted in Figure 8.9), thus ensuring the stability of the system as a whole. See, for example, McCombie and Thirlwall (1994) for discussion of the inelasticity of trade to price competition in the context of Kaldorian growth theory.

# 9

# Balance-of-payments-constrained growth I: Thirlwall's law and extensions

## 9.1 Introduction

The model of export-led growth with cumulative causation (ELCC) covered in the previous chapter implies that – under certain conditions – some countries can achieve ever-widening 'virtuous circles' of faster productivity growth, improving competitiveness, rising exports and rapid output growth, while other countries are doomed to suffer 'vicious circles' of slower productivity growth, worsening competitiveness, stagnant exports and sluggish output growth. However, these models do not consider the fact that faster growth of national income is likely to lead to more rapid increases in imports, which can put a strain on a country's balance of payments (BP). The original Kaldorian export-led growth (ELCC) model ignores the role of imports in counterbalancing exports and does not impose the restriction that the current account of the BP must be balanced in the long run. Based on this critique, Thirlwall (1979) developed an alternative model of growth in an open economy – also situated within the Kaldorian tradition – that has become known as the theory of 'balance-of-payments-constrained growth' (BPCG) or 'Thirlwall's law'.[1]

In Thirlwall's approach, virtuous circles based on rapid export growth with cumulative causation may be impossible to sustain in the long run because the resulting rapid growth of national income could make a country's imports rise too fast to be compatible with equilibrium in the BP. If rapid income growth generates rising BP deficits, this forces adjustments in domestic expenditures that eventually limit the growth of output (and income) to a rate that is consistent with BP equilibrium (Thirlwall and Dixon, 1979). According to McCombie and Thirlwall (1999, p. 49), 'We mean by the term balance-of-payment constraint that a country's performance in overseas markets, and the response of the world financial markets to this performance, constrains the rate of growth of the economy to a rate which is below that which internal conditions . . . would warrant.' This view is elaborated by Thirlwall and Hussain (1982) as follows:

> for most countries the major constraint on the rate of growth of output is likely to be the balance of payments position because this sets the limit to the growth of demand to which supply can adapt. Most countries, apart from the oil producing countries of the Middle East, can absorb foreign exchange without difficulty; and most cannot earn enough. It is true, of course, that the world as a whole cannot be balance of payments constrained, but it only requires one country or bloc of countries not to be so constrained, for all the rest to be so. There cannot be many less-developed countries that could not utilise resources more fully given the greater availability of foreign exchange. (Thirlwall and Hussain, 1982, p. 498)

The BPCG and ELCC approaches do coincide in certain respects, and both have Kaldorian roots. Both maintain the post-Keynesian belief that aggregate demand is paramount in determining a nation's growth, even in the long run, and see the demand-side constraints for most countries as lying primarily in the international domain rather than the domestic economy.[2] Both agree that increasing the growth rate of exports is key to raising a country's long-run growth rate of output, but for different reasons and with different causal mechanisms. The BPCG model stresses the need for exports to provide the foreign exchange earnings needed to pay for imports without running trade (current account) deficits, instead of the potential for cumulative causation in export performance and productivity growth emphasized in the ELCC model. The latter model focuses on changes in relative cost competitiveness driven by endogenous technological progress as driving export success (or failure), whereas BPCG asserts that such changes either dissipate in the long run (as relative prices remain constant on average in the long run) or else have small effects on trade flows (so-called elasticity pessimism). Thus, qualitative competitiveness matters in both theoretical approaches, but cost competitiveness and real exchange rates (RERs) matter only in the ELCC model.[3] In contrast, the BPCG approach puts special emphasis on the income elasticity of import demand, which determines how much imports increase in response to faster growth of output, and is thus inversely related to the growth rate that is consistent with BP equilibrium.

This is the first of two chapters on the BPCG model. This chapter begins by presenting the most basic version of the BPCG model, including the standard solutions for Thirlwall's law and its policy implications, in section 9.2.

Section 9.2 also discusses some of the key assumptions of this approach and contrasts Thirlwall's model with a neoclassical alternative. Section 9.3 considers extensions of the basic model that incorporate international capital flows, structural change and repercussion effects (the last in a model with two large countries). Section 9.4 discusses how partial pass-through of exchange rate changes into prices of traded goods and endogenous productivity growth in the form of Verdoorn's law can be incorporated into extended versions of the BPCG model. Section 9.5 concludes. Various critiques of the BPCG approach as well as several alternative models and efforts to reconcile it with other theoretical approaches will be considered in Chapter 10.

## 9.2 Thirlwall's law

### 9.2.1 The basic Thirlwall model

In this section, we consider the most basic version of the BPCG model first articulated by Thirlwall (1979). The basic model assumes that BP equilibrium is the most salient constraint on long-run growth in an open economy, because in the long run a country's trade must be balanced on average (cases with sustained trade imbalances financed by net capital flows will be considered in the following section). In effect, the model assumes that there are two goods: a domestically produced good that can be either purchased at home or exported and a foreign produced import good. Assuming that the two goods are imperfect substitutes for each other, the 'law of one price' does not apply (the goods may sell at different prices, measured in the same currency, and their *relative* prices may affect the demand for home versus foreign products).[4]

The model also assumes that supplies of exports and imports are infinitely elastic, so that the quantities traded are uniquely determined by the demand for each. Using the standard constant-elasticity form for mathematical convenience, the export demand function is given by

$$X = X_0\left(\frac{EP_f}{P}\right)^{\varepsilon_X} Y_f^{\eta_X} \tag{9.1}$$

where $X$ is the quantity of exports, $X_0$ is a positive constant, $E$ is the nominal exchange rate (in home currency per unit of foreign currency), $P_f$ is the 'foreign' (or rest-of-world) price level in foreign currency, $P$ is the home price level (in domestic currency), $Y_f$ is foreign (world) income, and $\varepsilon_X$ and $\eta_X$ are the price and income elasticities of export demand, respectively (defined so that $\varepsilon_X, \eta_X > 0$). All variables (except $X_0$) are understood to be functions of time, but time subscripts or parentheses are omitted to avoid notational clutter.[5] Note that $EP_f/P$ is the RER or relative price of foreign goods (that is, how much home goods have to be given up to buy foreign goods), so a rise (fall) in this ratio implies a real depreciation (appreciation) of the home currency. Intuitively, equation (9.1) says that exports increase when foreign goods become more expensive (home goods become relatively cheaper) and when foreign income rises.

Similarly, import demand is given by

$$M = M_0\left(\frac{EP_f}{P}\right)^{-\varepsilon_M} Y^{\eta_M} \tag{9.2}$$

where $M$ is the quantity of imports, $M_0$ is a positive constant, $Y$ is home country national income, and $\varepsilon_M$ and $\eta_M$ are the price and income elasticities of import demand, respectively (defined so that $\varepsilon_M, \eta_M > 0$). According to equation (9.2), imports increase when foreign goods become relatively cheaper compared with domestic goods and also when domestic income rises.

In the simplest model, in which there are no net capital (financial) flows in the long run (also no transfers), BP equilibrium requires balanced trade in goods and services, that is, the value of exports must equal the value of imports, measured in a common currency:[6]

$$PX = EP_fM \tag{9.3}$$

If we take natural logarithms of this equilibrium condition and differentiate with respect to time, we can convert it into growth rate form. Following our practice in earlier chapters, we use a lower-case letter to represent the growth rate of the corresponding quantity variable and a circumflex (ˆ) to represent the rate of increase in a nominal variable. Thus, the equilibrium condition in growth rate form can be written as

$$\hat{P} + x = \hat{E} + \hat{P}_f + m \tag{9.4}$$

Similarly, converting the export and import demand functions (9.1) and (9.2), respectively, into growth rate form yields (note the constants disappear since they are not functions of time):

$$x = \varepsilon_X(\hat{E} + \hat{P}_f - \hat{P}) + \eta_X y_f \tag{9.5}$$

$$m = -\varepsilon_M(\hat{E} + \hat{P}_f - \hat{P}) + \eta_M y \tag{9.6}$$

where $\hat{E} + \hat{P}_f - \hat{P}$ is the rate of real depreciation of the home currency. Then, substituting equations (9.5) and (9.6) into (9.4) and rearranging, the condition for maintaining balanced trade in the long run can be expressed as:

$$(\varepsilon_X + \varepsilon_M - 1)(\hat{E} + \hat{P}_f - \hat{P}) - \eta_M y + \eta_X y_f = 0 \qquad (9.7)$$

In this expression, the term $(\varepsilon_X + \varepsilon_M - 1)$ will be positive if the Marshall–Lerner (ML) condition $(\varepsilon_X + \varepsilon_M > 1)$ is satisfied; this is essentially the condition required (under certain simplifying assumptions) for a real depreciation of the currency to improve the trade balance, as explained in Appendix 9.1. Whether this condition is normally satisfied is the subject of much debate, as we will discuss below.

Now, the obvious question is, which variable(s) adjust to make this equilibrium condition hold in the long run? Following Thirlwall, we assume here that the price and income elasticities ($\varepsilon_i$ and $\eta_i$, $i = X, M$) are exogenously given and remain constant over long periods of time (we will discuss alternative views later in this chapter and the next one). We also assume that foreign income growth ($y_f$) is exogenously given, which requires that the country is too small to have appreciable 'repercussion effects' on rest-of-world income (this assumption will be relaxed later in this chapter).[7] On these assumptions, the only two possibilities are that either domestic income growth ($y$) or the rate of change in the RER $(\hat{E} + \hat{P}_f - \hat{P})$ – or some combination of the two – must adjust to satisfy (9.7).

Thirlwall's law is based on the Keynesian assumption that income or output is the adjusting variable, not relative prices or the RER (an alternative neoclassical solution will be discussed in section 9.2.3 below). Thus, taking the rate of change in the RER or relative prices $(\hat{E} + \hat{P}_f - \hat{P})$ as exogenously given (remember that this would be the long-run average trend rate of change), we can solve (9.7) for the growth rate of domestic income (output) that maintains balanced trade in the long run:

$$y_B = \frac{(\varepsilon_X + \varepsilon_M - 1)(\hat{E} + \hat{P}_f - \hat{P}) + \eta_X y_f}{\eta_M} \qquad (9.8)$$

where $y_B$ can be called the BP-constrained growth rate or, following Thirlwall (1979), the 'BP-equilibrium growth rate'.

Furthermore, Thirlwall (1979) and his followers have argued that the relative price (RER) change $(\hat{E} + \hat{P}_f - \hat{P})$ should have a negligible impact in the long run for either one of two reasons. On the one hand, empirical support for the ML condition holding in many countries is mixed at best. If we assume instead what might be called 'elasticity pessimism', we could assert that $\varepsilon_X + \varepsilon_M \approx 1$, which means that the price elasticities of export and import demand are not high enough for a devaluation to improve the trade balance.[8] In this case, the first term (representing relative price effects) in the numerator of (9.8) drops out, so this solution simplifies to

$$y_B = \frac{\eta_X y_f}{\eta_M} \tag{9.9}$$

On the other hand, even if ML is satisfied so that $\varepsilon_X + \varepsilon_M > 1$, relative price effects can also be ruled out if the relative price of foreign and domestic goods (the RER) does not change significantly in the long run, so that we can assume $\hat{E} + \hat{P}_f - \hat{P} = 0$. This would occur if, for example, domestic price changes closely mirror changes in prices of foreign goods converted to domestic currency in an open economy – especially when a nominal depreciation causes (perhaps after some time lag) an offsetting increase in domestic inflation. In this case, equation (9.8) again simplifies to (9.9). However, here yet another, more dramatic simplification is possible. Substituting $\hat{E} + \hat{P}_f - \hat{P} = 0$ into the export function in growth rate form (equation 9.5), the latter reduces to $x = \eta_X y_f$, so equation (9.9) can be rewritten as

$$y_B = \frac{x}{\eta_M} \tag{9.10}$$

The solutions (9.9) and (9.10) constitute the two alternative versions of Thirlwall's law.[9] Following Perraton (2003), we will refer to (9.9) as the 'strong form' and (9.10) as the 'weak form' of this law. In addition, (9.8) can be considered the most general solution for the BP-equilibrium growth rate, but advocates of the BPCG model generally believe that relative price effects should be negligible in the long run for one of the two reasons stated above (either elasticity pessimism or a constant RER in the long run), and therefore either the weak or strong form of Thirlwall's law should prevail. It should also be noted that the solution (9.10) is a clear analogue to the dynamic version of Harrod's (1933) foreign trade multiplier, equation (8.17) in Chapter 8, with the key difference that in Thirlwall's version the parameter in the denominator is the income elasticity of import demand rather than the marginal propensity to import. Hence, one can think of Thirlwall as having generalized Harrod's foreign trade multiplier by allowing for non-unitary income elasticities (since a constant marginal propensity to import assumes $\eta_M = 1$).

Equation (9.10) is remarkable in its stark simplicity. Assuming that a country has to maintain balanced trade and that the relative prices of its products and foreign goods don't change in the long run, the country's long-run average growth rate should equal the ratio of the growth rate of its exports to the income elasticity of its demand for imports. Moreover, the solution for $y_B$ in (9.10) is very easy to estimate empirically for any given country: all one needs to do is to estimate the import demand function (9.6) to obtain an estimate of $\eta_M$ and combine this with the average growth rate of exports $x$ calculated from the country's trade data; one does not need to estimate the export function (9.5) in order to perform this calculation. In contrast, if the solution (9.9) is used, then both the export and import demand functions (9.5) and (9.6) must be estimated econometrically to retrieve the two income elasticities $\eta_X$ and $\eta_M$, and these estimates must then be combined with data on foreign income growth $y_f$ (which of course is necessary for estimating equation 9.5 anyway).

Both the weak and strong versions of Thirlwall's law can be considered to be applications of the supermultiplier concept, discussed in Chapter 7, to an open economy facing a BP constraint. This is most obvious in equation (9.10), where export growth $x$ is the exogenous component of aggregate demand that drives the growth of domestic income consistent with BP equilibrium, $y_B$. Implicitly, the sum of domestic consumption, investment and government expenditures must adjust endogenously to ensure that domestic demand grows at this rate if BP imbalances are to be avoided in the long run. In equation (9.9), it is rather foreign income growth $y_f$ that constitutes the exogenous driving force for domestic demand, but clearly it operates through the export channel, and it should be recalled that exports must grow at the rate $x = \eta_X y_f$ if the RER has no tendency to change (so that $\hat{E} + \hat{P}_f - \hat{P} = 0$) in the long run. So, in effect, both versions of Thirlwall's law depict exports as the exogenous component of aggregate demand that drives growth in a BP-constrained open economy.

## 9.2.2 Key assumptions of the model

Like any economic theory, Thirlwall's law rests on certain key assumptions.[10] First, the structure of the markets for exported and imported goods must be as shown in Figure 9.1: the supply curves are horizontal (infinitely elastic) while the demand curves are downward sloping (thus the price elasticities of demand are finite). Essentially, prices are effectively fixed by cost conditions in the supplier country (home for exports, rest of world for imports), while quantities are purely demand-determined. This type of market structure may be called (following Branson, 1983, p. 48) a 'Keynesian small economy', that is, one which takes the prices of its imports and the demand curve for its exports as given (this will be contrasted with a more classical type of 'small

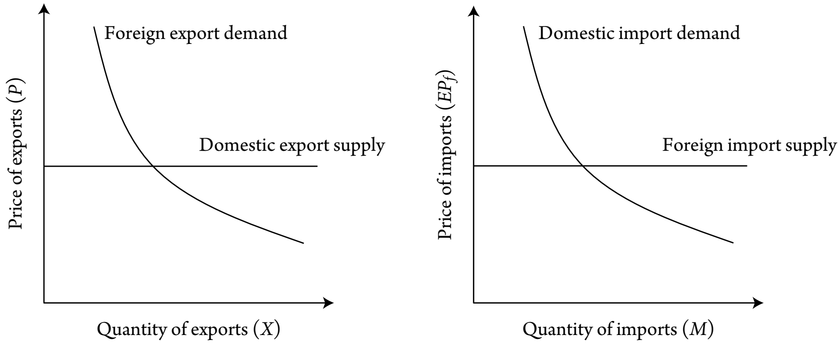



**Figure 9.1** Markets for exports and imports in a 'Keynesian small economy' (infinitely elastic supplies with prices fixed in the seller's currency)

open economy', which is a price-taker in the markets for both imports and exports, in section 10.2.4 in Chapter 10).

This set of assumptions is not controversial for imports, because most countries (except the very largest) are too small to influence the world (foreign) price of their imports and hence do, in fact, take the import price in foreign currency ($P_f$) as exogenously given. However, the assumption of an infinitely elastic supply of exports is more controversial. This assumption is most likely to be accurate when the home country is an exporter of industrial goods for which there are no binding supply constraints (especially no constraint on labour supply, since industrial capacity could be increased in the long run). This could be the case, for example, if a country largely exports manufactured products in industries that have significant excess capacity (à la Kalecki and Steindl), or in which factories can easily be replicated (at constant average cost) as demand expands. However, this assumption is of more doubtful relevance for exporters of primary commodities, for which supplies may be limited by natural resource constraints leading to increasing costs, and it may also be questioned for small developing countries with limited industrial capacity.[11] We will return to this point when we discuss the small country model of Razmi (2016a) in section 10.5.1 of Chapter 10.

Second, the BPCG model makes a number of simplifying assumptions about pricing, which may or may not hold in reality. Most importantly, it assumes that prices are fixed in the seller's currency – the home currency for exports and the foreign currency for imports. However, many export goods (especially primary commodities) are sold in global markets in which prices are set in an international currency like the US dollar. Furthermore, the model assumes that changes in the nominal exchange rate are always fully passed through into the other currency (home currency for imports, foreign currency for exports). However, a large literature has shown that firms may instead 'price to market' and therefore only partially pass through changes in nominal exchange rates into prices in the other currency. Thus, even if manufactured goods are sold based on cost-plus-markup pricing, firms have the option of adjusting their markups – cutting them when the seller's currency appreciates and raising them when the seller's currency depreciates (as assumed in the open economy neo-Kaleckian model in section 4.4.3 of Chapter 4) – which can affect prices of imports as well as exports (for example, if foreign sellers cut their markups in order to prevent import prices from rising too much in home currency terms following a depreciation of the home currency).[12] Finally, the basic model assumes that there is only one price for domestic products regardless of whether they are sold at home or exported; in reality, prices of exports may differ from prices of goods sold at home because of either qualitative differences or price discrimination (that is, charging different prices in different markets, at home and abroad). Alternative specifications of pricing will be considered later in this chapter and the next.

### 9.2.3 A neoclassical solution: comparison and critique

Returning to the most general form of the equilibrium condition in equation (9.7), an obvious alternative solution would be to allow the RER, or relative price of foreign goods, to be the adjusting variable for maintaining BP equilibrium, while taking the growth rate of domestic output as exogenously given.[13] To see how such a neoclassical solution would work, we can solve equation (9.7) for the rate of change in the RER:

$$\hat{E} + \hat{P}_f - \hat{P} = \frac{\eta_M y - \eta_X y_f}{\varepsilon_X + \varepsilon_M - 1} \tag{9.11}$$

In evaluating this expression, it is important to note that the denominator will be positive if and only if the ML condition holds, that is, the sum of the price elasticities (recall these have been defined to be positive) must exceed unity: $\varepsilon_X + \varepsilon_M > 1$ (see Appendix 9.1). Although this condition is less important for Thirlwall's post-Keynesian analysis, with its focus on output (growth) as the adjusting variable, it is vital for the neoclassical solution, with its focus on relative price (RER) adjustment. For purposes of understanding the neoclassical view, therefore, we will stipulate that this condition holds (the empirical debate over whether the ML condition is normally satisfied in reality will be covered in Chapter 10).

In such a neoclassical analysis, it would also be assumed that the growth rates of domestic and foreign income ($y$ and $y_f$) are exogenously given at fixed 'natural rates', defined as in earlier chapters as the growth rate of the labour force plus the growth rate of labour productivity, $y = y_N = n + q$ and $y_f = y_{f,N} = n_f + q_f$, in the long run. Then, assuming that the ML condition holds, a higher natural rate of growth at home ($y_N$) or a higher income elasticity of import demand ($\eta_M$) would require the RER to depreciate more quickly, in order to offset the faster growth of imports that would otherwise be implied and thereby prevent a trade deficit from emerging. In contrast, a higher natural rate of growth abroad ($y_{f,N}$) or a higher income elasticity of export demand ($\eta_X$) would require the RER to depreciate more slowly or to appreciate gradually over time (the latter would occur if $\hat{E} + \hat{P}_f - \hat{P} < 0$).[14]

The contrast between Thirlwall's law and the neoclassical view of exchange rate adjustment can be understood with the help of a diagram. To facilitate this comparison, we will again assume that the ML condition is satisfied (elasticity pessimism does not hold). We can rewrite the solution for the rate of change in the RER that maintains BP equilibrium (equation 9.11) in slope-intercept form as

$$\hat{E} + \hat{P}_f - \hat{P} = \frac{-\eta_X y_f}{\varepsilon_X + \varepsilon_M - 1} + \frac{\eta_M}{\varepsilon_X + \varepsilon_M - 1} y \tag{9.12}$$

which can be graphed as shown in Figure 9.2. To heighten the contrast between the two views, we designate $y_N$ as the exogenous natural rate of growth, and we assume that $y_B < y_N$ so that the country's growth is BP constrained relative to this benchmark. (Of course, as discussed extensively

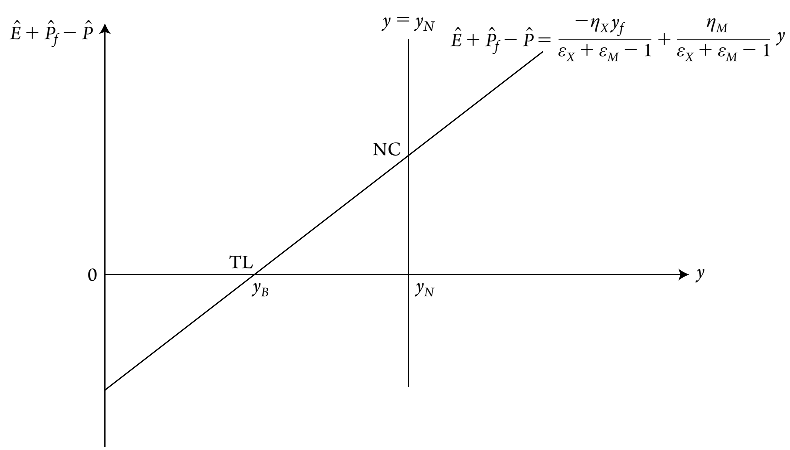



**Figure 9.2** Thirlwall's law (TL) and neoclassical (NC) solutions compared

elsewhere in this book, in reality the supply-side determinants of the natural rate of growth are endogenous, which implies that an exogenously given $y_N$ does not generally exist, but we assume an exogenous $y_N$ here for illustrative purposes; BPCG models in which $y_N$ adjusts endogenously are considered in section 10.4 of Chapter 10.) In Figure 9.2, the Thirlwall's law solution (either version 9.9 or 9.10) occurs at point TL, where $\hat{E} + \hat{P}_f - \hat{P} = 0$ and $y = y_B$, while the neoclassical solution (9.11) occurs at point NC where $y = y_N$ but (in the case shown here) $\hat{E} + \hat{P}_f - \hat{P} > 0$, that is, a continuous real depreciation is required in the long run.

Although Figure 9.2 is intended to represent the two polar views of pure output (growth) adjustment and pure relative price (RER) adjustment, the diagram reveals that in principle these are just two extreme cases, and there is a continuum of output growth rates and relative price changes (represented by the upward-sloping line corresponding to equation 9.12) that are consistent with BP equilibrium in the long run. Essentially, this line represents the possible trade-offs between real currency depreciation and slower output growth in a country that is BP constrained.[15] To be clear, Thirlwall's post-Keynesian view does not require that relative price (RER) effects are absolutely zero, only that they are *relatively small* so that the outcome is *closer* to the point where $y = y_B$ than the point where $y = y_N$. In the words of Thirlwall and Hussain (1982, p. 498, emphasis added), 'it is *largely* real income (and employment) that adjusts to bring the value of imports and exports into line with one another to preserve balance of payments equilibrium' (note their use of 'largely' rather than 'entirely').

Moreover, there are many reasons to question the validity of the neoclassical solution. First, as discussed in earlier chapters, there are many theories of long-run growth in which economies either do not necessarily grow at their so-called natural rates, or else the 'natural rate of growth' itself is endogenous because the growth rates of labour supply and/or labour productivity can be endogenous – including in response to demand conditions. Thus, we cannot generally assume that the long-run growth rates (both home and foreign) are fixed independently of demand conditions – including the impact of BP constraints – so that they could be used to determine the long-run trend in the RER as in equation (9.11). Nevertheless, for any given country that is too small to affect the growth of the rest of the world, actual foreign growth $y_f$ can be taken as exogenously given by that country, and hence its own growth rate $y$ could be the variable that has to adjust in order to maintain BP equilibrium as specified in equation (9.7). Second, neither the inflation rates ($\hat{P}$ at home and $\hat{P}_f$ abroad) nor the nominal depreciation of the exchange rate ($\hat{E}$) will necessarily adjust in such a way as to guarantee balanced trade in the long run, if these are driven by factors such as domestic labour market conditions, firms' pricing policies, international financial flows and the central bank's monetary policies. As discussed in more depth in Chapter 10, the empirical study by Alonso and Garcimartín (1998–99) found that relative prices of exports and imports do not adjust significantly in the 'right' direction in response to trade imbalances, but output growth rates do.

The comparison of Thirlwall's law with a neoclassical alternative in Figure 9.2 does reveal one potential 'pitfall' (in the phrase of Palley, 2002c) in the BPCG approach, which is the apparent conclusion that a country could have persistent growth at a BP-equilibrium rate that is significantly different from its natural rate of growth (which, as will be recalled, is the rate consistent with full employment of labour or a constant unemployment rate). For example, in the case shown in Figure 9.2, $y_B < y_N$, which implies that the country would have perpetually increasing unemployment in the long run. Conversely, $y_B > y_N$ would imply that employment would be growing so rapidly that the labour force would eventually be exhausted. Alternative mechanisms that could reconcile these two long-run equilibrium growth rates, as proposed by Palley (2002c), Setterfield (2006b), Oreiro (2016) and others, will be discussed in Chapter 10 (section 10.4). But for the moment, we can simply 'flag' the fact that the potential discrepancy between the BP-equilibrium growth rate and the natural rate of growth adds a third sort of Harrodian 'problem' to the two problems noted in Chapter 3 (possible divergences between the natural and warranted growth rates, and between the warranted and actual growth rates).

## 9.2.4 Policy implications

In spite of their stark simplicity, the basic Thirlwall's law solutions have powerful policy implications. First and foremost, Thirlwall's model stresses the importance of exports in the growth process, but for very different reasons from the ELCC models covered in Chapter 8. In the latter models, export growth is important because it is a critical ingredient in a process of cumulative causation in which growing sales to external markets make it possible for a country to increase its productivity faster through dynamic economies of scale, and thereby to achieve ever-widening competitive advantages. However, this type of cumulative causation is ruled out in either solution for Thirlwall's law (equation 9.9 or 9.10) by the assumption that relative prices do not matter in the long run, which means that faster growth of productivity either does not translate into sustained advantages in cost competitiveness (if relative prices are constant in the long run) or alternatively – even if the country does become more competitive – there is no resulting gain because the quantity responses of exports and imports to a lower relative price (real depreciation) are too small (elasticity pessimism).

Rather, the reason why exports are so important in Thirlwall's approach is because they are vital to offset the otherwise constraining impact of rising import demand when an economy grows faster – an effect that is more severe, the higher is the income elasticity of imports $\eta_M$. Indeed, many empirical studies tend to find that $\eta_M > 1$ in most countries. This implies that when countries attempt to speed up their growth through domestic means, they are likely to need additional imports more than proportionately to the rise in income, hence implying growing trade deficits unless exports rise fast enough to prevent that from happening. This problem is especially acute in less developed countries, which may lack domestic sources for key products such as capital equipment and high-technology goods, or which may face constraints of inadequate resources or institutions that limit domestic supplies of basic food and energy products. Less developed countries, in particular, typically face a binding foreign currency constraint, since they must earn foreign exchange in order to pay for necessary imports (and to service international debts). But the same problems may also be found in more advanced or emerging economies that have large appetites for imported goods. As Thirlwall (2002, p. 53) states:

> It may be possible to initiate a consumption-led, investment-led or government expenditure-led growth, for a short time, but each of these components of demand has an import component . . .. If there are no export earnings to pay for the import content of other components of expenditure, demand will have to be constrained.

Second, the Thirlwall approach blends demand-side and supply-side determinants of growth in a particularly compelling way for open economies. On the one hand, it is evidently demand-side constraints that must be envisioned to keep actual output growing at the BP-equilibrium rate ($y_B$) when a higher growth rate would invite a rising trade deficit (as, for example, if output grew at the rate $y_N$ shown in Figure 9.2). These constraints could be imposed, for example, by contractionary fiscal policies that are introduced in response to increased trade (current account) deficits, as found for example by Summers (1988) and Epstein and Gintis (1992) – although Artis and Bayoumi (1990) found contrary evidence. The Thirlwall model fits into the post-Keynesian approach of demand-driven growth broadly defined, but instead of focusing on the domestic constraints imposed by firms' investment demand (as in the neo-Robinsonian and neo-Kaleckian models), it focuses on the external constraints imposed by a nation's export performance relative to its propensity to import. Also, Thirlwall's law (explicitly in the strong form and implicitly in the weak form) implies that growth in an open economy is strongly affected (on the demand side) by the growth of the foreign countries that constitute any given nation's main export markets.

On the other hand, supply-side factors are also implicit in Thirlwall's law. As mentioned above, one reason why a country may have a high income elasticity of demand for its imports could be that it has limited domestic capacity to produce certain key goods (whether consumption, investment or intermediate goods) – hence, relieving those supply constraints by increasing domestic capacity, eliminating domestic bottlenecks or improving domestic institutions could help to lower $\eta_M$ and thereby increase $y_B$ (see Cimoli and Porcile, 2014). On the export side, the solution (9.9) – which applies regardless of whether constant relative prices or elasticity pessimism is assumed – makes it clear that it is vital for a country to have a high foreign income elasticity of demand for its exports, $\eta_X$. For this purpose, what is essential is the composition of a nation's exports – do they consist of staple foods and basic industrial goods (for example, cotton textiles) for which income elasticities are low, or do they consist of advanced manufactures (for example, electronic products, transportation equipment) for which income elasticities tend to be higher? In order to ensure that exports are of the latter type, countries must not only invest in the capital stock required to produce them, but also must pay attention to other supply-side 'inputs' such as education, research and development (R&D), infrastructure and so on. Thus, although it may appear paradoxical, Thirlwall's law implies that a country may require certain types of supply-side policies in order to relieve the demand-side BP constraint on its growth.

In addition, the BPCG approach allows – indeed, requires – that some countries (a few large ones) are not BP constrained, but rather are growing at the maximum rate made possible by the expansion of their productive capacity (or their domestic demand). Thirlwall (1979) argued that Japan was such a case, after finding that its long-run average growth rates was significantly lower than its BP-equilibrium growth rate in the period from the 1950s through to the mid-1970s. Similar considerations may apply to China in its years of most rapid growth (roughly 1980–present), as well as to major resource exporters such as Saudi Arabia or a large global demand-driver like the US (which has been able to sustain large current account deficits indefinitely with little apparent penalty).

Third, and here Thirlwall's analysis is contrary to some of the models of cumulative causation covered in Chapter 8, what matters in the BPCG framework is the non-price competitiveness of a country's goods compared with foreign products, as reflected in the income elasticities of exports and imports, rather than their cost competitiveness, which would be reflected in relative prices. By assuming either elasticity pessimism or else constant relative prices in the long run, the BPCG model implies that cost competitiveness is unimportant and non-price competition (in terms of product quality, product as opposed to process innovation, and so on) is what matters most to long-run growth. As McCombie (1989, p. 611, emphasis added) has stated,

> The estimated values of [the income elasticities of exports and imports] show considerable variation between countries and this reflects differences in the various aspects of non-price competition – the quality, reliability etc., of manufactured goods. Thus, *it is the supply characteristics (which determine the degree of non-price competitiveness) that are crucial in determining a country's growth rate relative to that of the rest of the world.*

Fourth, the BPCG model implies that domestic demand-side policies cannot generally be effective in the long run, because any domestic stimulus (for example, through increased government spending) would lead to a BP (trade) deficit, which would then require a reversal of the stimulus policy. This does not mean that government policies are unimportant – measures that could enhance non-price competitiveness (for example, technical training, R&D subsidies and infrastructure investment) may be quite effective – but policies that focus solely on expanding domestic aggregate demand can only be of short-run benefit.

Last but not least, the Thirlwall model has important and controversial implications for international trade policy. Contrary to the Ricardian and neoclassical models of comparative advantage in which trade liberalization is mutually beneficial to all nations, the Thirlwall model depicts a more mercantilist world in which export promotion is generally beneficial but countries have to be cautious about opening up their import markets too much lest they dissipate the gains from increased exports through higher imports. Of course, this is based on a generally Keynesian approach in which the chief considerations are not microeconomic efficiency (the optimal use of given resources), but rather full employment and sustainable long-run growth. According to the BPCG model, it is not irrational for countries to seek lower trade barriers in foreign export markets while being reluctant to open their own markets too much to imports. This does not imply that countries should not engage in trade liberalization, but when they do they need to make sure that the gains on the export side will not be outweighed by losses on the import side. The BPCG model also makes it clear that export promotion should not be confused with trade liberalization – the latter merely means tearing down existing barriers to trade, while the former may entail active government efforts to help export industries (especially those with high income elasticities) get established and succeed.

Nevertheless, the Thirlwall approach by no means supports a blanket protectionist policy or 'trade war' approach. Indeed, it could not, since each country's import restrictions (tariffs, quotas and so on) constitute barriers to other countries' exports. Hence, a world in which all countries impose high tariffs and other barriers on each other would be one in which exports could not thrive and the type of growth envisioned in the BPCG model could not flourish (although if economies become sufficiently closed, they could grow through domestic means, but even then a lack of export opportunities could make them founder on BP constraints if they need certain types of crucial imports such as raw materials, energy or capital goods). Moreover, the political economy fears that protectionism in the real world may sometimes be driven more by special interest politics than by strategic development considerations are not unfounded; nor are microeconomic concerns that protectionism may sometimes shield inefficiency and give disincentives to innovate or raise productivity. Still, the Thirlwall model suggests that trade liberalization has macroeconomic consequences, and that if it is not done correctly it may fail to achieve the promised growth gains or even possibly tighten BP constraints, as many studies have found in various developing countries.[16] Therefore, what the model implies is that trade liberalization should be carried out in a thoughtful and strategic manner, with reciprocal reductions in trade barriers designed to allow all countries sufficient growth of their exports to make up for the increases in their imports – and that any trade strategy needs to be accompanied by industrial and technological policies to enhance non-price competitiveness and promote favourable shifts in income elasticities (higher for exports, lower for imports).

## 9.3 Extensions of the model

The basic BPCG model presented in the previous section is highly simplified in several key respects. Although this starkness is part of its power, it remains to be seen how the major conclusions and policy implications are altered, qualified or amplified if some of the simplifying assumptions are dropped. As one would expect, therefore, the BPCG literature is replete with many different kinds of extensions. For reasons of space, we will confine our presentation in this section and the next to models incorporating the following phenomena: international 'capital' (financial) flows; multisectoral economies with structural change; imports of intermediate goods used in export production; two or more large countries; partial pass-through of exchange rate changes into prices of traded goods; and cumulative causation à la Verdoorn's law from Chapter 8.[17] Of course, these various extensions can be combined into a wide array of more complex models as appropriate for a particular country or situation, but in this section we will cover each extension separately as a distinct modification to the basic framework from section 9.2. In every case considered here, the solution can be seen as a modified version of Thirlwall's law; models of BP constraints that differ more radically from the Thirlwall's law solutions are covered in Chapter 10.

### 9.3.1 International capital flows[18]

One key limitation of the original Thirlwall model was the assumption that trade in goods and services must be balanced in the long run, which ignores the possibility that long-term capital flows could enable some countries to sustain trade surpluses or deficits over long periods of time. As more and more countries have opened up their capital markets since the 1980s, persistent trade imbalances have become common among many countries, including advanced economies and natural resource exporters as well as developing nations. As shown in Table 9.1, a wide range of countries – rich and poor, large and small – have experienced large, sustained imbalances in their net exports (measured as a percentage of gross national product, GDP) on average over the period 2000–16.[19] An obvious and important extension of Thirlwall's model, therefore, is to allow for chronically imbalanced trade matched by net flows of financial capital.[20]

Thirlwall and Hussain (1982) introduced one way of incorporating capital flows by focusing on the growth rate of net inflows. Let us define net capital inflows measured in domestic currency as $NCF$ (where $NCF < 0$ would indicate net outflows). Since Thirlwall and Hussain do not otherwise include net interest payments on international debts, we will define $NCF$ to be net of such payments – in other words, $NCF$ equals net new borrowing minus net outflows of interest on existing international debt.[21] Then the condition for BP equilibrium becomes

$$PX + NCF = EP_fM \tag{9.13}$$

where we can think of the left-hand side of (9.13) as total receipts of foreign exchange. Then, converting to growth rates and defining the share of export revenue in total BP receipts as $\theta = PX/(PX + NCF)$,[22] we obtain the following BP equilibrium condition in growth rate form:

$$\theta(\hat{P} + x) + (1 - \theta)ncf = \hat{E} + \hat{P}_f + m \tag{9.14}$$

**Table 9.1** Indicators of external imbalances for selected countries, averages for 2000–16

| Advanced economies | $NX/Y$ | $\theta = PX/EP_fM$ | Emerging market and developing economies | $NX/Y$ | $\theta = PX/EP_fM$ |
|---|---|---|---|---|---|
| Singapore | 24.0 | 1.14 | Kuwait | 29.4 | 1.98 |
| Norway | 12.4 | 1.44 | Saudi Arabia | 18.9 | 1.62 |
| Switzerland | 9.1 | 1.17 | Bahrain | 18.0 | 1.31 |
| Netherlands | 8.6 | 1.13 | Malaysia | 16.2 | 1.20 |
| Sweden | 5.9 | 1.15 | Angola | 16.1 | 1.35 |
| Denmark | 5.9 | 1.13 | Russia | 9.8 | 1.45 |
| Germany | 5.2 | 1.15 | Nigeria | 9.6 | 1.29 |
| Finland | 3.2 | 1.11 | Thailand | 5.8 | 1.09 |
| Korea, Rep. | 3.0 | 1.09 | Chile | 4.6 | 1.15 |
| Czech Republic | 2.7 | 1.03 | Botswana | 4.5 | 1.14 |
| Hungary | 2.2 | 1.03 | China | 3.8 | 1.17 |
| Canada | 1.2 | 1.03 | Argentina | 3.7 | 1.25 |
| New Zealand | 1.2 | 1.04 | Poland | –1.5 | 0.96 |
| Japan | 0.5 | 1.05 | Mexico | –1.6 | 0.94 |
| Australia | –1.3 | 0.95 | Costa Rica | –2.4 | 0.96 |
| Slovak Republic | –1.3 | 0.98 | Turkey | –3.1 | 0.88 |
| Spain | –1.5 | 0.95 | Colombia | –3.2 | 0.93 |
| United Kingdom | –2.3 | 0.92 | India | –3.2 | 0.85 |
| United States | –3.9 | 0.75 | Cameroon | –3.3 | 0.94 |
| Lithuania | –4.3 | 0.92 | Egypt | –5.3 | 0.78 |
| Portugal | –5.6 | 0.86 | Bulgaria | –7.4 | 0.89 |
| Greece | –7.7 | 0.79 | Tanzania | –8.1 | 0.71 |
| Latvia | –8.3 | 0.85 | Mongolia | –9.7 | 0.82 |
| | | | Kenya | –10.5 | 0.69 |
| | | | Uganda | –11.7 | 0.59 |
| | | | Guatemala | –12.3 | 0.66 |
| | | | Ghana | –15.3 | 0.70 |
| | | | Honduras | –18.6 | 0.66 |
| | | | Jamaica | –18.7 | 0.65 |
| | | | Jordan | –27.1 | 0.64 |

Notes: $NX/Y$ is net exports of goods and services as a percentage of GDP; $\theta = PX/EP_fM$ is the ratio of the value of exports to the value of imports for goods and services. Countries were selected to be representative of those with relatively large surpluses or deficits in various global regions; very small countries and ones with small imbalances or missing data were omitted.

Source: World Bank, *World Development Indicators Online Database*, http://data.worldbank.org/data-catalog/world-develop ment-indicators, data accessed 26 July 2018, and authors' calculations.

where *ncf* is the growth rate of *NCF*. In this equation, the left-hand side is the weighted average of the growth rates of export earnings and net capital inflows, while the right-hand side is the growth rate of import expenditures.

If we substitute equations (9.5) and (9.6) for export and import demand in growth rate form into the equilibrium condition (9.14) and solve for domestic income growth *y*, we obtain the following solution for the BP-equilibrium growth rate:[23]

$$y_B = \frac{[\theta\varepsilon_X + \varepsilon_M - 1)](\hat{E} + \hat{P}_f - \hat{P}) + \theta\eta_X y_f + (1 - \theta)(ncf - \hat{P})}{\eta_M} \tag{9.15}$$

This is the most general solution of the Thirlwall–Hussein model for the BP-equilibrium growth rate with capital flows. The first term in the numerator is the relative price effect, where the satisfaction of the ML condition is more difficult in a country that receives net capital inflows (that is, one with an initial trade deficit) because the price elasticity of export demand $\varepsilon_x$ is multiplied by the share of exports in total receipts $\theta$, and $\theta < 1$ in a country that has a trade deficit (so the price elasticities need to be somewhat higher than in the case of balanced trade for the condition $\theta\varepsilon_X + \varepsilon_M > 1$ to be satisfied, as discussed in Appendix 9.1). The second term is the growth rate of foreign income multiplied by the income elasticity of export demand and weighted by the same share $\theta$, while the third term is the growth rate of real net capital inflows $(ncf - \hat{P})$ weighted by the share of net capital inflows in total receipts $(1 - \theta)$.

If we assume that relative price effects are negligible in the long run, because either $\hat{E} + \hat{P}_f - \hat{P} \approx 0$ (constant relative prices) or $\theta\varepsilon_X + \varepsilon_M \approx 1$ (elasticity pessimism), equation (9.15) simplifies to the following expression for the strong form of Thirlwall's law with net capital flows:

$$y_B = \frac{\theta\eta_X y_f + (1 - \theta)(ncf - \hat{P})}{\eta_M} \tag{9.16}$$

And, as before, under the assumption of constant relative prices, $x = \eta_X y_f$ and we can transform (9.16) into the corresponding weak-form solution:

$$y_B = \frac{\theta x + (1 - \theta)(ncf - \hat{P})}{\eta_M} \tag{9.17}$$

All of these solutions (9.15)–(9.17) highlight the fact that faster growth of net capital inflows *ncf* raises the BP-equilibrium growth rate by allowing more rapid growth of imports without risking an overall BP deficit,[24] for any given rate of export growth (in the weak form – or foreign growth rate weighted by the income elasticity of exports in the strong form). However, when net capital inflows are large (or rather, rapidly growing), Thirlwall and Hussain warn that it may not be possible to rule out relative price effects – even though these may still have relatively little impact in the long run – because sustained net capital inflows may induce a persistent real appreciation of the currency ($\hat{E} + \hat{P}_f - \hat{P} < 0$), and then if the extended ML condition with imbalanced trade ($\theta\varepsilon_X + \varepsilon_M > 1$) is satisfied, the relative price effects do not drop out of equation (9.15). In this situation, the positive impact of the net capital inflows on the BP-equilibrium growth rate is at least partially offset by the tendency of such inflows to cause the currency to appreciate, thereby making the country's goods less competitive in global markets.[25]

However, it may be questioned whether any of the solutions for $y_B$ in equations (9.15)–(9.17) truly constitute long-run equilibria. Thirlwall and Hussain (1982) did define *ncf* as the growth rate of 'permanent, sustainable capital inflows', but did not explicitly address what makes them sustainable. Nevertheless, if the growth rate of net financial inflows *ncf* could be at any given level, it could be high enough to imply that the share of BP receipts accounted for by such inflows $(1 - \theta)$ would be rising steadily over time (which would not be consistent with a long-run steady state), and it could also imply that the country would be accumulating rising international debts relative to GDP (which in turn would imply rising debt service burdens). Countries that have chronically rising current account deficits accompanied by increasing foreign debt–GDP ratios may succeed in growing rapidly (that is, faster than their BP-equilibrium rates) for some period of time, but such borrowing-led booms often end badly in a financial crisis. Even if no crisis results, rapid growth of capital inflows is unlikely to persist for a long period of time, especially if the debt-to-GDP ratio soars or indicators of debt service burdens (for example, the interest outflow as a share of export earnings or GDP) deteriorate. Essentially, there is no guarantee that the ratio $\theta$ stabilizes in the long run in any of equations (9.15)–(9.17), and hence these equations could be regarded as describing medium-run growth paths that may or may not be sustainable in the long run.

In response to this concern, BPCG theorists have developed an alternative way to model capital flows that are sustainable in the long run. McCombie and Thirlwall (1997) proposed a model in which the debt–income ratio must stabilize at a constant level, while Moreno-Brid (1998–99) considered the case of a constant, sustainable ratio of the current account balance (deficit or surplus) to national income (GDP). Both of these models imply the following BP equilibrium condition in growth rate form:[26]

$$\theta(x - y) = \hat{E} + \hat{P}_f - \hat{P} + m - y \tag{9.18}$$

where $\theta = PX/(PX + NCF) = PX/EP_fM$ as before. If we then substitute equations (9.5) and (9.6) for export and import demand in growth rate form into (9.18) and solve for $y$, we obtain a general expression for the BP-constrained growth rate with sustainable capital flows:[27]

$$y_B = \frac{(\theta\varepsilon_X + \varepsilon_M - 1)(\hat{E} + \hat{P}_f - \hat{P}) + \theta\eta_X y_f}{\eta_M - 1 + \theta} \tag{9.19}$$

Then, if relative price effects are negligible in the long run because of either elasticity pessimism $(\theta\varepsilon_X + \varepsilon_M \approx 1)$ or constant relative prices $(\hat{E} + \hat{P}_f - \hat{P} \approx 0)$, equation (9.19) reduces to

$$y_B = \frac{\theta\eta_X y_f}{\eta_M - 1 + \theta} \tag{9.20}$$

and under the latter assumption only this is equivalent to

$$y_B = \frac{\theta x}{\eta_M - 1 + \theta} \tag{9.21}$$

These two solutions for $y_B$ correspond to the strong and weak forms of Thirlwall's law, respectively, modified to incorporate *sustainable* levels of net capital inflows.

There are several things to note about the solutions (9.20) and (9.21). First, the growth rate of actual net capital flows (*ncf*) does not appear in these solutions; these solutions only include the ratio $\theta$, which represents the proportional trade surplus (and hence is inversely related to the extent to which the country requires net capital inflows). In this model, actual net capital flows have to adjust through some (unspecified) endogenous mechanism in order to maintain a sustainable trade imbalance (current account deficit as a percentage of GDP). In fact, it can easily be seen that in order to obtain either solution (9.20) or (9.21), it must be true that $ncf - \hat{P} = y_B$, that is, net capital inflows must grow (in real terms) at the BP-equilibrium growth rate of income in the long run.[28]

Second, the higher is the ratio $\theta$ (indicating a smaller proportional trade deficit that needs to be financed by net capital inflows), the higher is $y_B$ (assuming that $\eta_M > 1$, as suggested by most empirical studies for most countries).[29]

Thus, in the long run, a greater proportion of net capital inflows (in the sense of a lower $\theta$ ratio) does *not* necessarily increase the BP-equilibrium growth rate; on the contrary, as long as $\eta_M > 1$, a higher proportion of net capital inflows (lower $\theta$) actually reduces $y_B$. The reason is that, if $\eta_M > 1$, then real net capital inflows (which have to grow at the same rate as output in the long run) must be growing more slowly than imports and the low growth rate $ncf - \hat{P}$ will therefore hold down the BP-equilibrium growth rate per equation (9.16) or (9.17) until it equals (9.20) or (9.21), respectively.

McCombie and Thirlwall (1997) and Thirlwall (2011) argue that the deviations of (9.20) and (9.21) from the corresponding solutions without capital flows (equations 9.10 and 9.11) are likely to be relatively small. To see their point, consider the average ratio of $\theta = 0.75$ recorded for the US in 2000–16, which is the lowest ratio for any major advanced economy in that period – only some developing countries have lower $\theta$ ratios (see Table 9.1). If we assume that $\eta_M = 2$ and $x = 6$ (which are realistic orders of magnitude for the US economy), then $y_B = 2.57$ per cent with capital inflows ($\theta = 0.75$) whereas it would be 3.00 per cent without capital inflows (if trade were balanced and $\theta = 1$). The difference of 0.43 percentage points in the BP-equilibrium growth rate appears small, as McCombie and Thirlwall claim. However, such an apparently small difference in the sustainable *growth rate* can translate into quite a large difference in *cumulative income changes* if we consider the compounding of economic growth over a long period of time. Using this same example, GDP would increase by a factor of 4.4 times over a 50-year period with an annual growth rate of 3.00 per cent, compared with only 3.6 times with an annual growth rate of 2.57 per cent, resulting in a cumulative loss of 0.8 (80 per cent) of initial GDP in forgone national income a half-century later.

Nevertheless, McCombie and Thirlwall are right to conclude that capital inflows are unlikely to significantly improve a country's growth in the long run, and as long as the income elasticity of import demand is greater than one, such inflows are if anything likely to diminish long-run growth if they have to adjust to sustainable proportions. Thus, we may agree with the summary of the argument by Gouvêa and Lima (2010, p. 173) when they write, 'a major conclusion of the broader literature that considers the possibility of sustainable unbalanced trade is that capital flows are unable to allow an individual country to increase its growth rate above that given by the original Thirlwall's law by very much or for very long'.[30]

To recapitulate, rapid inflows of foreign capital can help a country grow faster than it otherwise could in the medium run per equations (9.15)–(9.17), subject to the qualification that such inflows could lead to RER appreciation that could potentially counteract some of the benefits of the capital inflows via negative price effects on net exports. But if such a debt-led growth boom leads to a rising ratio of the current account deficit to GDP (or, similarly, a rising ratio of external debt to GDP), it may not be sustainable in the long run. In order for the country to stabilize its current account deficit (or its external debt) as a percentage of GDP, it would have to reduce its net capital inflows until they grow (in real terms) at the same rate as GDP itself in the long run, in which case the country's BP-equilibrium growth rate would be held *below* the rate that would result if the country's trade were balanced. Hence, the augmented BPCG model implies that relying on net capital inflows to boost growth is not likely to be a viable long-run strategy, even if it may have a temporary payoff during a period in which such inflows grow more rapidly than is sustainable in the long run.

### 9.3.2 Multisectoral models with structural change and intermediate imports

As discussed earlier, the composition of a country's exports is vital in the BPCG approach because the growth rate of those exports – or, alternatively, their foreign income elasticity – depends on whether the country exports goods that sell in more dynamic or more stagnant global markets. Similarly, the composition of imports can also be important, as the income elasticity of imports will be higher if a country imports goods whose demand is highly income-sensitive (for example, energy products, capital goods or luxury consumption goods), and this in turn would depress the growth rate compatible with BP equilibrium.

The point that export composition matters has been formalized by Araujo and Lima (2007) and Gouvêa and Lima (2010) in a multisectoral BPCG model. What we discuss here is the simplified version of the multisectoral model presented by Gouvêa and Lima (2013), which distils the essence of the Araujo–Gouvêa–Lima approach.[31] In a nutshell, this model transforms the strong-form solution (9.9) into the ratio of the weighted averages of the industry-level income elasticities of demand for exports and imports, multiplied by the foreign growth rate:

$$y_{B,t} = \frac{y_{f,t}\sum_{j=1}^{\overline{N}} \alpha_{j,t}\eta_{X,j}}{\sum_{j=1}^{\overline{N}} \beta_{j,t}\eta_{M,j}} \tag{9.22}$$

where $j$ indexes the industry or good, $t$ indexes time, $\alpha_{j,t}$ and $\beta_{j,t}$ are the shares of good $j$ in total exports and imports (respectively) at time $t$, $\eta_{X,j}$ and $\eta_{M,j}$ are the income elasticities of export and import demand for each good $j$, there are $\overline{N}$ total industries or goods, and – importantly – both the foreign growth rate $y_{f,t}$ and the domestic BP-equilibrium growth rate $y_{B,t}$ are time-varying.

Structural change is reflected in the fact that the composition of exports and imports (as reflected in the industry shares $\alpha_{j,t}$ and $\beta_{j,t}$) is assumed to be time-varying, whereas the income elasticities are presumed to be more or less permanent features of the products to which they pertain (and hence are not modelled as time-varying). Thus, for example, if a country has rising shares of highly income-elastic exports (say, computers and other electronic products), then (holding foreign growth and the composition of imports constant) its BP-equilibrium growth rate will be rising over time. On the other hand, if a country produces exports with low income elasticities (say, textiles and apparel, footwear and food products), while it has a persistently large proportion of highly income-elastic imports (such as capital goods), then that country will have a low BP-equilibrium growth rate. Thus, the multisectoral BPCG model makes the importance of supply-side factors more explicit by showing the need for a country to upgrade its export structure in the direction of goods with higher foreign income elasticities and to bolster its ability to produce domestic substitutes for highly income-elastic imports. Or, in terms of the model parameters, a country's development policies should focus on striving to increase the shares $\alpha_{j,t}$ for exports with high income elasticities $\eta_{X,j}$ and to decrease the shares $\beta_{j,t}$ for imports with high income elasticities $\eta_{M,j}$.

Aside from its disaggregation of industrial sectors, the multisectoral BPCG model also differs from the more traditional, aggregative version in another way. In the traditional version, the BP-equilibrium growth rate is assumed to be stable over long periods of time and hence provides what might be called (in classical terms) a 'strong attractor' for actual growth in the long run (even though it has always been understood that actual growth could deviate from the BP-equilibrium rate in short-run periods, when trade might not be balanced). In contrast, the multisectoral analysis emphasizes the time-varying nature of the BP-equilibrium growth rate in a situation of ongoing structural change, in which the composition of a country's trade is continuously evolving. For this reason, the multisectoral BPCG model is especially applicable to developing and emerging market nations, in which such structural change is an essential feature of their growth (see Szirmai, 2012; Cimoli and Porcile, 2014; Rodrik, 2014, among many others).

One key type of structural change involves the shift that many developing nations have made (following the path blazed by the industrialized countries earlier) away from their traditional specializations in exports of primary commodities and towards specializations in exports of manufactures. One of the motivations for such a shift is the likelihood that most manufactured exports have higher income elasticities than primary commodities; another is the greater prospects for scale economies and productivity growth. On the face of it, such a transformation seems likely to raise the BP-equilibrium growth rate. However, such an outcome is not guaranteed if the new exports are highly intensive in imported intermediate goods, thereby negating at least some of the gains in terms of relieving the BP constraint. As a result of the increasing importance of so-called vertical trade in intermediate inputs and semi-finished products in 'global value chains', many developing countries have ended up doing largely assembly-oriented manufacturing that relies heavily on imported inputs and contains relatively little domestic value added, rather than producing vertically integrated manufactures that use mostly domestically produced inputs.

To analyse this issue, Blecker and Ibarra (2013) and Ibarra and Blecker (2016) adapted the multisectoral BPCG framework to incorporate imports of intermediate goods that are used in export production. To keep the model tractable (and because of limitations in the Mexican data used by Blecker and Ibarra for their empirical estimates), their model is limited to two export sectors and two import sectors. The two exported goods, denoted by subscripts, are manufactures ($m$), which are produced using imported intermediate goods ($i$), and primary products or 'other' goods ($o$), which (for simplicity) are produced using only domestic inputs. (Thus, with apologies for any possible confusion, the variable $m$ means imports but a subscript $_m$ refers to manufactures.) The two imported goods, also denoted by subscripts, are final (consumption and capital) goods ($c$) and intermediate goods ($i$). For three of these goods (manufactured exports and both types of imports), it is assumed (as before) that their supplies are infinitely elastic, they are priced in the seller's currency and their output is strictly demand-determined. In contrast, the real quantity of other exports (primary commodities) grows at the exogenously given rate $x_o$, while their price (denominated in foreign currency) increases at the exogenously given rate $\hat{P}_{o,f}$. This specification assumes that the quantities and prices of primary commodity exports are determined by conditions in global commodity markets as well as domestic supply constraints, and cannot be modelled in the same way as industrial (manufactured) exports.

On these assumptions, the demand function for manufactured exports can be written in growth rate form as

$$x_m = \varepsilon_m(\hat{E} + \hat{P}_f - \hat{P}) + \eta_m y_f \tag{9.23}$$

where $\varepsilon_m$ and $\eta_m$ are (respectively) the relative price (RER) and income elasticities of demand for manufactured exports, and $\hat{P}$ is now understood to be the home inflation rate for manufactured goods while $\hat{P}_f$ is the foreign inflation rate for competing manufactures (assuming that home and foreign manufactured goods are imperfect substitutes). The demand function for imports of intermediate goods incorporates the assumption that these imports are purchased in part for the production of manufactured exports. Written in growth rate form,

$$m_i = -\varepsilon_i(\hat{E} + \hat{P}_f - \hat{P}) + \eta_i y + \mu_i x_m \tag{9.24}$$

where $\varepsilon_i$ and $\eta_i$ are (respectively) the relative price (RER) and income elasticities of demand for intermediate imports and $\mu_i$ is the elasticity of demand for imports of intermediate inputs with respect to manufactured exports. In contrast, the demand function for imports of final (consumption and capital) goods is more similar to the demand function for aggregate imports (equation 9.6) used earlier:[32]

$$m_c = -\varepsilon_c(\hat{E} + \hat{P}_f - \hat{P}) + \eta_c y \tag{9.25}$$

where $\varepsilon_c$ and $\eta_c$ are (respectively) the relative price (RER) and income elasticities of demand for final imports. Note that these equations assume for simplicity that all imports have the same prices and all import-competing domestic goods have the same prices, regardless of whether they are intermediate or final goods.[33]

Assuming no capital flows, the BP equilibrium condition (expressed in growth rate form and with all exports and imports valued in foreign currency) is

$$\alpha_m(\hat{P} - \hat{E} + x_m) + (1 - \alpha_m)(\hat{P}_{o,f} + x_o) = \beta_i(\hat{P}_f + m_i) + (1 - \beta_i)(\hat{P}_f + m_c) \tag{9.26}$$

where $\alpha_m$ is the share of manufactures in the value of total exports and $\beta_i$ is the share of intermediate goods in the value of total imports. Essentially, the left-hand side is the weighted average growth rate of the value of the two types of exports, while the right-hand side is the weighted average growth rate of the value of the two types of imports. Substituting (9.23)–(9.25) into (9.26) and solving for the home country growth rate $y$, we obtain:

$$y_B = \frac{(\alpha_m - \mu_i\beta_i)\eta_m y_f + (1-\alpha_m)(\hat{P}_{o,f} - \hat{P}_f + x_o) + [(\alpha_m - \mu_i\beta_i)\varepsilon_m + \beta_i\varepsilon_i + (1-\beta_i)\varepsilon_c - \alpha_m](\hat{E} + \hat{P}_f - \hat{P})}{\beta_i\eta_i + (1-\beta_i)\eta_c} \qquad (9.27)$$

which is the most general expression for the BP-equilibrium growth rate $y_B$ in the model with intermediate imports. Time subscripts have been suppressed here to avoid cluttering the notation, but it should be understood that the shares $\alpha_m$ and $\beta_i$ and the BP-equilibrium growth rate $y_B$ are all time-varying. Note that $(\alpha_m - \mu_i\beta_i)\varepsilon_m + \beta_i\varepsilon_i + (1-\beta_i)\varepsilon_c - \alpha_m > 0$ must hold in order for a faster rate of RER depreciation (higher $\hat{E} + \hat{P}_f - \hat{P}$) to increase the BP-equilibrium growth rate $y_B$; this inequality is the equivalent of the ML condition for this model.[34]

As before, relative price effects can be considered negligible in the long run if either relative prices are constant in the long run, so that $\hat{E} + \hat{P}_f - \hat{P} = 0$, or elasticity pessimism prevails, which would mean that the ML condition for this model is not satisfied, in which case $(\alpha_m - \mu_i\beta_i)\varepsilon_m + \beta_i\varepsilon_i + (1-\beta_i)\varepsilon_c - \alpha_m \approx 0$. In either case, (9.27) simplifies to the strong-form solution with intermediate imports:

$$y_B = \frac{(\alpha_m - \mu_i\beta_i)\eta_m y_f + (1-\alpha_m)(\hat{P}_{o,f} - \hat{P}_f + x_o)}{\beta_i\eta_i + (1-\beta_i)\eta_c} \qquad (9.28)$$

Alternatively, if (and only if) relative prices are constant, we can deduce from equation (9.23) that $x_m = \eta_m y_f$ when $\hat{E} + \hat{P}_f - \hat{P} = 0$, in which case we obtain the weak-form solution with intermediate imports:

$$y_B = \frac{(\alpha_m - \mu_i\beta_i)x_m + (1-\alpha_m)(\hat{P}_{o,f} - \hat{P}_f + x_o)}{\beta_i\eta_i + (1-\beta_i)\eta_c} \qquad (9.29)$$

In this context, structural change consists in shifts in the relative proportions of manufactures in total exports ($\alpha_m$) and of intermediate goods in total imports ($\beta_i$). In regard to exports of primary products, the model highlights the key role of changes in their terms of trade compared with imports of industrial goods ($\hat{P}_{o,f} - \hat{P}_f$) as well as the growth in their quantity $x_o$, which depend on domestic supply constraints and global commodity market conditions that are taken as exogenously given.

An export strategy focused on primary commodities is risky for well-known reasons, including volatility in their terms of trade, the potential for environmental degradation, eventual resource depletion and the lack of stimulus to technological innovation. Nevertheless, this model also illustrates why a shift towards manufactured exports – even ones with relatively high income elasticities – may not be a panacea for developing countries. If manufactured exports are based on assembly operations that require high proportions of imported intermediate inputs, then faster growth of manufactured exports (even if those exports have high income elasticities) may bring only limited gains (if any) in terms of relieving the BP constraint. Mathematically, if the share of manufactured exports $\alpha_m$ rises but the share of intermediate imports $\beta_i$ also increases, and if the elasticity of intermediate imports with respect to those exports $\mu_i$ is high or rising, there may be little or no gain (and possibly even a deterioration) in the BP-equilibrium growth rate as the weighting factor $(\alpha_m - \mu_i\beta_i)$ in the numerator of equation (9.28) or (9.29) could either rise very little or possibly fall, while the rise in $\beta_i$ has an ambiguous effect in the denominator (which will increase if $\eta_i > \eta_c$ and conversely).[35]

However, the Blecker–Ibarra model is an incomplete model of the role of imports of intermediate goods, because it does not account for the impact of the costs of those imports on the prices of domestically produced goods (including manufactured exports) assembled using imported intermediates as inputs. This limitation is addressed in a different approach to modelling intermediate imports in a BPCG framework by Ribeiro et al. (2016, 2017a, 2017b), which will be covered in Chapter 10 (section 10.6).

### 9.3.3 A model of two large countries

Up to this point, we have considered only the perspective of a single 'home' country – described by the small Keynesian open economy model discussed earlier – which trades with a 'foreign' country that is really the entire rest of the world (or all of a country's major trading partners). Thus, all foreign variables (rates of change in foreign prices $\hat{P}_f$ and income $y_f$) have been taken as exogenously given. However, the BPCG model can easily be extended to consider the situation of two large countries (or blocs of countries) that trade with each other, in which case their reciprocal income effects on each other must be taken into account.

Here, we present a simplified version of the two-country (or two-bloc) BPCG model of McCombie (1993). The two countries (blocs) are designated as *A* and *B*, for which one could think of various real-world examples (northern and southern Europe, North America and East Asia, or the advanced economies of the global 'North' and the developing economies in the 'South'). We simplify by assuming constant relative prices in the long run, so that relative price or RER effects can be ignored. This means that our analysis here is strictly long run in nature. Also for simplicity, we return to the original assumptions that each country produces a single good and there are no sustained capital flows, so trade must be balanced in the long run.[36]

Assuming there are only two countries or groups, each one's exports must equal the other's imports. Assuming again that export supplies are infinitely elastic, the growth rate of each country's exports is determined by the other country's import demand function, which in the absence of relative price effects implies that $x^A = m^B = \eta_M^B y^B$ and $x^B = m^A = \eta_M^A y^A$, where the countries are indicated by superscripts. There is thus a single BP equilibrium condition for the two countries, which (on the assumption of constant relative prices) is simply that each country's exports and imports must grow at the same rate in the long run: $x^A = m^B = x^B = m^A$. Using these import (equal to export) demand functions, BP equilibrium requires that

$$y^A = \frac{\eta_M^B}{\eta_M^A} y^B \tag{9.30}$$

Equation (9.30) is represented by the solid BP line in panel (a) of Figure 9.3. Points above and to the left of this line represent growing trade deficits for *A* and surpluses for *B*, while points below and to the right represent the opposite.

Where the two economies end up on this diagram (which may be on or off the BP line) depends on aggregate demand in both of them (unless one of them is supply constrained, as discussed below). We can specify the growth of each country's income (output) as follows:

$$y^A = k_A^A g_A^A + k_X^A x^A = k_A^A g_A^A + k_X^A \eta_M^B y^B \tag{9.31}$$

$$y^B = k_A^B g_A^B + k_X^B x^B = k_A^B g_A^B + k_X^B \eta_M^A y^A \tag{9.32}$$

where the *A* and *B* superscripts indicate the countries, $g_A$ is the growth rate of domestic autonomous spending (for either country), $k_A$ is the domestic autonomous spending multiplier (again for either country), $k_X$ is the export multiplier and the other variables are as previously defined. Here, we again use the fact that each country's exports equal the other country's imports

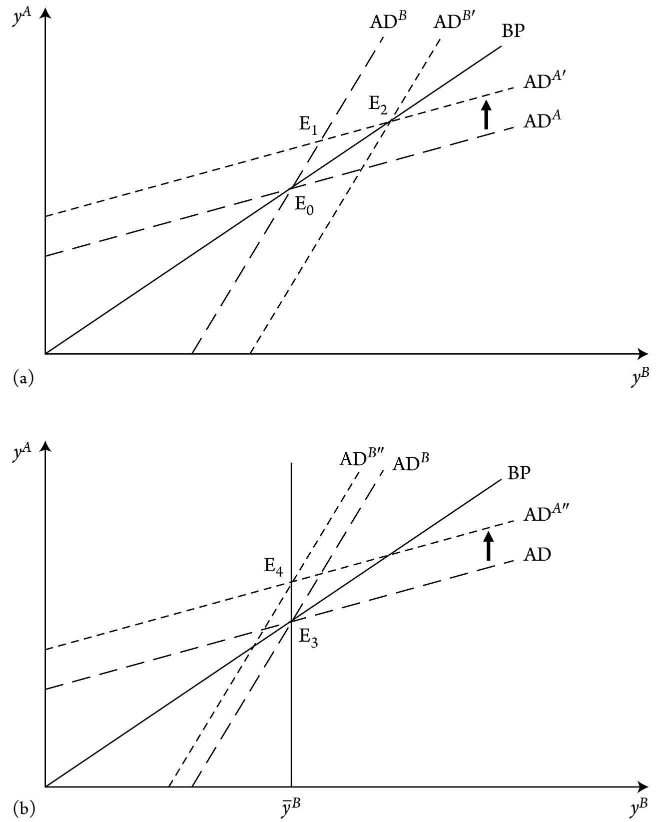



**Figure 9.3** The BPCG model with two large countries and no relative price (RER) effects: (a) the case of no supply or other exogenous constraints; (b) assuming a supply or other exogenous constraint in country *B*

while assuming no relative price effects for simplicity. These two equations are drawn as the dashed AD lines in Figure 9.3; the point where they intersect represents the simultaneous solution of these equations for the actual growth rates of the two countries.[37]

Two important policy messages follow from this analysis. First, consider what happens if both countries are demand constrained and one of them (*A*) increases the growth of its domestic autonomous spending $g_A^A$ (for example, through a government stimulus or a private investment boom). For the sake of discussion, let us assume that the countries start on the lines $AD^A$ and $AD^B$ in panel (a) of Figure 9.3, which intersect at equilibrium point $E_0$ on the BP line, so that trade is initially balanced. Then, when $g_A^A$ rises, the aggregate demand line for country *A* shifts up from $AD^A$ to $AD^{A'}$, and the new equilibrium point is $E_1$, at which *A* has a widening trade deficit and *B* has a growing surplus. Such imbalanced trade would have to be financed by temporary capital flows from *B* to *A*.

Because such a configuration is likely to be unsustainable in the long run, something has to adjust to bring the two countries back to a situation of balanced trade. The more optimistic scenario is that country *B* would follow *A*'s example and raise its own domestic autonomous spending growth rate $g_A^B$ so that its aggregate demand line would shift to the right to $AD^{B\prime}$, in which case the two countries could reach equilibrium point $E_2$ where output growth is higher in both and trade is balanced. This optimistic solution can be called 'global Keynesianism', which can be successful (assuming that no supply constraints are breached in either country) if carefully calibrated increases in autonomous spending in both countries enable them to grow more rapidly simultaneously without either one having a trade deficit. However, if *B* refuses to match *A*'s expansionary stance, sooner or later *A* will be forced to reverse its increase in autonomous spending until its aggregate demand line returns to $AD_0^A$, in which case the growth equilibrium falls back to the original point $E_0$ (this might be called the global austerity case, where both countries grow more slowly). For example, in the aftermath of the global crisis of 2007–09, coordinated fiscal expansions could have led to a more rapid and robust global recovery (thus reaching a point like $E_2$), but unfortunately the adoption of fiscal austerity in many key countries forced the world economy into a suboptimal solution (with slower growth and continued trade imbalances).

Second, we need to consider the case where one country's growth is, as McCombie (1993, p. 489) put it, 'resource- or policy-constrained'. Suppose, for example, that country *B* is growing at the rate made possible by the expansion of its productive capacity on the supply side, or else its policy makers are committed to a fixed growth target for whatever reason. In this case, *B*'s growth rate becomes exogenously fixed, for example at $\bar{y}^B$ as shown in panel (b) of Figure 9.3. In this case, the only global growth equilibrium that is sustainable in the long run is point $E_3$, where *A*'s BP-constrained growth rate is $y^A = (\eta_M^B/\eta_M^A)\bar{y}^B$. Now, if *A* raises the growth of its domestic autonomous spending $g_A^A$ so that its aggregate demand line shifts up to $AD^{A\prime\prime}$, the global growth equilibrium would temporarily shift to point $E_4$. Since *A* would be buying more imports from *B* and *B* does not want to (or cannot) increase its output growth, *B* would be forced to cut back on domestic autonomous spending in order to release the required resources for expanded export production, so *B*'s aggregate demand curve would shift to the left until it reaches $AD^{B\prime\prime}$. Although point $E_4$ thus represents a temporary equilibrium in terms of growth rates, it remains in the region where *A* has a rising trade deficit and *B* has a growing surplus with capital flowing from *B* to *A*, and hence is not sustainable in the long run.

As long as $B$ is constrained to grow at the rate $\bar{y}^B$, then, $A$ will eventually be forced to reverse the increase in its autonomous spending, and the only sustainable long-run equilibrium is at point $E_3$. In this sense, country $A$ is truly BP constrained by $B$'s inability (or unwillingness) to raise its growth rate above $\bar{y}^B$. Only if $B$ either succeeds in relaxing its supply constraints by increasing the growth of its productive capacity (for example, through higher rates of domestic investment and technological progress) or else modifies its policy stance and becomes willing to be more expansionary will it be possible for $A$ to grow faster without incurring a trade deficit that would be unsustainable in the long run.[38] This analysis thus has important implications for various real-world situations, such as the impact on southern European countries when northern European countries (led by Germany) pursue restrictive macro policies that restrain their growth, or the impact of China's development of its productive potential on its various trading partners.

## 9.4 Partial pass-through and cumulative causation

All the extensions considered up to this point modify the strong and weak versions of Thirlwall's law, that is, they change the long-run BP-equilibrium growth rate that prevails when relative price effects are negligible. The last two extensions considered in this chapter are grouped together here because they only alter the relative price effects in the model, and hence they only matter over a time frame in which these effects are significant (short to medium run, or in the long run only if relative price effects do not dissipate over a longer time horizon).

### 9.4.1 Partial pass-through of exchange rate changes

Partial pass-through refers to the qualification made earlier that prices of exported and imported goods do not necessarily remain in constant proportions to the prices of domestic and foreign products, respectively, as firms may adjust the prices of traded goods in order to meet the competition in a given national market. One implication of such 'pricing-to-market' behaviour is that a change in the exchange rate will be only partially (instead of fully) passed through into the relative prices of traded goods, hence the term 'partial pass-through'. In order to allow for the possibility of such behaviour, let us define $P_X$ as the price of exports and $P_M$ as the price of imports, both measured in home currency, while $P_d$ is the price of domestic goods sold at home also in home currency (as before, $P_f$ is the price of foreign goods in foreign currency). Both the traditional ML analysis and the basic version of the BPCG model effectively assume that $P_X = P_d$ and $P_M = EP_f$ or, in growth rate form, $\hat{P}_X = \hat{P}_d$ and $\hat{P}_M = \hat{E} + \hat{P}_f$. However, sellers of traded goods often wish to take prices in the target market (the foreign market for home exports and the domestic market for imports) into account in setting the prices of those goods.

Follow the approach taken by Lavoie (2014), based on the earlier work of Godley (1999) and Godley and Lavoie (2007),[39] one convenient way to represent this is by the following pair of equations in rate-of-change form:

$$\hat{P}_X = \lambda_X(\hat{E} + \hat{P}_f) + (1 - \lambda_X)\hat{P}_d \tag{9.33}$$

$$\hat{P}_M = \lambda_M(\hat{E} + \hat{P}_f) + (1 - \lambda_M)\hat{P}_d \tag{9.34}$$

Thus, full pass-through (as assumed by ML and Thirlwall) assumes $\lambda_X = 0$ and $\lambda_M = 1$, but more generally we can allow for partial pass-through by assuming that $0 \leq \lambda_X \leq 1$ and $0 \leq \lambda_M \leq 1$. Now, the export and import demand equations are rewritten as functions of the appropriate relative prices for exported goods sold abroad (compared with foreign goods) and imported goods sold at home (compared with domestic goods), which in growth rate form are

$$x = \varepsilon_X(\hat{E} + \hat{P}_f - \hat{P}_X) + \eta_X y_f \tag{9.35}$$

$$m = -\varepsilon_M(\hat{P}_M - \hat{P}_d) + \eta_M y \tag{9.36}$$

Before discussing the model solution, it is useful to show how the terms of trade $(P_X/P_M)$ are related to the RER $(EP_f/P_d)$ in this specification. Written in rate-of-change form, equations (9.33) and (9.34) imply that

$$\hat{P}_X - \hat{P}_M = (\lambda_X - \lambda_M)(\hat{E} + \hat{P}_f - \hat{P}_d) \tag{9.37}$$

The absolute value of the elasticity of the terms of trade with respect to the RER thus equals $|\lambda_X - \lambda_M|$. With full pass-through (in which case $\lambda_X = 0$ and $\lambda_M = 1$), this elasticity equals unity, and as pass-through becomes more partial (that is, as $\lambda_X$ rises and $\lambda_M$ falls), the elasticity decreases (in absolute value). Thus, with partial pass-through, a nominal depreciation ($\hat{E} > 0$) will bring about a less than proportional fall in a country's terms of trade.

The main change in the model with partial pass-through is a modification of the ML elasticities condition. The rate of change in the trade balance ratio can be written as the difference between the rates of change in the value of exports and the value of imports as follows:

$$\hat{\theta} = \hat{P}_X + x - (\hat{P}_M + m) \tag{9.38}$$

where the exchange rate is omitted because both export and import prices are measured in the same (home) currency units. After substituting (9.33) and (9.34) into (9.35) and (9.36) and then substituting the resulting expressions along with (9.37) into (9.38), we can derive the effect of a change in the rate of nominal depreciation on the rate of change in the trade balance:

$$\frac{\partial \hat{\theta}}{\partial \hat{E}} = (1 - \lambda_X)\varepsilon_X + \lambda_M \varepsilon_M + \lambda_X - \lambda_M \tag{9.39}$$

Thus, a more rapid depreciation makes the trade balance increase faster if $(1 - \lambda_X)\varepsilon_X + \lambda_M \varepsilon_M > \lambda_M - \lambda_X$ or, in other words, if the weighted sum of the price elasticities of export and import demand (where the weights are the pass-through coefficients) exceeds the absolute value of the elasticity of the terms of trade with respect to the RER. This may appear easier to satisfy than the standard ML condition (for which $\lambda_M - \lambda_X = 1$), since $\lambda_M - \lambda_X$ is likely to be closer to zero than to unity here, but partial pass-through also lowers the weights on the price elasticities so no general conclusions can be stated. It is easy to see, however, that it is possible to have cases in which this condition may be satisfied even though the traditional ML condition is not.[40]

Under this same condition, a depreciation will raise the BP-equilibrium growth rate if relative price effects are significant. To see this, note that the BP equilibrium condition in this model is obtained by setting $\hat{\theta} = 0$ in equation (9.38), which after making appropriate substitutions yields the solution:

$$y_B = \frac{[(1 - \lambda_X)\varepsilon_X + \lambda_M \varepsilon_M + \lambda_X - \lambda_M](\hat{E} + \hat{P}_f - \hat{P}_d) + \eta_X y_f}{\eta_M} \tag{9.40}$$

If the generalized price elasticity (ML) condition $(1 - \lambda_X)\varepsilon_X + \lambda_M \varepsilon_M > \lambda_M - \lambda_X$ is satisfied, a real depreciation of the currency $(\hat{E} + \hat{P}_f - \hat{P}_d > 0)$ will cause $y_B$ to increase, albeit at the cost of a falling terms of trade (since $\hat{P}_X - \hat{P}_M < 0$ in this situation). However, if this condition does not hold or the RER is constant $(\hat{E} + \hat{P}_f - \hat{P}_d = 0)$ in the long run, then (9.40) reduces to the standard strong-form solution (9.9), and with a constant RER it also yields the standard weak-form solution (9.10).

## 9.4.2 Cumulative causation and Verdoorn's law

Another modification to the BPCG model that only matters as long as relative price effects are significant is to allow for cumulative causation through endogenous productivity growth, as in the ELCC model covered in Chapter 8. To avoid mixing up issues, we revert here to the basic model of prices in which there is full pass-through so that there is only one price for domestic output and exports, with inflation rates $\hat{P} = \hat{P}_d = \hat{P}_X$ for home goods and $\hat{P}_M = \hat{E} + \hat{P}_f$ for imports. We will adopt two equations from the ELCC model, starting with the one for markup pricing in rate-of-change form:

$$\hat{P} = \tau' + \hat{W} - q \qquad (9.41)$$

where $\tau'$ is the proportional rate of change in the price–cost margin (one plus the markup), $\hat{W}$ is the growth rate of nominal wages and $q$ is the rate of labour productivity growth.[41] For present purposes, we abstract from changes in markups for simplicity and assume $\tau' = 0$, in which case (9.41) simplifies to $\hat{P} = \hat{W} - q$, which is the same as equation (8.20) in Chapter 8. Productivity growth is assumed to be endogenous according to Verdoorn's law, which was given in equation (8.21) and is reproduced here for convenience:[42]

$$q = q_0 + \rho y \qquad (9.42)$$

where $q_0$ is a shift factor representing autonomous technological dynamism (including catch-up possibilities) and technology policies (R&D subsidies, intellectual property rights and so on), $\rho$ represents the Verdoorn effect (dynamic increasing returns or positive feedbacks) and $q_0, \rho > 0$.

Substituting equations (9.41) and (9.42) into equations (9.4)–(9.6) and solving for $y$, we obtain the following solution for the BP-equilibrium growth rate:[43]

$$y_B = \frac{(\varepsilon_X + \varepsilon_M - 1)(\hat{E} + \hat{P}_f - \hat{W} + q_0) + \eta_X y_f}{\eta_M - \rho(\varepsilon_X + \varepsilon_M - 1)} \qquad (9.43)$$

where $\hat{P}_f$, $\hat{W}$ and $y_f$ are all taken as exogenously given. The denominator must be positive in order to have an economically sensible solution in which greater price competitiveness or higher foreign growth leads to a rise in $y_B$. Therefore, assuming that the ML condition holds $(\varepsilon_X + \varepsilon_M > 1)$, we must also assume $\rho < \eta_M/(\varepsilon_X + \varepsilon_M - 1)$. In other words, the Verdoorn effect cannot be too large, or there cannot be too much cumulative causation. As

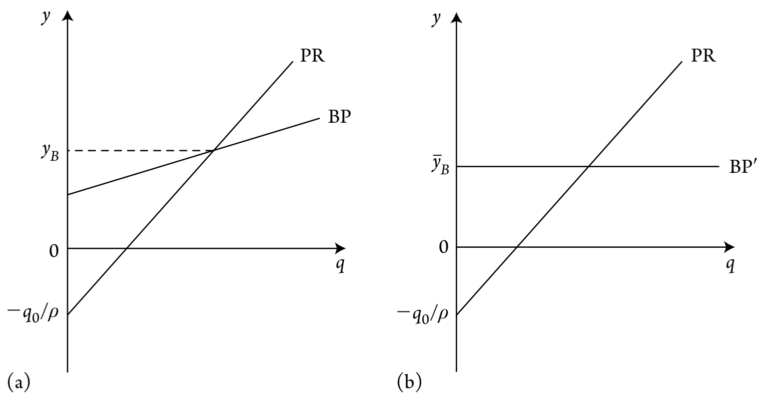



**Figure 9.4** BP-equilibrium output growth with endogenous productivity growth (Verdoorn's law): (a) assumes relative prices (RERs) affect the BP constraint; (b) assumes relative prices (RERs) do not affect the BP constraint

long as ML is satisfied and relative prices are not constant, equation (9.43) will apply, in which case the BP-equilibrium growth rate is an increasing function of both the autonomous component of productivity growth $q_0$ and the Verdoorn coefficient $\rho$. Moreover, since productivity growth is endogenous (and is an increasing function of output growth), anything that relieves the BP constraint (for example, faster growth of foreign income $y_f$ or a rise in the income elasticity of exports $\eta_X$) will also lead to more rapid productivity growth. Hence, the solution (9.43) embodies cumulative causation between export competitiveness, output growth and productivity growth, albeit via a somewhat different feedback mechanism compared with the ELCC model in Chapter 8.[44]

To show the similarity with the ELCC model, we can illustrate the BPCG model with Verdoorn's law as shown in Figure 9.4, which is drawn in a way that is deliberately analogous to Figure 8.2. In panel (a) of Figure 9.4, the upward-sloping BP line represents BP equilibrium as given by the solution of equations (9.4)–(9.6) and (9.41) for output growth $y$ as a function of productivity growth $q$:

$$y = \frac{(\varepsilon_X + \varepsilon_M - 1)(\hat{E} + \hat{P}_f - \hat{W} + q) + \eta_X y_f}{\eta_M} \tag{9.44}$$

As in Figure 8.2, the PR line in Figure 9.4 is the 'productivity regime' representing the Verdoorn equation (8.21) or (9.42). The intersection of BP and PR represents the simultaneous solution of equations (9.42) and (9.44), which is equivalent to equation (9.43). Either an upward shift in the BP line or a rightward shift in PR would increase the equilibrium values of both $q$ and $y_B$. In effect, the BP relation replaces the demand regime (DR) curve of Figure 8.2 (or, one could assume that demand has to adjust to make the DR curve pass through the equilibrium of PR and BP, for example through fiscal policy targeted on keeping trade balanced). However, if either elasticity pessimism holds ($\varepsilon_X + \varepsilon_M \approx 1$) or relative prices are constant ($\hat{E} + \hat{P}_f - \hat{P} = 0$), then the BP equilibrium condition becomes the horizontal line BP′ shown in panel (b) of Figure 9.4 (this corresponds to either equation 9.9 or 9.10, neither of which depends on productivity growth $q$). Thus, when relative price effects are eliminated, cumulative causation (endogeneity of $q$) plays no role in determining the long-run BP-equilibrium growth rate, which we can designate as $\bar{y}_B$ to show that it is independent of $q$. In this case, $q$ remains a function of $y$, but the former has no impact on $\bar{y}_B$, as originally demonstrated by Thirlwall and Dixon (1979).[45]

To conclude, over any time horizon short enough for relative prices to change persistently (and assuming that price elasticities are high enough to satisfy the pertinent ML condition), cumulative causation effects can be felt within the BPCG model, but in a time frame long enough for relative prices to be stationary (or if the price elasticities are not high enough to satisfy ML) those effects drop out of the model. In either case, relaxing the BP constraint (shifting BP upward) always improves productivity growth (the long-run equilibrium occurs at a higher point on PR, indicating a higher equilibrium $q$) as long as Verdoorn's law operates, but a positive feedback of higher productivity growth onto BP-equilibrium output growth (indicating cumulative causation) only occurs when relative price effects are significant (that is, when BP is upward sloping). In section 10.6 in Chapter 10, we will consider a model that more explicitly shows how an economy can transition from an equilibrium similar to equation (9.43), in which Verdoorn effects matter, to a long-run equilibrium characterized by Thirlwall's law.

## 9.5 Conclusions

Since the pioneering work of Thirlwall (1979), the BPCG approach has developed into a rich framework for analysing growth in open economies in which the need to maintain BP equilibrium is the dominant constraint in the long run. Starting with a starkly simple model, which yields the basic solutions for Thirlwall's law (in its weak and strong forms), this framework has been extended to incorporate numerous complexities including capital flows, structural change and large countries (the last with repercussion effects). Perhaps most importantly, the BPCG approach has strong policy implications. Thirlwall's law implies that trade liberalization and export promotion efforts are likely to yield disappointing dynamic gains (and possibly even dynamic losses) if adequate attention is not paid to ensuring that countries are able to increase their export growth by enough to offset their increased openness to imports. The model also implies that what matters most for long-run growth is not the degree of trade openness per se, but rather the structural transformations that enable a country to raise the ratio of the (weighted average) foreign income elasticity of its exports to the (weighted average) domestic income elasticity of its imports.

However, the standard BPCG models have also been subject to debate for various reasons, ranging from doubts about some of the empirical methods used for testing their predictions to questions about the applicability of the model under different structural conditions. According to some critics, for example, the standard BPCG model of a Keynesian small economy does not apply to truly small countries, which should be treated as pure price-takers, while the model only works for large countries if the terms of trade (relative prices) are allowed to adjust. Another line of debate has centred on how the model needs to be modified if the BP-equilibrium growth rate has to be reconciled with the natural rate of growth in the long run. Still other economists have sought to identify ways in which Kaldor–Verdoorn cumulative causation can affect long-run growth in the presence of a BP constraint, without having to allow for implausible continuous variations in relative prices.

In recent years, the deliberate de-emphasis on relative price or RER effects in the traditional BPCG model has been challenged even by some economists sympathetic to this framework. In all the formulations covered in this chapter, relative prices or RERs enter the models in rate-of-change form, and on the assumption that these variables don't change continuously in the same direction over very long periods of time, relative price changes are generally ignored for the purposes of long-run analysis. That is exactly why, for example, cumulative causation cannot affect the long-run BP-equilibrium growth rate in a standard BPCG model, since Kaldor–Verdoorn effects (positive feedbacks from output growth to productivity growth) operate by lowering one country's prices of traded goods *continuously* relative to other countries. The assumption that relative prices either don't change or have negligible effects in the long run is the basis for the policy implication that what matters to long-run growth is *only* non-price (qualitative) competitiveness (as reflected in the income elasticities of exports and imports), rather than cost competitiveness (as reflected in relative prices or the RER).

However, Chapter 8 cited empirical evidence showing that levels of RERs (or relative unit labour costs) do matter to long-run growth, even if their rates of change do not. This type of evidence has led to a variety of new modelling efforts aimed at incorporating levels of relative prices or RERs into a broader BPCG framework. These efforts are analogous to how Boggio and Barbieri (2017) revived the earlier approach of Beckerman (1962) by incorporating levels of relative labour costs into an export-led growth framework, as covered in Chapter 8. The next chapter will explore these contemporary debates over the validity of the BPCG model as well as several new extensions, alternatives and reconciliations that have been proposed to rectify the model's perceived deficiencies in these various dimensions.

## ? STUDY QUESTIONS

1) What are the key assumptions, conclusions and policy implications of the BPCG model? How do these compare with the same for the ELCC model from Chapter 8?
2) In what sense is long-run growth export-led in the BPCG model? How does this compare with the mechanism of export-led growth in ELCC?
3) Does trade liberalization necessarily improve a country's BP-equilibrium growth rate? Apply both the basic BPCG model and any relevant extensions in your analysis.
4) How can the post-Keynesian BPCG model and a neoclassical model of real exchange rate (RER) adjustment be considered to be two alternative 'solutions' to a common underlying framework of BP adjustment?
5) Does capital mobility significantly alter the results and implications of the BPCG model? Discuss.
6) Using a model of two large countries, compare the prospects for a domestic demand-driven increase in long-run growth if the demand stimulus is applied only in one country versus coordinated between the two countries. Also consider the case of a demand expansion in one large country if the other faces a resource or policy constraint.
7) Explain how the BPCG model can be modified for the cases of (a) partial exchange rate pass-through and (b) incorporating Verdoorn productivity effects. What assumption(s) of a standard BPCG model has (have) to be suspended for these factors to matter to the BP-equilibrium growth rate?

## NOTES

1 Although Thirlwall developed this model quite independently, he subsequently discovered that his model had several intellectual antecedents or anticipations (as recounted in Thirlwall, 2011). The idea of relative growth rates being driven by differences in income elasticities of demand for exports and imports in the context of trade between the developed 'centre' and less developed 'periphery' was proposed by Prebisch (1950) and later formalized by Rodríguez (1977, p. 227 n. 74). In addition, Thirlwall's law can be seen as a dynamic counterpart to Harrod's (1933) foreign trade multiplier, which was covered in Chapter 8, while the related idea of a 'foreign exchange gap' was found in the two-gap models of Chenery and Bruno (1962) and Chenery and Strout (1966). Also, Houthakker and Magee (1969) stated the fundamental idea that unequal income elasticities of exports and imports require adjustments of either exchange rates or income growth in order to avoid growing trade imbalances over time.

2 However, both views are at odds with the neo-Keynesian and neo-Kaleckian theories that put more emphasis on domestic investment as constraining profits and growth, as discussed in Chapters 3 and 4. Razmi (2016a) produces a partial reconciliation of these views by constructing a model of a

'small open economy' with a BP constraint, in which investment plays a key role (see section 10.5.1 in Chapter 10).

**3** As discussed in Chapter 8, Kaldor emphasized the importance of relative costs and prices in his earliest work on export-led growth, and then changed his mind in light of 'Kaldor's paradox' – which led him to become more sympathetic to a BPCG view. Some extensions of the BPCG approach allow for relative price or RER effects, as discussed later in this chapter and the next, but we are referring here to the original version of Thirlwall (1979).

**4** Note that these assumptions are similar to the characterization of exports and imports in the open economy neo-Kaleckian model covered in Chapter 4 (section 4.4.3).

**5** Alternatively, $X_0$ could be modelled as incorporating a time trend representing (at least in an ad hoc way) supply-side improvements in a country's export capacity over time: $X_0(t) = X_0(0)e^{\lambda t}$, where $t$ is time, $e$ is the base of the natural logarithm and $\lambda$ is the rate of increase in export capacity (for example, Blecker, 1992; Ibarra and Blecker, 2016).

**6** In principle, Thirlwall defines BP equilibrium as balance on current account, but in a simplified model in which there are no unilateral transfers or net investment income flows, this is equivalent to balanced trade in goods and services.

**7** Nevertheless, the country is not a theoretical 'small country' defined as a price-taker in its export market; see Chapter 10 for further discussion.

**8** In the more extreme case in which $\varepsilon_X + \varepsilon_M < 1$, the trade balance would actually deteriorate as a result of a currency depreciation (increase in the RER as defined here). In this case, the responses of the quantities of exports and imports would be so small that they would be outweighed by the higher cost of imported goods.

**9** The origin of the phrase 'Thirlwall's law' is somewhat murky. Thirlwall (2011, p. 310 n. 1) generously attributes this expression to Skolka (1980, p. 14), who used the German phrase 'das Thirlwallschen Gesetz'. However, this expression translates more literally as 'the Thirlwall law', since the name is used as an adjective. Thirlwall (1979, pp. 46, 50) had already characterized his results as showing that 'a new economic law might almost be formulated', and stated that the implications of equation (9.10) 'might almost be stated as a fundamental law'. Davidson (1990–91, 1992) helped to popularize the English phrase 'Thirlwall's law'. We are indebted to Torsten Niechoj for supplying a copy of Skolka's article in the original German and for discussion of the English translation.

**10** It should be noted that most of these assumptions are the same as the ones on which the conventional ML analysis of the conditions for a currency devaluation to improve the trade balance are based (see Appendix 9.1).

**11** The notion that manufactures and primary products have different supply conditions and market structures goes back to Ricardo (1821 [1951]) and Kalecki (1954 [1968]), among others.

**12** For one version of a BP-constrained growth model with partial exchange rate pass-through and flexible markups, see Blecker (1998). An alternative version of the model with partial pass-through by Godley and Lavoie (2007), who distinguish prices of traded goods from domestic prices, is discussed in section 9.4.1 below.

**13** This solution is discussed explicitly in Krugman (1989), who attributes it to the earlier work of Johnson (1958). The same idea is implicit in some of the discussion in Houthakker and Magee (1969).

**14** Houthakker and Magee (1969) pioneered in econometrically estimating differences in income elasticities of export and import demand across countries, and argued that these could give rise to chronic trade imbalances if countries grow at similar rates. Such results, which have been confirmed many times since then, are the basis for the frequently heard arguments that the US (which is usually found to have a relatively high $\eta_M$) needs to depreciate the dollar in real terms in order to avoid growing trade deficits, while surplus countries like Germany, Japan and China need to appreciate their currencies in order to avoid growing trade surpluses because they are typically estimated to have a relatively high $\eta_X$.

**15** Although they did not develop a model of the type discussed here, the idea of such a trade-off was implicitly recognized in the argument of Houthakker and Magee (1969) that countries with unfavourable income elasticities (that is, relatively higher for imports than for exports) would feel pressure to either devalue their currencies or restrict their output growth in order to avoid rising trade deficits. In particular, they found that the income elasticity of US exports was relatively low, which implied that 'there will have to be differences in relative growth rates or in inflation, or exchange rates will have to be adjusted' in order to prevent a 'secular . . . deterioration' in the US trade balance (Houthakker and Magee, 1969, pp. 120–21).

**16** According to these studies, the income elasticity of import demand $\eta_M$ typically increased proportionately more than the rate of export growth $x$ after trade liberalization. See, for example, Moreno-Brid (1998, 1999), Pérez Caldentey and Moreno-Brid (1999), Pacheco-López and Thirlwall (2004), Santos-Paulino and Thirlwall (2004), and Pacheco-López (2005).

**17** McCombie and Thirlwall (1994) and Thirlwall (2011) discuss many of these extensions along with additional ones, which are not covered here for reasons of space. See McCombie and Thirlwall (2004) for a collection of many of the original articles in this genre.

**18** What we call here (in traditional terminology) international 'capital flows' are really international transactions in financial assets and liabilities and do not generally involve movements of physical or productive capital. Indeed, in contemporary BP accounting, what used to be called the 'capital account' is often called the 'financial account', in which net increases in domestic liabilities to foreigners (so-called capital inflows) have a positive sign and net increases in domestic ownership of foreign assets (so-called capital outflows) have a negative sign (and net decreases have the opposite signs). Although the financial account includes foreign direct investment, which is the acquisition of significant ownership shares of foreign enterprises, a large part of it consists of trade in purely financial assets and liabilities such currencies, bonds, bank loans, equity shares and so on.

**19** Of course, the current account includes net flows of investment income and unilateral transfers, as well as trade in goods and services. However, since the formal modelling in the BPCG approach usually focuses on exports and imports of goods and services, we present data on net exports in this table.

**20** In principle, net capital or financial flows (including official reserve transactions) offset current account imbalances, rather than trade imbalances per se, but if we ignore transfer payments and remittances of migrant labour and include net interest payments in the financial flows, we can say that net financial inflows (or outflows) must match net imports (or net exports) of goods and services. See Alleyne and Francis (2008) for a BPCG model that incorporates transfers or remittances.

**21** Thus, *NCF* here does not equal the financial account balance in the BP, but rather equals the financial account balance minus net outflows of interest. See Moreno-Brid (1998–99) for a model that makes interest payments explicit in a BPCG framework. In countries where transfer payments (foreign aid or remittances) are important, these could also be included, as in the 'net financial inflows' variable used by Blecker (2009).

**22** Using equation (9.13), we can see that $\theta$ as defined here is equivalent to the trade balance or net export ratio $\theta = PX/EP_fM$, as defined in Appendix 9.1.

**23** Of course, in the special case in which trade is balanced and there are no net capital flows (so that $\theta = 1$ and $ncf = 0$), equation (9.15) reduces to (9.8). In the same special case, equation (9.16) similarly reduces to (9.9) and (9.17) reduces to (9.10).

**24** The 'overall BP' here refers to the BP excluding official reserve transactions by central banks, that is, the sum of the current account balance plus the non-official capital (financial) account balance. In systems with fixed or managed exchange rates, official reserve transactions are carried out to maintain pegs or targets for the exchange rate, and can be regarded as accommodating all other ('overall' in this sense) BP flows.

**25** This is similar to the 'Dutch disease' phenomenon, in which booming exports of primary commodities can bring such an inflow of foreign exchange that the RER appreciates sharply, thereby making other exports (especially of manufactured goods) less competitive. In today's world, commodity booms are often accompanied by speculative capital inflows so that the two mechanisms work together to produce an overvaluation of the currency that depresses industrial exports.

**26** See Moreno-Brid (1998–99) for the derivation. The more traditional assumption of balanced trade (a current account balance equal to zero) in the long run is the special case in which $\theta = 1$ and equation (9.18) becomes (9.4).

**27** For economically meaningful results (a positive growth rate), the denominator of this expression must be positive. This is likely to be true, because the only way the denominator could be negative would be for a country to have a very low income elasticity of import demand ($\eta_M << 1$) and a very large trade deficit ($\theta << 1$), which seems like an unlikely combination. If $\eta_M > 1$, as seems to be empirically true in most countries, then $\eta_M - 1 + \theta > 0$ regardless of the size of $\theta$ (since $\theta > 0$).

**28** To see this mathematically, using the weak-form solution as an example (the mathematics are parallel for the strong form), set the solutions for $y_B$ in equations (9.17) and (9.21) equal to each other and solve for $(ncf - \hat{P})$, which (after simplification) yields:

$$ncf - \hat{P} = \frac{\theta x}{\eta_M - 1 + \theta} = y_B$$

**29** From (9.21), it can be seen that

$$\frac{\partial y_B}{\partial \theta} = \frac{(\eta_M - 1)x}{[\eta_M - 1 + \theta]^2}$$

the sign of which (assuming $x > 0$) depends only on the sign of $(\eta_M - 1)$. As long as $\eta_M > 1$, $\theta$ and $y_B$ move in the same direction.

**30** An important qualification – which is consistent with Thirlwall's approach – is that if the net capital inflows are invested in the development of vibrant export industries, thereby increasing the income elasticity of export demand and/or the growth rate of exports, then such inflows could have a lasting benefit. However, such a possibility depends more on the composition or use of the capital flows, not their overall growth rate, and there are alternative ways for countries to promote export industries without relying too heavily on capital inflows (see, for example, Amsden, 1989; Wade, 1990; Lee, 2013). The 'new developmentalism' approach of Bresser-Pereira et al. (2015) also emphasizes that countries should not rely on significant net capital inflows as a means to promote industrial development, but instead should emphasize competitive RERs and reliance on domestic savings.

**31** The original versions of Araujo and Lima (2007) and Gouvêa and Lima (2010) were more complicated because they embedded their BPCG analysis in the framework of a Pasinetti (1981, 1993) model of structural change, a full presentation of which would take us too far afield here. For other perspectives on structural change see section 8.2.2 in Chapter 8.

**32** Although some imported capital goods could be used to increase export capacity, they are not likely to be correlated with current export production, and hence we simplify by assuming that they are not related to manufactured exports. In econometric estimates for Mexico, Ibarra (2011) and Blecker and Ibarra (2013) found that imports of final goods (either consumption and capital goods separately, or the two combined) are not significantly affected by current or lagged manufactured exports, but the share of intermediate imports $\beta_i$ increased and this offset Mexico's gains from increased export growth after the country liberalized its trade.

**33** These assumptions about import prices are made for simplicity, but also accord with the data limitations faced by Blecker and Ibarra (2013) and Ibarra and Blecker (2016) in their econometric work on Mexico.

**34** To see the analogy to the standard ML condition, note that the condition for this expression to be positive can be rewritten as stating that a weighted sum of the price of RER elasticities (all defined so as to be positive) must be greater than unity: $[1 - (\mu_i\beta_i/\alpha_m)]\varepsilon_m + (\beta_i/\alpha_m)\varepsilon_i + [(1 - \beta_i)/\alpha_m]\varepsilon_c > 1$.

**35** For the Mexican case, Blecker and Ibarra (2013) and Ibarra and Blecker (2016) found that $\eta_c > \eta_i$, so the rise in $\beta_i$ decreased the denominator. Nevertheless, these authors also found that the overall impact of a rising $\beta_i$ on $y_B$ was negative in Mexico, since the impact on the numerator outweighed the impact on the denominator.

**36** The two-country model can easily be generalized to allow for changes in relative prices (RERs) and imbalanced trade with net capital flows, as shown by McCombie (1993). These cases are omitted here for reasons of space.

**37** Since the $y_A$ line is drawn as flatter than the BP equilibrium line, the diagram assumes that the slope of the former is less than that of the latter, or $k_X^A\eta_M^B < \eta_M^B/\eta_M^A$, which implies $0 < \eta_M^A k_X^A < 1$. Similarly, the fact that $y_B$ is drawn as steeper than BP implies that $0 < \eta_M^B k_X^B < 1$, and together these inequalities imply that $0 < 1 - \eta_M^A\eta_M^B k_X^A k_X^B < 1$ and $1/(1 - \eta_M^A\eta_M^B k_X^A k_X^B) > 0$. See McCombie (1993, p. 487 n. 9).

**38** In the case where $B$ has a policy constraint, we can think of this constraint as giving the country an exogenously fixed growth rate of domestic autonomous spending, $\bar{g}_A^B$. Using equations (9.30) and (9.32), we can see that in this situation $A$'s BP-equilibrium growth rate becomes proportional to $\bar{g}_A^B$:

$$y^A = \frac{\eta_M^B k_A^B}{\eta_M^A (1 - \eta_M^B k_X^B)} \bar{g}_A^B$$

where $\eta_M^B k_X^B < 1$.

39 Blecker (1998) dealt with partial pass-through of nominal exchange rate changes via adjustments of markups by foreign and domestic producers, but he did not incorporate the distinction between export and import prices on the one hand and domestic and foreign prices on the other. Implicitly, of course, the markups must be flexible in the Godley–Lavoie approach as well, but they are not modelled explicitly.

40 Using the numerical example from Lavoie (2014, p. 523), assume $\lambda_X = 0.2$ and $\lambda_M = 0.5$, which implies $\lambda_X - \lambda_M = -0.3$ with an absolute value of 0.3. Suppose that the price elasticities are both 0.4 so that the standard ML condition would not be satisfied (since $0.4 + 0.4 = 0.8 < 1$). Using Lavoie's partial pass-through coefficients, the relevant comparison instead is $[(1 - 0.2) \times 0.4] + (0.5 \times 0.4) = 0.52 > 0.30$, so the modified ML condition is satisfied.

41 Recall that labour productivity is defined as $Q = 1/a_0$, where $a_0$ is the labour–output ratio, and its growth rate is $q = -\hat{a}_0$.

42 Recall that this is a simplified, aggregative version of Verdoorn's law, whereas the original version was confined to the manufacturing sector as discussed in Chapter 8. In the presentation here, we focus on a single country and hence we omit the parallel specification of the foreign country (rest of world) covered in section 8.3.2.

43 A similar solution is given as equation (16.9) in Blecker (2013b), except in a model that also allows for imbalanced trade financed by net capital inflows ($\theta < 1$). Here, balanced trade ($\theta = 1$) is assumed.

44 In the ELCC model, faster export growth increases domestic output growth through standard Keynesian multiplier effects (the foreign trade multiplier). In the present model, faster export growth increases the BP-equilibrium growth rate by allowing for more rapid growth of imports without incurring a trade deficit when the home country's output grows faster.

45 Also note that the growth of domestic autonomous demand ($g_A$) cannot matter here in the long run, as it does in the extended ELCC model covered in section 8.6 of Chapter 8, due to the imposition of the BP constraint.

## Appendix 9.1 The Marshall–Lerner condition

This condition, named after Alfred Marshall (1842–1924) and Abba Lerner (1903–82), shows how high the price elasticities of export and import demand need to be in order for a devaluation of a country's currency to improve its trade balance (increase net exports), holding other factors constant (a caveat that is important, as we shall see). Although the word 'devaluation' is usually applied to increasing a fixed exchange rate (that is, raising the amount of domestic currency required to buy a unit of foreign currency), while 'depreciation' is used to refer to a market-driven rise in a flexible exchange rate (defined in the same units), we will use the two terms interchangeably here.

To see the problem that this analysis addresses, suppose we measure net exports or the trade balance (we use these terms interchangeably) as the difference between the value of exports and the value of imports in home currency units:

$$NX = PX - EP_f M$$

where the home currency price of exports ($P$) and foreign currency price of imported goods ($P_f$) are taken as given (recall from the text that prices are assumed to be fixed in the seller's currency). Suppose the home country currency depreciates, which means that $E$ (the domestic currency price of foreign exchange) rises. We would expect that, to some extent, the quantity of exports $X$ would increase (as domestic goods become relatively cheaper) and the quantity of imports $M$ would decrease (as foreign goods become relatively more expensive) – how much so will depend on the price elasticities $\varepsilon_X$ and $\varepsilon_M$ in equations (9.1) and (9.2). While these quantity changes help to raise $NX$, they are offset to some degree by the rise in $EP_f$, that is, the price of imported goods converted to domestic currency, which makes any given quantity of imports more expensive for domestic purchasers. Thus, the problem is whether the changes in the quantities are sufficient to outweigh the 'valuation effect' of the increased domestic currency price of imports.

The Marshall–Lerner (ML) condition shows us the minimal levels of the price elasticities of export and import demand that are required for the net effect of a rise in $E$ on $NX$ to be positive, that is, for the quantity responses to outweigh the valuation effect. Assuming that supplies are infinitely elastic (as shown in Figure 9.1) and prices are fixed in the sellers' currencies (so we can take $P$ and $P_f$ as given), and also assuming for simplicity that trade is initially balanced ($NX = 0$), we can substitute equations (9.1) and (9.2) into the definition of $NX$ and find that

$$\partial NX/\partial E > 0 \text{ if and only if } \varepsilon_X + \varepsilon_M > 1.$$

In other words, the price elasticities of demand for exports and imports (recalling that these were defined so as to be positive) must sum to more than unity, which is considered a relatively weak condition since neither the demand for exports nor the demand for imports need be price-elastic (elasticity greater than one) by itself; only the sum of the two elasticities must exceed one. It must be emphasized that this is only a partial or direct effect of the depreciation (as indicated by the use of the partial derivative signs), because it holds constant other factors that could also be affected by the depreciation. One of these factors is the cost of imported inputs, which can affect domestic prices (including for exported goods); the impact of a devaluation on the costs of imported intermediate goods will be addressed in Chapter 10 (section 10.6).

The ML condition can easily be generalized to the more realistic case in which a country that devalues has an initial trade imbalance. To do this, it helps to define the trade balance in ratio form as $\theta = PX/EP_fM$, where trade is in surplus, balanced or in deficit as $\theta > 1$, $\theta = 1$ or $\theta < 1$. By using the definition of $\theta$ in the solution for $\partial NX/\partial E$, it is easily seen that the more general ML condition is that $\partial NX/\partial E > 0$ if and only if $\theta\varepsilon_X + \varepsilon_M > 1$. Thus, if a country has a trade deficit ($\theta < 1$) at the time when it devalues, as is often the case, the price elasticities would have to be somewhat higher in order for the trade balance (measured in domestic currency) to improve than they would have to be if the country started out with balanced trade.

However, the ML condition does not generalize so easily when other assumptions are violated. For example, if changes in nominal exchange rates are passed through only partially into export prices (in foreign currency) or domestic currency prices of imports, then the analysis becomes more complex and any improvement in the trade balance depends on the degree of 'partial pass-through' as well as the price elasticities of demand. Also, if the assumption of infinitely elastic supplies (horizontal supply curves) for exports and imports is dropped, the ML condition has to be replaced with the much more general (and much more complicated) Bickerdicke–Metzler–Robinson condition, which incorporates both supply and demand elasticities (see Robinson, 1947; Dornbusch, 1980). In addition, if income distribution is affected by a currency depreciation (for example, because profit markups are endogenous as in the open economy neo-Kaleckian model of Chapter 4), then the analysis becomes further complicated and one must take the resulting shift in the profit (or wage) share into account (see section 4.4.3).

Finally, whether the ML condition is typically satisfied for most countries in the real world is hotly debated, as will be discussed in the next chapter. One point to flag for future reference in this regard is that estimates of the price elasticities $\varepsilon_X$ and $\varepsilon_M$ that assume horizontal (infinitely elastic) supply curves for exports and imports could possibly suffer from simultaneity bias, in which case they would be biased downward, if the supply curves are in fact upward sloping. Hence, empirical estimates of $\varepsilon_X$ and $\varepsilon_M$ that do not control for possible endogeneity of the RER $(EP_f/P)$ or relative prices may be biased against confirming that the ML condition holds.

# 10 Balance-of-payments-constrained growth II: critiques, alternatives and syntheses

## 10.1 Introduction

The balance-of-payments-constrained growth (BPCG) model presented in the previous chapter has become one of the most commonly used 'workhorses' in heterodox macroeconomics in recent years. The BPCG modelling approach has become especially popular for studying economic development issues and the convergence (or lack of convergence) of developing countries in the global 'South' to the (per capita) income levels achieved by the advanced economies in the global 'North'. Nevertheless, the BPCG framework has also become the subject of considerable controversy and debate. Critics have focused on a number of perceived weaknesses of this approach and have proposed a variety of alternative models in response. Even economists sympathetic to the BPCG approach have developed new versions of the model (beyond the extensions covered in the previous chapter) in response to what they acknowledge are deficiencies or omissions in the standard models. Concern has also arisen about differences between the BPCG framework and other heterodox approaches covered earlier in this book. This has prompted efforts at theoretical reconciliation, especially with the other branch of Kaldorian growth theory – the export-led cumulative causation (ELCC) model covered in Chapter 8.

In particular, the de-emphasis on relative prices, real exchange rates (RERs) and cost competition in the BPCG approach has come under criticism from a variety of sources. As discussed in Chapter 9, the traditional BPCG model (as incarnated in Thirlwall's law) rejects the idea that persistent or continuous changes in relative prices or real exchange rates could be a plausible explanatory factor for determining the long-run growth rate consistent with balance-of-payments (BP) equilibrium. However, as discussed in Chapter 8 (section 8.7.3), mounting empirical evidence suggests that the *levels* of relative prices, the RER or relative unit labour costs (RULC) – as opposed to their rates of change – do, in fact, matter to long-run growth performance.[1] For larger countries, adjustments in relative prices may be necessary accompaniments to the income adjustments that bring about BP equilibrium. For smaller countries, RER levels may affect investment and the growth of export capacity. In all countries, RER levels may influence the industrial structure and rate of technological progress. Recognition of this last point has led to a proliferation of new BPCG models that allow for some long-run influence of the RER, particularly (and somewhat paradoxically) by affecting the aggregate *income* elasticities of exports and imports – sometimes indirectly, by affecting the composition of the goods produced. Some of these newer models can be seen as extending the traditional 'Thirlwall's law' solution, while others are clearly intended to serve as alternatives to it. Some of these modified or alternative models imply that endogenous technological progress of the sort found in the ELCC models of export-led growth can also play a role even if BP constraints are taken into account and that both price and non-price competition can matter to long-term economic performance.

The rest of this chapter is organized as follows. Section 10.2 discusses several critiques of the BPCG theoretical model and empirical tests thereof. Section 10.3 focuses particularly on the argument that the income elasticities of demand for imports and exports – which are crucial but exogenous parameters in standard renditions of Thirlwall's law – should be treated as endogenous. Section 10.4 discusses efforts to analyse how the BP-equilibrium growth rate can be reconciled with the natural rate of growth, which (as noted in earlier chapters) is the rate of output growth required to ensure a constant employment rate in the long run. Section 10.5 presents alternative models of BP constraints for small countries and what we will call 'medium-large' countries. Section 10.6 covers a new way of reconciling the BPCG model with the ELCC model from Chapter 8, which also incorporates a new approach to wage and price adjustment in moving from the short and medium run to the long run. Section 10.7 concludes with some observations on future directions for research in this paradigm.

## 10.2 Critiques and defences of Thirlwall's law[2]

Thirlwall's law has been subjected to various critiques over the years since its first formulation in 1979. This section will discuss some of the most frequent critiques and the responses by defenders of Thirlwall's approach, as well as the implications of the critiques being accepted or rejected for both theory and policy. In order to focus on the core issues, the discussion in this section is deliberately limited to the most basic types of BPCG models from section 9.2 in Chapter 9, omitting the extensions covered in sections 9.3 and 9.4 (although some of those will be resurrected later in this chapter). These critiques and responses (as well as those extensions) have led to a proliferation of new types of models of BP constraints, examples of which will be covered later in this chapter.

### 10.2.1 Testing a near-tautology?

One of the earliest criticisms, which has recently been revived, questions whether the most common empirical methods for testing Thirlwall's law are merely testing something that is virtually a tautology or identity. Most empirical tests of this law have focused on determining whether actual, long-run average growth rates are close to (or insignificantly different from) the BP-equilibrium growth rates predicted by the BPCG model. Using a wide variety of econometric methodologies, numerous empirical studies have found that actual average growth rates are close to BP-equilibrium growth rates for the vast majority of countries and (long-run) time periods considered (see Thirlwall, 2011 for a survey and references).[3] However, McCombie (1981) and more recently Clavijo and Ros (2015a) and Razmi (2016a) have claimed that such tests of Thirlwall's law are testing a near-identity that is likely to be satisfied for almost any country regardless of whether its growth is BP constrained in the sense of Thirlwall or not. We follow Clavijo and Ros's presentation here, but the essence of all these arguments is the same.

Consider the strong and weak versions of Thirlwall's law represented by equations (9.9) and (9.10) in Chapter 9, which are reproduced here for convenience.

$$y_B = \frac{\eta_X y_f}{\eta_M} \tag{10.1}$$

$$y_B = \frac{x}{\eta_M} \tag{10.2}$$

According to the near-tautology (or near-identity) argument, econometric estimates of the income elasticities $\eta_X$ and $\eta_M$ are likely to approximate the ratios of the growth rate of each trade variable (the volume of exports or imports) to the corresponding income growth rate (foreign or domestic), that is, $\eta_X \approx x/y_f$ and $\eta_M \approx m/y$ – especially if relative price effects are negligible in the long run. Then, it is easy to see that either equation (10.1) or (10.2) is equivalent to

$$\frac{y_B}{y} \approx \frac{x}{m} \tag{10.3}$$

Hence, a statistical test of whether $y = y_B$ is equivalent to a test of whether $x = m$.

Thus, Thirlwall's law will appear to be confirmed (that is, the null hypothesis that $y = y_B$ will not be rejected) for any data set in which $x \approx m$, that is, the quantities of exports and imports grow at approximately the same rate. However, the longer the time period considered, the more likely it is that the latter will be true in almost any country's data. For example, even though the US trade balance for goods and services shifted from a surplus of 0.5 per cent of gross domestic product (GDP) in 1967 to a deficit of 2.9 per cent of GDP in 2015, nevertheless the average annual growth rates of real exports and imports were nearly identical over the 1968–2015 period (5.67 and 5.53 per cent respectively for goods and services, or 5.82 and 5.95 per cent for goods alone).[4] As a very large economy that experienced a shift of −3.4 percentage points of GDP in its net exports over this almost half-century period, it is difficult to maintain that the US economy was BP constrained in any meaningful sense. Clearly, the US was able to obtain sufficient net financial inflows to sustain an increasing trade deficit as its growth exceeded a rate that would have been consistent with maintaining balanced trade (Blecker, 2013a). Moreover, the US economy is usually regarded as a demand-driver for the entire global economy, constrained mainly by its own domestic aggregate demand (consumption, investment and government spending). As noted in Chapter 9 (section 9.2.4), Thirlwall has always acknowledged that some countries must be unconstrained by their BP in order for other countries to be so constrained, and the largest economies like the US and China are prime candidates for *not* being BP constrained.[5] Nevertheless, any standard empirical test of Thirlwall's law for the US will appear to confirm the law because $x \approx m$ in the US long-run data.

Thirlwall (1981) responded to this critique – in an argument reiterated more recently by McCombie (2011) – by pointing out that econometric estimates of the income elasticities $\eta_X$ and $\eta_M$ need not equal the observed ratios of growth rates $x/y_f$ and $m/y$, respectively, as long as relative prices are controlled for in the estimated demand functions for exports and imports. That is, of course, a valid point. However, McCombie (2011, p. 357) – who has long since accepted Thirlwall's response to his own critique of 1981 – argues that relative price effects on export or import demand are small or insignificant in most studies,[6] implying 'that it is not relative prices that, for example, cause imports to adjust, but changes in income in a Keynesian manner'. McCombie (2011) is certainly correct in saying that 'inclusion of the [relative price] term means that the law is not an identity', but if anything it would seem logical that if relative price effects are small or insignificant, then the estimated income elasticities should be *closer* to the observed ratios $x/y_f$ and $m/y$ than if relative price effects were large and significant. In any event, the contention of the critics is that this method of testing Thirlwall's law amounts to testing a near-identity, not an exact one, and it remains to be seen empirically how much standard estimates of income elasticities differ from the ratios of growth rates.

Another defence comes from Pérez Caldentey (2015, p. 58), who argues that identifying the mechanism that achieves equilibrium between the growth rates of exports and imports constitutes the 'essence' of Thirlwall's law, which 'establishes that it is through variations in the level (or growth rate) of income that an equilibrium between $x$ and $m$ is achieved, not by an adjustment in relative prices'.[7] This is indeed a correct statement of the essence of Thirlwall's law; nevertheless, finding that $x \approx m$ does not necessarily demonstrate that adjustment occurs through variations in income growth rather than relative prices or other mechanisms (for example, supply-side adjustments). In short, this defence does not vindicate the use of tests of equality between actual and BP-equilibrium growth rates as meaningful tests of the causal story implied by Thirlwall's law.

It is important to recall, however, that the 'near-identity' critique is *only* a criticism of certain types of *empirical tests* of Thirlwall's law. This critique suggests that statistical tests of equality between the actual and BP-equilibrium growth rates have weak power to reject the null hypothesis that these growth rates are equal. Even if it is accepted, this critique does *not* necessarily disprove Thirlwall's law; it simply implies that other, more powerful statistical tests are required to validate it. These more powerful tests are found in the (relatively fewer) studies that have more directly tested what Pérez Caldentey correctly calls the essence of the law: whether BP equilibrium is achieved through adjustments in national incomes rather than relative prices.

Along these lines, Alonso and Garcimartín (1998–99, pp. 266, 276) noted that the hypothesis of Thirlwall's law 'cannot be tested through the degree of correlation between [the] actual and Thirlwall's Law rate[s] of growth', and instead proposed to explicitly 'test the balance-of-payments constraint hypothesis by identifying the variable by means of which the balance-of-payments equilibrium is achieved'. Alonso and Garcimartín tested the two alternative hypotheses (income adjustment versus relative price adjustment) for ten industrialized countries. They found that the income adjustment mechanism was statistically significant and had the right sign (income grows more rapidly in response to a rising trade surplus, or more slowly in response to a widening deficit) in eight of the ten countries considered (the US and France were the two exceptions); in contrast, the price adjustment mechanism was statistically insignificant for all countries studied.

Some additional evidence along these lines comes from studies that have applied cointegration methods to test the BPCG hypothesis, beginning with the work of Moreno-Brid (1999) on Mexico and Pérez Caldentey and Moreno-Brid (1999) on Central America. More recently, Lima and Carvalho (2008) found that national income and exports are significantly cointegrated with each other but not with the RER, using annual time-series data for Brazil. This study covers a long period (1930–2004), but given its use of annual data, the adjustment processes for which it tests are likely medium run in nature. Earlier, Razmi (2005, p. 668) found that price variables were sometimes significant in cointegrating vectors for India, but in some estimates they adjust 'in the "wrong" – that is, disequilibrating – direction', and he also found that equilibrium for the import relationship was reached 'in approximately four years'. In any case, the results of cointegration studies need to be interpreted with caution, because the finding that a cointegrating vector exists does not suffice to prove the direction of causality between the variables. Also, to the best of our knowledge, no one has formally tested for adjustments over longer data frequencies (for example, five- or ten-year periods), which would require the use of panel data for large numbers of countries.[8]

### 10.2.2 Foreign income growth and domestic capital accumulation

Assuming that income elasticities of export and import demand are relatively stable over time, as most BPCG theorists have assumed,[9] the strong version of Thirlwall's law (equation 10.1) implies that we should observe a significant positive correlation between individual countries' growth rates and foreign growth rates. Razmi (2016a) finds that this is not the case, using a data set comprising 167 countries with data averaged for five-year periods to focus on long-run relationships.[10] First, he shows graphically that the raw correlation between individual country growth rates and world growth rates is not generally positive; for most of the countries in his sample, there is simply no correlation (the individual country growth rates are independent of world growth), and for almost a third of the sample the correlation is anomalously negative.

More formally, Razmi (2016a) tests for the statistical significance of foreign growth effects by estimating an econometric model explaining national growth rates, using panel data for the 167 countries with five-year time periods. The foreign (world) growth variable always has a positive sign and is statistically significant in most estimates, but the magnitude of its coefficient drops notably when the domestic capital accumulation rate is included in the model – and the world growth rate becomes insignificant when the generalized method of moments (GMM) is used to control for endogeneity. When both variables are included, the coefficient on the capital accumulation rate always exceeds the coefficient on the world growth rate. A counter-argument (by defenders of the BPCG approach) could be that investment is endogenous and responds (via the accelerator mechanism) to domestic income growth, which in turn is driven by exports – in which case the direction of causality between capital accumulation and income growth would be the opposite of what Razmi's model assumes. Also, foreign income growth retains a positive coefficient (which is significant in some specifications) in Razmi's estimates, even after other variables are controlled for, and a trade-weighted measure of foreign income (or foreign expenditures on imports) for each country might be a better measure than total world income for identifying foreign income effects.

## 10.2.3 Level versus rates of change in relative prices

As discussed in Chapter 9, relative price (RER) effects are usually dismissed in the BPCG literature by asserting either 'elasticity pessimism' ($\varepsilon_X + \varepsilon_M \approx 1$) or else that relative prices (RERs) do not change significantly in the long run (in which case $\hat{E} + \hat{P}_f - \hat{P} \approx 0$).[11] Empirical evidence on elasticity pessimism is mixed, however. Some studies (for example, Cline, 1989) find that the Marshall–Lerner (ML) condition ($\varepsilon_X + \varepsilon_M > 1$) is satisfied for most countries, while others (for example, Alonso and Garcimartín 1998–99) find that it does not hold (and elasticity pessimism is validated) for most countries.[12] In contrast, it is quite easy to claim that $\hat{E} + \hat{P}_f - \hat{P} \approx 0$ should hold in the long run. While there are substantial shifts in RERs over short and medium horizons, there is increasing evidence that RERs are mean-reverting over very long periods, especially when non-linearities and shifting long-run equilibria are taken into account (Taylor and Taylor, 2004), and it's not credible to view relative prices as continuously rising or falling in the very long run. As McCombie (2011, p. 358, emphasis in original) states, 'even if the Marshall-Lerner conditions are satisfied, to increase permanently the *growth* of exports and to reduce the growth of imports would require a *continuous* depreciation of the currency, which is implausible'. Thus, the assumption of constant relative prices (or a stationary RER) has been the primary basis for assuming that the relative price (RER) effects in equation (9.8) are negligible and hence relying on equation (10.1) or (10.2) (the same as equations 9.9 and 9.10) instead.

But even if relative prices (RERs) don't change continuously in the long run, this does not necessarily imply that the *level* of the RER may not have a significant impact on a nation's growth. As discussed in Chapter 8 (section 8.7.3), many empirical studies (for example, Rodrik, 2008; Rapetti et al., 2012; Berg et al., 2012; Razmi, 2016a), using a variety of methodologies, have found that RER levels (or degrees of undervaluation relative to estimated equilibrium levels) have significant effects on economic growth in many countries – especially developing countries and, in some studies, industrialized countries as well. Razmi (2016a) expressly compares the level and rate of change in the RER in his econometric model and, for the most part, finds that the RER level (measured as the degree of undervaluation) is more statistically significant in explaining countries' growth (averaged over five-year periods) than the rate of change in the relative price (especially in estimates using GMM or two-stage least squares methods to control for simultaneity). Similarly, as discussed in Chapter 8 (section 8.7.3), Boggio and Barbieri (2017) find that levels of RULC are statistically significant in explaining changes in countries' export market shares, even though rates of change in RULC are not, in the framework of an export-led growth model of the Beckerman (1962) type. Thus, even if it is not realistic for a country to gain a continually increasing competitive advantage by continuously depreciating its RER or lowering its RULC, it is entirely plausible – indeed, empirically supported – that countries that have maintained undervalued levels of their RERs (or low RULC) for substantial periods of time have obtained sustained long-term growth benefits (including, according to Berg et al., 2012, longer durations of rapid growth 'spells') as a result.

### 10.2.4 Country size

As noted in Chapter 9, the standard BPCG model assumes a 'Keynesian small economy' in the sense of Branson (1983, p. 48): a country that has infinitely price-elastic supplies of both exports and imports. The assumption of an infinitely elastic supply of imports is not controversial for most countries, except the very largest (for example, US, China) whose demand can possibly affect global prices. But critics of the traditional BPCG approach have argued that assuming an infinitely elastic supply of exports is unrealistic for most countries, which are small players (and hence price-takers) in their export markets. In an argument anticipated by McGregor and Swales (1985, p. 21), more recently Ros (2013), Clavijo and Ros (2015a) and Razmi (2016a) contend that the 'small economy' model is more appropriate for many if

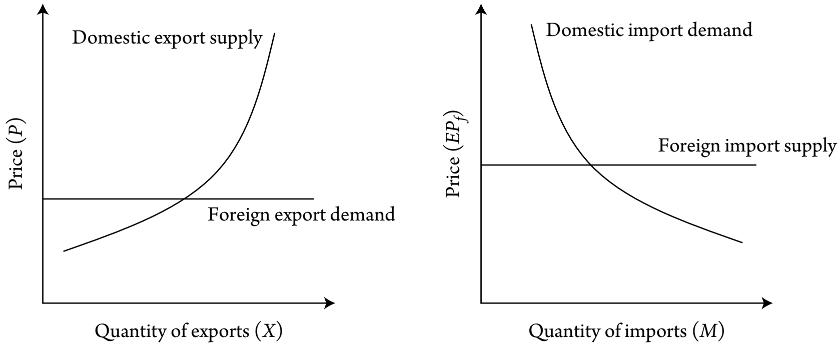



**Figure 10.1** Markets for exports and imports in a price-taking 'small economy' (infinitely elastic demand for exports and supply of imports)

not most countries (especially developing nations): they are price-takers in both export and import markets, which implies that they have infinitely elastic supplies of imports and infinitely elastic *demand* for exports. This in turn implies that the equilibrium quantity of exports must be determined by supply constraints in the exporting country's industries.

The market structures for the small country case are shown in Figure 10.1. The import market is the same as in the Keynesian small open economy discussed earlier (compare Figure 9.1): the supply of imports is infinitely elastic, the price of imports is fixed in foreign currency and the demand for imports depends on national income and the RER. However, in a small country it is demand rather than supply that is infinitely elastic in the export market, as shown by the horizontal export supply curve in Figure 10.1. Essentially, a small country can sell all the export goods that it can possibly produce in global markets at a given world price in foreign currency. The equilibrium quantity of exports is therefore supply constrained, as the small country's ability to produce exported goods is limited by its natural resources, capital stock, technology, infrastructure and so on. In this situation, the strong version of Thirlwall's law cannot apply because it requires a downward-sloping demand curve (with a finite price elasticity) for exports. Small countries may still be subject to BP constraints on their growth, but if so these constraints depend critically on the (exogenous) world terms of trade for the countries' exports and the accumulation of capital in their export industries, rather than the growth rate of foreign income.[13] An alternative BPCG model for a small country is discussed in section 10.5.1, below.

Nevertheless, one may question how accurately the conventional small country model describes the long-run situation for actual smaller nations in the global economy. Of course, export supplies are limited at any point in time by the capacity of a country's export industries, which depends on the available capital, skills and technology. But over long periods of time, capital can be accumulated, skills can be acquired and technologies can be improved – often in mutually complementary and reinforcing ways. Many typical export factories (for example, apparel sweatshops, electronics assembly plants) can easily be relocated or replicated, creating what is effectively a highly (if not infinitely) elastic long-run supply function for exports (although the very smallest countries may still be limited in their export capacity by virtue of their size). Also, the price-taker specification assumes that domestic and foreign goods are perfect substitutes, which may or may not be a good approximation to reality depending on the nature of a country's export products.

Furthermore, even if individual exporting countries can be regarded strictly as small economies facing infinitely elastic export demand curves at exogenously given world prices, they may collectively be subject to a 'fallacy of composition' or 'adding-up constraint' (Blecker and Razmi, 2008, 2010). That is, if a large number of such countries attempt to export similar products simultaneously, they may be large enough as a group that an increase in their total supply will depress world prices of their export products. Viewing such exporting economies as a group, they collectively face a downward-sloping export demand curve, and because of the easy ability to ramp up production of similar products across a wide spectrum of exporting nations, total supplies from such a group of countries may effectively be infinitely elastic in the long run. Technically speaking, these countries individually still face infinitely elastic demand curves for their own exports (so the small country model still applies to each country by itself), but as a practical matter they cannot expect to sell any quantity of exports they can produce at a constant price when rival competitor nations are engaged in similar export-promotion efforts in the same product lines at the same time.

However, the fallacy-of-composition argument does not rescue the traditional solutions for Thirlwall's law, in which relative prices are irrelevant. When many countries are exporting similar types of products, what might be called their 'cross exchange rates' (their relative prices or RERs relative to competing exporting nations) become important determinants of their export success and, hence, their growth (Blecker, 2002b; Blecker and Razmi, 2008, 2010). In these cases, too, levels of RERs can be important determinants of long-term growth outcomes. The fallacy-of-composition argument also implies the importance of countries finding distinctive product niches in their export production, instead of seeking to replicate the export industries that are already established (or likely to be established) in other exporting nations.

## 10.3 Endogeneity of income elasticities

All of the BPCG models discussed up to this point – with the sole exception of the multisectoral version covered in section 9.3.2 in Chapter 9 – have taken the income elasticities of export and import demand ($\eta_X$ and $\eta_M$) as exogenously given parameters, which effectively constrain long-run growth according to Thirlwall's law (either equation 10.1 or 10.2). However, this assumption was questioned long ago by Krugman (1989), and more recently has been modified in different ways in new analyses by various economists. This section will consider Krugman's argument and later approaches to endogenizing the income elasticities in turn.

### 10.3.1 Krugman's argument: reversing the causality

Without explicitly citing Thirlwall (1979) or any of the subsequent BPCG literature, Krugman (1989) noted the key empirical regularity that has long been cited in support of Thirlwall's law: the fact that proportional differences in countries' growth rates are closely approximated by the relative income elasticities of their respective exports (where home country imports are the same as foreign country exports). Based on the strong form of Thirlwall's law in equation (10.1), if we assume that a 'home' country's growth rate equals its BP-equilibrium growth rate (that is, $y = y_B$) in the long run, then we should find that

$$\frac{y}{y_f} = \frac{\eta_X}{\eta_M} \tag{10.4}$$

Krugman dubbed this the '45-degree rule' because, if the ratios on the two sides of equation (10.4) are plotted on a graph for a sample of countries, the data points should lie approximately along the diagonal of the first quadrant. Krugman verified that this was true using both the estimates of income elasticities from Houthakker and Magee (1969) – the same ones originally cited by Thirlwall (1979) – and more recent estimates that Krugman made using data from the 1970s and 1980s.

Then, Krugman observed that relative purchasing power parity (PPP) tends to hold (that is, RERs have constant trends, or $\hat{E} + \hat{P}_f - \hat{P} \approx 0$) in the long run, in which case the neoclassical solution in which differences in income elasticities are offset by secular trends in RERs (as in equation

9.11 in Chapter 9) is not supported empirically. But Krugman refused to consider the possibility that income growth rates could be the adjusting variables:

> I am simply going to dismiss a priori the argument that income elasticities determine economic growth, rather than the other way around. It just seems fundamentally implausible that over stretches of decades balance of payments problems could be preventing long term growth, especially for relatively closed economies like the U.S. in the 1950s and 1960s. Furthermore, we all know that differences in growth rates among countries are primarily determined in [*sic*] the rate of growth of total factor productivity, not differences in the rate of growth of employment; it is hard to see what channel links balance of payments due to unfavorable income elasticities to total factor productivity growth. (Krugman, 1989, p. 1037)

Krugman then argued that, on the contrary, conventionally measured income elasticities could in fact adjust to differences in growth rates, based on an adaptation of one of his models of international trade with monopolistic competition (Krugman, 1980). In that model, countries trade in differentiated products produced by firms under conditions of increasing returns to scale, and the number of goods (each produced by a single firm) produced in each country is proportional to the country's size (as measured by its labour force, since labour is the only input and full employment is assumed). Krugman (1989) assumed that growth comes in the form of an (exogenously) increasing labour force rather than rising total factor productivity, but the important point is that it occurs at an exogenously given natural rate.[14] The intuition for Krugman's (1989) model was well summarized in Thirlwall's (1991, p. 25) response as follows:

> If, then, the labor force grows at different rates between countries, the faster-growing country will be able to increase its share of world markets by increasing the number of goods faster than other countries, allowing it to sell more without a reduction in its relative prices, therefore giving the fast-growing country an apparently higher income elasticity of demand for its exports.

Mathematically, if there are two countries, home and foreign, with balanced trade, their exports and imports must increase at the same rate, which is an income-weighted average of their respective growth rates:

$$x = m = y\left(\frac{Y_f}{Y + Y_f}\right) + y_f\left(\frac{Y}{Y + Y_f}\right) \qquad (10.5)$$

Here, each country's growth is weighted by the share of the other country in world income because the other country is the market for each one's exports. Then, the conventionally measured income elasticities would be

$$\eta_X = \frac{x}{y_f} = \left(\frac{y}{y_f}\right)\left(\frac{Y_f}{Y + Y_f}\right) + \left(\frac{Y}{Y + Y_f}\right) \tag{10.6}$$

and

$$\eta_M = \frac{m}{y} = \left(\frac{Y_f}{Y + Y_f}\right) + \left(\frac{y_f}{y}\right)\left(\frac{Y}{Y + Y_f}\right) \tag{10.7}$$

and the ratio of equation (10.6) to (10.7) yields (after some manipulations and cancellations) equation (10.4) or, equivalently, the strong form of Thirlwall's law (10.1). But in this case, the income elasticities are functions of the growth rates, not the other way around.

Krugman's argument can be criticized on many grounds (see Thirlwall, 1991; McCombie, 2011). However, by far the most fundamental problem is his assumption of exogenously given long-run growth rates. As discussed in earlier chapters, the so-called natural rate of growth (equal to the sum of the growth rates of labour supply and labour productivity) is an endogenous variable in numerous theories, both mainstream and heterodox. In some of the classical-Marxian models discussed in Chapter 2, labour supply is the adjusting variable. Alternatively, the rate of labour-saving technical progress or labour productivity growth is treated as endogenous in many models, ranging from the neoclassical endogenous growth theory (NEGT) discussed in Chapter 1 to the Kaldorian ELCC models covered in Chapter 8. The latter models explicitly make productivity growth a function of output growth per Verdoorn's law, in which case the former is not independent of the latter. Hence, Krugman's assumption that the natural rate of growth must be treated as exogenous is untenable, and without that assumption his causal story breaks down.

Curiously, Krugman's (1989) article was published just around the time when mainstream growth theorists were beginning to abandon the Solow–Swan assumption of an exogenously given rate of technological change. As discussed in Chapter 1 (section 1.4.1), the NEGT literature, beginning with Romer (1986, 1990), showed various ways in which the rate of technological progress – and hence the natural rate of growth – could become endogenous in response to changes in variables such as saving rates and stocks of human capital in models that have thoroughly neoclassical pedigrees (for example, by assuming intertemporal optimization by households, profit-maximization by firms and other elements of neoclassical rationality). Ironically, some of the NEGT models derive their results by assuming increasing returns, as Krugman did in his innovative models of trade with monopolistic competition (Krugman, 1979, 1980). But Krugman (1989) did not anticipate how the incorporation of increasing returns into neoclassical growth models could upset conventional Solow–Swan thinking about the exogeneity of the long-run equilibrium (natural) rate of growth, just as their incorporation into trade models upset traditional Heckscher–Ohlin thinking about comparative advantages.

Of course, one must still show how export performance and the BP could constrain long-run growth while influencing the process of technological change (hence making the natural rate of growth endogenous). The ELCC model in Chapter 8 showed one important channel: faster growth of output – itself induced by rapid growth of exports – can lead to faster growth of labour productivity via induced technological change and dynamic increasing returns (Verdoorn's law). We have already seen (in section 9.4.2 in Chapter 9) that this same mechanism can operate in the presence of a BP constraint, albeit in a medium-run time frame in which relative prices can change; more BPCG models incorporating Verdoorn's law will be covered later in this chapter. Similarly, NEGT models for open economies have identified various ways in which a country's external orientation can influence its rate of technological progress – for example, whether a country specializes in innovative activities or not, as in the model of Grossman and Helpman (1991). And, as noted in Chapter 1 (section 1.5.2), there are now some NEGT models that identify channels through which depressed aggregate demand can lower potential output, for example by diminishing investment in human or physical capital or by reducing the rate of technological innovation. Thus, if BP constraints force a country to restrain domestic demand and slow its output growth, the results could be the same as these various models would predict: there would be fewer opportunities to exploit increasing returns, investment in technological innovation and human capital would suffer, and productivity growth would diminish accordingly.

The new models of international trade with heterogeneous firms (starting with Melitz, 2003) also introduce an important element that was not found in Krugman's early models of trade with monopolistic competition, which assumed for simplicity that all firms had identical costs and productivity. By recognizing that firms differ in their cost functions and productivity levels, this new literature shows that increased openness to trade or increased penetration of export markets can lead to endogenous increases in average productivity levels because trade selects for the most productive firms, inducing them to grow and expand while forcing less productive firms to exit (at the cost of possible job losses and/or increased inequality in some versions of these theories). Indeed, these new trade models can be seen as providing a type of microeconomic rationale for the Kaldorian idea of positive feedbacks from export growth and structural change to productivity increases at the aggregate level, albeit one based more on firm-level compositional shifts within industries rather than Kaldorian 'dynamic increasing returns'. For all these reasons, Krugman's (1989) assertion that the BP constraint could not have feedback effects on the long-run growth rate is untenable.

Nevertheless, to paraphrase Marx's (1867 [1976], p. 103) famous remark about the Hegelian dialectic, there may be a 'rational kernel within the mystical shell' of Krugman's argument about income elasticities being endogenous variables. Even without conceding the exogeneity of the output growth rate in the long run, it can surely be acknowledged – in the Kaldorian spirit of 'circular and cumulative causation' – that the average, aggregate income elasticities of export and import demand could also be endogenous in the long-term evolution of a nation's economy. As discussed in Chapter 8, the process of economic growth typically involves profound structural changes, in which employment and production shift from agriculture to manufacturing and later (or simultaneously) to services. Then, as discussed in relation to the multisectoral Thirlwall's law in section 9.3.2, structural changes that affect the industrial composition of a country's exports and imports will change the weighted average income elasticities, since the weights are industry shares of exports or imports. Thus, the aggregate-level income elasticities should not be taken as immutably fixed parameters in a BPCG model, but rather should be treated as endogenous variables that can change over long periods of time as industrial structures and trade specializations evolve.

## 10.3.2 Heterodox alternatives: relative wages and structural change

Several recent articles have exploited the idea of endogenous income elasticities in extended BPCG models that incorporate structural change. One notable example is Cimoli and Porcile (2014), who proposed a model of North–South trade designed to incorporate the neo-Schumpeterian idea of technological gaps.[15] Since a discussion of their complete model would be beyond the scope of this chapter, we will focus here on the way they endogenize the income elasticities in the BCPG component of their model. Consider the equation for Krugman's 45-degree rule (which is equivalent to the strong form of Thirlwall's law), but viewing the home country as the developing South ($s$) and the foreign country as the developed North ($n$):

$$\frac{y^s}{y^n} = \frac{\eta_X}{\eta_M} \tag{10.8}$$

where $y^s$ and $y^n$ are the BP-equilibrium growth rates for the two regions and $\eta_X$ and $\eta_M$ refer to the income elasticities of Southern exports and imports, respectively.

Cimoli and Porcile then assume that the ratio of the income elasticities is a function of the diversification and technological complexity of the productive structure of the Southern economy, which they assume is an increasing function of the number of goods that the South produces $N^s$:

$$\eta_X/\eta_M = \Phi(N^s), \Phi' > 0 \tag{10.9}$$

Assuming for simplicity that all goods are tradable, $N^s$ is also the number of goods that the South can export, which in turn is determined by the Ricardian model of trade with multiple commodities originally developed by Dornbusch et al. (1977) and later extended by Cimoli (1988), Dosi et al. (1990) and Cimoli and Porcile (2010). In this model, any good $i$ is produced in the South if the South's relative productivity in the good is greater than the South's relative wage, or

$$\frac{a_{0i}^n}{a_{0i}^s} \geq \frac{W^s}{EW^n} \tag{10.10}$$

where $a_{0i}^j$ is the labour coefficient for good $i$ in country $j$ (so $j$'s labour productivity in good $i$ is $1/a_{0i}^j$), $E$ is the nominal exchange rate (in Southern currency per unit of Northern currency), $W^j$ is the nominal wage in country $j$, and $W^s/EW^n$ is the relative Southern wage (which can be thought of as the RER in terms of labour instead of goods, that is, the relative cost of Southern labour). If there is a large total number of goods $\overline{N}$, and the goods are arrayed from greatest to least comparative advantage of the South, then goods 1 to $N^s$ are produced in the South while goods $N^s + 1$ to $\overline{N}$ are produced in the North, and under Cimoli and Porcile's assumptions a rise in $N^s$ (that is, a shift in global production to the South) will raise the ratio of income elasticities $\eta_X/\eta_M$ and the BP-equilibrium relative growth rate of the South, $y^s/y^n$. This occurs fundamentally because when the South acquires a more diversified production structure, it not only moves into exporting goods with higher income elasticities, but also is able to reduce the need for many imports by producing more imported goods at home (which reduces the income elasticity of import demand).[16]

There are two ways in which the South could increase the number of goods it produces $N^s$. First, the South could improve its productivity by lowering the labour coefficients $a^s_{0i}$ for some goods that it formerly imported, either through the adoption of foreign technology or the improvement of domestic technology. Second, for any given set of labour coefficients, the South could depreciate its RER in terms of wages (via a rise in $E$ or fall in $W^s/W^n$). Thus, the level of the RER in terms of wages affects the number of goods produced, which in turn alters the relative income elasticity of export demand to import demand. This model thus introduces a channel whereby the level of a key relative price (the relative wage, or RER in terms of labour) can influence the long-run BP-equilibrium growth rate, even though relative prices do not continuously change in the long run. Hence, the Cimoli and Porcile (2014) model can help to reconcile the BPCG model with the empirical findings, referred to earlier, of positive effects of RER undervaluation (or lower RULC) *in levels* on long-run growth rates.

However, Ribeiro et al. (2016, 2017a, 2017b, 2018) introduce several sceptical notes about the efficacy of devaluation policies or undervalued RERs for promoting higher BP-equilibrium growth rates. First, even in the short run, the competitive benefits of an RER depreciation can be offset by the resulting increased costs of imported intermediate inputs used for domestic production, on the one hand, and redistributive effects (raising the profit share and lowering the wage share), on the other. Increased costs of imported intermediate goods can lessen or, in the extreme case, offset the competitive gains from a currency devaluation (especially if accompanied by higher domestic wages and inflation). The redistributive effects can have a different impact depending on whether an economy has wage-led or profit-led demand (as noted in section 4.4.3 of Chapter 4). If demand is profit-led so that the increase in the profit share is expansionary, then a devaluation that increases the profit share will increase demand for imports via an income effect, which offsets the competitive benefits of a depreciated RER for the trade balance (net exports). For the long run, Ribeiro et al. (2016) develop a model in which the income elasticity ratio $\eta_X/\eta_M$ depends on the wage share of national income and the inverse of the technological gap (the ratio of the stock of technical knowledge in the South to the North), both of which are also endogenous variables. This model (the details of which are too complex to try to summarize in a short space here) generates a 'technological progress regime' that can be either wage-led (if a higher wage share leads to a more rapid pace of labour-saving technical change) or profit-led (if a higher profit share creates strong incentives to invest in productivity-enhancing innovations), depending on various parameter values and initial conditions (since the model is nonlinear). As a result, the long-run BP-equilibrium growth rate $y_B = (\eta_X/\eta_M)y_f$ can either rise or fall in response to a currency devaluation.

## 10.4 Reconciling the BP-equilibrium and natural growth rates

One obvious question about Thirlwall's law is whether the standard solution for the BP-equilibrium growth rate is consistent with the natural rate of growth, which – as discussed extensively in earlier chapters – is the rate that maintains either full employment of the labour force or a constant employment rate in the long run. If the BP-equilibrium growth rate is greater than the natural rate of growth, the economy would eventually run up against a labour supply constraint, so that growth at the BP-equilibrium rate would become infeasible. In contrast, if the former growth rate is less than the latter, the economy would suffer perpetually increasing unemployment, which seems equally implausible. This is another version of an issue covered earlier in this book (especially in Chapters 1, 6 and 8), which is how to reconcile the growth rate made possible by aggregate demand conditions (in this case, the BP-equilibrium rate) with the growth made possible by aggregate supply conditions (growth of the labour force and labour productivity).[17]

To see this difficulty more formally, consider the following simplified BPCG model with endogenous productivity growth as outlined by Palley (2002c) and Setterfield (2006b), converted into a common notation:

$$x = \eta_X y_f \tag{10.11}$$

$$m = \eta_M y_D \tag{10.12}$$

$$x = m \tag{10.13}$$

$$y_N = q + n \tag{10.14}$$

$$q = q_0 + \rho y_D \tag{10.15}$$

$$y_D = y_N \tag{10.16}$$

Equations (10.11) and (10.12) are the demand functions for exports and imports in growth rate form, omitting the relative price (RER) terms for simplicity (and because they may be considered to be of little importance in the long run), while (10.13) is the BP equilibrium condition in real terms (omitting prices and the exchange rate). As usual, export growth depends on (exogenously given) foreign income growth $y_f$, but import growth depends on the growth rate of home country aggregate demand $y_D$. Writing the import demand function this way implicitly assumes that actual output growth always equals demand growth in the short run (so we can avoid having to introduce an additional variable for the actual growth rate). Equation (10.14) defines the natural rate of growth $y_N$ as the sum of the growth rates of labour supply ($n$) and labour productivity ($q$), similar to previous chapters. Equation (10.15) is Verdoorn's law, which endogenizes labour productivity growth; it is similar to equation (8.21) in Chapter 8 or (9.42) in Chapter 9, but now expressed as a function of $y_D$ instead of $y$. Equation (10.16) states the equilibrium condition that aggregate demand must grow at the natural rate of growth of output in a long-run equilibrium.

The first three of these equations (10.11 to 10.13) solve for the strong form of Thirlwall's law, seen as the solution for the growth rate of aggregate demand that is consistent with BP equilibrium:

$$y_D = (\eta_X/\eta_M)y_f \tag{10.17}$$

The remaining three equations (10.14 to 10.16) can be solved separately to find the (natural but endogenous) growth rate of output that equates $y_N = y_D$ incorporating Verdoorn's law:

$$y_N = \frac{q_0 + n}{1 - \rho} \tag{10.18}$$

Obviously, these two growth rates are different in general; they would only be equal under the special condition that $(q_0 + n)/(1 - \rho) = (\eta_X/\eta_M)y_f$, which (it would appear) could only hold by accident.[18] In mathematical terms, the model described by the six equations (10.11)–(10.16) is overdetermined, because it contains only five endogenous variables ($x$, $m$, $y_D$, $y_N$ and $q$). If (10.17) exceeds (10.18), the economy would be constrained by supply-side limits to grow at the natural rate (10.18). In the opposite case, the economy would grow at the BP-equilibrium rate (10.17), but would face perpetually increasing unemployment.

There are several possible solutions to this dilemma. Porcile and Spinola (2018) note one possibility: if the supply of labour is infinitely elastic, then the growth rate of labour supply $n$ would automatically adjust to equate (10.17) and (10.18). This could occur in a dual economy with 'surplus labour', in the sense of Lewis (1954), where the excess supply of labour in a backward or traditional sector is large enough to ensure that the modern sector has a horizontal labour supply curve at the going wage rate (determined by average productivity of labour in the backward sector, possibly including a premium to cover the costs of relocating to the modern sector). This could also occur in a high-wage country that allows workers to immigrate as needed to keep the effective domestic labour supply equal to the demand. The latter case might correspond to what has occurred in many Persian Gulf countries, where shortages of national labour have been relieved by reliance on immigration, thereby permitting rapid growth unconstrained by domestic labour supply. In either case (surplus labour or free labour mobility), $n$ would have to adjust to the equilibrium rate

$$n^* = -q_0 + (1 - \rho)(\eta_X/\eta_M)y_f \tag{10.19}$$

while output would grow at the BP-equilibrium rate given by the strong form of Thirlwall's law (equation 10.1) in the long run.

However, a perfectly elastic supply of labour is not realistic for most countries. Moreover, a Lewis model makes for a strange bedfellow with a BPCG model in which growth is driven by external demand, since Lewis assumed that – in classical fashion – growth was driven by domestic capital accumulation. Making the opposite assumption of an exogenously fixed rate of labour supply growth ($n = \bar{n}$), Palley (2002c) suggested another solution to the dilemma outlined above: he assumed that the income elasticity of import demand is endogenous and that it is an increasing function of the rate of capacity utilization: $\eta_M = \eta_M(u), \eta_M'(u) > 0$.[19] The rationale for the latter assumption is that when an economy is operating at a high rate of utilization, it will have to import proportionately more goods to relieve tighter (and more pervasive) domestic supply bottlenecks as income grows, while an economy with more excess capacity can expand its output with less need for imports. In this case, the two equilibrium growth rates (10.17) and (10.18) will be equal provided that the utilization rate settles at an equilibrium value $u^*$ such that the special condition cited above holds as an equilibrium condition:

$$\eta_M(u^*) = \frac{1 - \rho}{(q_0 + \bar{n})\eta_X y_f} \tag{10.20}$$

Of course, this solution will be economically meaningful only if $u^*$ lies within a sensible range: the equilibrium utilization rate not only must be positive and below unity, but also must be a rate that firms would be willing to maintain with 'normal' limits as discussed in Chapter 6.[20]

A deeper difficulty with this solution, as Palley admits, is that growth ends up being purely supply-determined in the long run: the demand-side BP-equilibrium solution given by equation (10.17) must adjust (subject to the preceding caveat about the level of $u^*$) to the supply-determined equilibrium solution (10.18), not the other way around. Aggregate demand matters only during the transition from one long-run equilibrium to another; although the BP constraint is satisfied in the long run, it is not really the binding constraint or determinative factor in the long run (so Thirlwall's law is satisfied in a technical sense, but not in spirit).

An alternative solution was suggested by Setterfield (2006b), who argued that the utilization rate is also likely to affect productivity growth. Specifically, he argued for a modified version of equation (10.15), in which the Verdoorn coefficient $\rho$ becomes an endogenous parameter and is an increasing function of the utilization rate:

$$q = q_0 + \rho(u) y_D \qquad (10.15')$$

with $\rho'(u) > 0$. The justification for this assumption is that firms are (for any given growth rate of demand $y_D$) more likely to seek to innovate (or to adopt existing innovations) in order to increase productivity when they are pressing more on capacity constraints, and less likely to do so when there is more slack in their operations.[21] Using the original import demand function (10.12), in which the income elasticity $\eta_M$ is exogenous, and again assuming exogenously given growth of the labour force at the rate $\bar{n}$, the natural rate of growth given by equation (10.18) now adjusts to the BP-equilibrium growth rate (10.17) through endogenous variations in the Verdoorn coefficient $\rho$. This will occur if there is an equilibrium utilization rate $u^{**}$ such that

$$\rho(u^{**}) = 1 - \frac{(\eta_X/\eta_M) y_f}{q_0 + \bar{n}} \qquad (10.21)$$

If such an equilibrium exists (and is feasible in the sense discussed above), growth will be fully demand-determined – and the BP constraint will be binding – in the long run, as envisioned in Thirlwall's law (strong form). It should be noted that this solution is similar in spirit to some of the models considered in Chapter 8, in which the natural rate of growth adjusts to a demand-determined growth rate in the long run; the difference here lies in the nature of the demand-side constraint (exports alone in ELCC models versus BP equilibrium in a BPCG model). Alternatively, if $\eta_M$ is made endogenous à la Palley and $\rho$ is made endogenous à la Setterfield at the same time, then the solutions (10.17) and (10.18) could be reconciled by long-run adjustments on *both* the supply and demand sides of the model (that is, adjustments in the BP-equilibrium *and* natural rates of growth), assuming that the utilization rate would reach the equilibrium level $\bar{u}$ defined by

$[\eta_X/\eta_M(\bar{u})]y_f = (\bar{n} + q_0)/(1 - \rho(\bar{u}))$, and that this level would be feasible as discussed above.

Yet another way of reconciling the BP-equilibrium and natural growth rates is proposed by Oreiro (2016), who introduces the role of RER adjustment into a modified version of the Palley and Setterfield models.[22] In contrast with Palley, Oreiro argues that the income elasticity of import demand should be linked to the RER, rather than to the utilization rate. Based on logic similar to that of Cimoli and Porcile (2014), Oreiro reasons that a higher (depreciated) RER will enable a country to produce a wider range of goods, and that a country with a more diversified domestic productive structure will have less need to import as its economy expands. Linearizing this relationship for mathematical convenience, Oreiro assumes that the income elasticity of import demand in equation (10.12) is given by $\eta_M = b_0 - b_1(EP_f/P)$, where $b_0$ and $b_1$ are positive constants. Substituting this expression into (10.12), the RER becomes the sixth variable needed to close the system of six equations (10.11)–(10.16), in which case the overdetermination disappears and a unique equilibrium exists provided that the RER is endogenous and settles at the following equilibrium level:

$$\left(\frac{EP_f}{P}\right)^* = \left(\frac{1}{b_1}\right)\left[b_0 - \eta_X y_f\left(\frac{1-\rho}{q_0 + \bar{n}}\right)\right] \qquad (10.22)$$

with the obvious restriction that the term in brackets must be positive for the solution to be economically meaningful (since the RER cannot be negative). Furthermore, Oreiro (2016) demonstrates that this solution for the equilibrium RER is stable if one assumes that the RER rises (depreciates) in response to an increasing trade deficit (the interested reader is referred to his article for the details).

A few caveats about Oreiro's solution are in order. First, the long-run equilibrium growth rate in Oreiro's model must be the solution (10.18) above, which is derived solely from the supply side of the model; aggregate demand and BP equilibrium play no role in determining the long-run equilibrium growth rate. More like in Palley's model than Setterfield's, in Oreiro's model the BP-equilibrium growth rate must adjust to the natural rate in the long run, not the other way around.[23] Second, the adjustment process that Oreiro specifies for the RER assumes that this variable adjusts in the right direction to eliminate trade imbalances, which is contrary to some of the empirical findings discussed earlier (especially Alonso and Garcimartín, 1998–99). Oreiro's model also ignores some of the complexities introduced by Ribeiro et al. (2016, 2017b), which as discussed earlier imply that the long-run effects of a currency devaluation on the BP-equilibrium growth rate could go in either direction. Third, Oreiro also solves for the capacity utilization rate that makes the capital stock grow at the same long-run equilibrium rate as output, but the long-run solution of his model is recursive and the utilization–accumulation relationship plays no role in determining the long-run equilibrium growth rate given by (10.18).[24]

In spite of these caveats, Oreiro's model illustrates another way in which the RER could matter to long-run growth: even though growth is not really BP constrained in the long run (since growth occurs at the natural rate, which is determined independently of the BP constraint), a particular level of the RER may still be necessary to reconcile the BP-equilibrium growth rate with the natural rate of growth. Most importantly, Oreiro's model implies an important policy lesson: if a country allows its RER to be overvalued relative to an equilibrium level such as that implied by equation (10.22), the country could be condemned to having a BP constraint imposed in a way that would prevent it from reaching its 'natural' or full-employment growth path.[25] Or, to put it another way, although a country that actually reaches its long-run equilibrium is not BP constrained in Oreiro's model, it is certainly possible that a country could be persistently out of equilibrium (and BP constrained) if its currency is chronically overvalued. Taken together, the models of Palley (2002c), Setterfield (2006b), Oreiro (2016) and Porcile and Spinola (2018) highlight a variety of mechanisms that can potentially reconcile BP-equilibrium growth with the natural (full-employment) growth rate; which of these mechanisms applies depends on the history and structure of a given country.

## 10.5 Alternative models of BP constraints for different size countries

As discussed earlier, the assumption of a Keynesian small economy in the standard BPCG model is controversial, especially for small countries that could be assumed to be price-takers for their exports. Moreover, standard BPCG models treat all imports as (imperfect) substitutes for domestic products, but many developing countries are highly dependent on imports of capital goods (machinery and equipment, which can be generalized to include intellectual property) for which no domestic substitutes are available. This section considers two alternative models that attempt to show how a BP constraint could operate in a small, price-taking country and in a medium-large country that is dependent on imported capital goods.

### 10.5.1 Razmi's small country model

Based on his arguments about country size discussed in section 10.2.4 above, Razmi (2016a) postulates a model of BP-constrained growth for a small country, in the sense of one that takes all foreign prices as given. Based on the market structure shown in Figure 10.1, the equilibrium quantity of exports is now determined on the supply side (while demand for exports is infinitely elastic). The export supply function used by Razmi is

$$X = X^s(P,K) = P^{\gamma_X}K^{\delta_X} \tag{10.23}$$

where $P$ is the home currency price of exports as well as of domestic goods, $K$ is the capital stock, and $\gamma_X$ and $\delta_X$ are the elasticities of export supply with respect to the price and capital stock respectively.[26] For simplicity, a constant term is omitted, the prices of exported goods are assumed to be the same as the prices of domestic products sold at home, and export supply is assumed to respond positively to the nominal export price regardless of the exchange rate or foreign prices.

In growth rate form, the export supply function becomes:

$$x = \gamma_X\hat{P} + \delta_X g \tag{10.24}$$

where $x$ and $\hat{P}$ are the growth rates of export quantities and domestic prices, respectively, and $g = \hat{K}$ is the growth rate of the capital stock as in earlier chapters. For imports, which are still demand-determined, we continue to use equation (9.2) or (9.6) in growth rate form; the latter is reproduced here for convenience:

$$m = -\varepsilon_M(\hat{E} + \hat{P}_f - \hat{P}) + \eta_M y \tag{10.25}$$

The equilibrium condition (assuming balanced trade or no financial capital flows in the long run for simplicity) is the same as equation (9.4) from Chapter 9, which is also reproduced here:

$$\hat{P} + x = \hat{E} + \hat{P}_f + m \tag{10.26}$$

Now, the RER plays an additional role in this model that is not seen in any of the BPCG models considered earlier – with the sole exception of Oreiro (2016).[27] Razmi assumes that investment in domestic industries is a positive function of the RER, so that a real depreciation makes home production more competitive globally and hence leads to greater capital accumulation in the home country.[28] Since the growth rate of the capital stock $g$ equals the ratio of (net) investment to capital, we can express Razmi's assumption on this point as follows:

$$g = g\left(\frac{EP_f}{P}\right), g' > 0 \tag{10.27}$$

As will be seen below, equation (10.27) is not essential to the mathematical solution of the rest of Razmi's model, but it is crucial to his policy conclusions and empirical tests so we will include here.[29]

Then, substituting equations (10.24) and (10.25) for export supply and import demand, respectively, into the BP equilibrium condition (10.26), with all variables expressed in growth rate form, we can solve for the BP-equilibrium growth rate of domestic output

$$y = \frac{\gamma_X \hat{P} + \delta_X g + \varepsilon_M (\hat{E} + \hat{P}_f - \hat{P})}{\eta_M} \tag{10.28}$$

At this point, Razmi utilizes two crucial implications about prices that follow from his assumptions about the elasticities of supply and demand for traded goods. On the one hand, an infinitely elastic supply of imports implies an exogenously given foreign price of imports, which he interprets by assuming that $P_f$ is constant and hence its rate of change is zero ($\hat{P}_f = 0$).[30] On the other hand, infinitely elastic demand for exports implies an exogenously given price of exports in foreign currency, which means that the ratio $P/E$ must be constant and hence the export (domestic) price must change at the same rate as the nominal exchange rate ($\hat{P} = \hat{E}$).[31] Under these assumptions and also incorporating the investment (accumulation) function (10.27), equation (10.28) becomes:

$$y_S = \frac{\gamma_X \hat{E} + \delta_X g(EP_f/P)}{\eta_M} \tag{10.29}$$

where $y_S$ denotes the solution of the small country model.

To draw out the implications of Razmi's solution, it is useful to compare equation (10.29) with the general solution (9.8) of the Thirlwall BPCG model, which is reproduced here:

$$y_B = \frac{(\varepsilon_X + \varepsilon_M - 1)(\hat{E} + \hat{P}_f - \hat{P}) + \eta_X y_f}{\eta_M} \tag{10.30}$$

The only commonality in these solutions is the appearance of the income elasticity of import demand in both denominators, which reflects the fact that both approaches treat imports the same way and both require that the BP constraint must be satisfied. However, the numerators are quite different. The Thirlwall solution (10.30) includes foreign income growth $y_f$ as a determinant of the home country's growth, which Razmi has challenged on empirical grounds as discussed in section 10.2.2 above. Equation (10.30) allows for the possibility that the *rate of change* in the RER (that is, the rate of real depreciation, $\hat{E} + \hat{P}_f - \hat{P}$) could affect the BP-equilibrium growth rate, although there is general agreement among both supporters and critics of the BPCG model that the RER is unlikely to change continuously in the very long run (in which case, equation 10.30 becomes 10.1 or 10.2, either the strong or weak version of Thirlwall's law). In contrast, Razmi's small country solution (10.29) implies that the *level* of the RER matters, not for its demand-side effects, but rather because of its influence on the rate of investment and hence on the country's export capacity.[32] As mentioned earlier, both Razmi (2016a) and other recent empirical studies (cited earlier) find econometric evidence showing that the level rather than the rate of change in the RER (or RULC) has a significant effect on comparative growth rates across countries.

In effect, Razmi argues that the traditional BPCG approach dismisses the importance of the RER too easily. By focusing on the rate of the change in the RER, BPCG advocates can credibly claim that it is unlikely that a country could maintain a continuous rate of real depreciation in the long run. Without disputing this, Razmi counters that it can still be important for a country to maintain a high *level* of its RER (a low real value of its currency) because of the encouragement this gives to investment (whether by domestic or foreign firms) in the country's tradable goods industries. In a sense, Razmi's model brings us full circle back to the focus on the rate of capital accumulation ($g$) in the theories of the classical economists, Marx and the early neo-Keynesians, which was lost sight of (or assumed to adjust automatically, as in the supermultiplier view) in both variants of the open economy Kaldorian approach (the ELCC and BPCG models). However, in Razmi's approach capital accumulation is driven by open economy considerations (international competitiveness as determined by the level of the RER) and the BP constraint remains in full force, even if it is modelled differently from the Thirlwall approach.

However, the relevance of the pure small country case can be questioned for the reasons given earlier: endogenous capital accumulation in export sectors effectively makes export supplies much more elastic in the long run than they are in the short run (with the stipulation that the supply constraints will tend to be overcome more rapidly in countries with more competitive RER levels), and the efforts of many countries to increase exports of similar products simultaneously can run afoul of the fallacy of composition. Moreover, home and foreign goods may not be perfect substitutes, as would be required for the law of one price to hold in the export sector (as is implied for the case of a pure price-taker). Thus, one should be cautious in regard to the applicability of a strict small country model beyond the very smallest countries, countries that export homogeneous types of commodities, or ones that are able to increase their own exports (perhaps of niche products) without other countries following suit. Nevertheless – and in spite of their very different assumptions and logic – Razmi's model is similar to those of Oreiro, Cimoli and Porcile, and various others in implying the importance of not allowing the RER to become overvalued to prevent the BP constraint from binding at a low equilibrium growth rate.

### 10.5.2 Ros's model of terms-of-trade adjustment in a medium-large country

While Razmi seeks to model a BP constraint in a small country, Ros (2013) deliberately sets out to construct a model for what he calls a 'large' country. However, Ros's definition of 'large' is different from the one used in section 9.3.3 in Chapter 9, according to which a country is large enough to have significant repercussion effects on the income of the foreign country (rest of world). Rather, Ros considers a country that is large enough to influence the terms of trade (relative price) for its export products, but not large enough to affect foreign income. To avoid confusion, we will refer to this as the case of a 'medium-large' country.

Another key difference in Ros's approach lies in his treatment of imports. In most BPCG models (except Razmi's on the export side), domestic and foreign goods are assumed to be imperfect substitutes. Ros makes the same assumption for exports, but not for imports. Rather, drawing upon the literature on North–South trade,[33] Ros assumes that imports include capital goods for which no domestic substitutes are available (in the extreme case, all investment requires imports). To account for demand for capital goods, he follows the approach adopted in some of the early neo-Keynesian growth models covered in Chapter 3: there is a uniform marginal propensity to save and all savings are invested (there is no independent investment function). Given these assumptions, Ros's model could plausibly be applied to relatively large emerging or developing nations that depend heavily on imports of capital equipment, and which are big enough to affect world prices of their specialized export products, but not so big as to have repercussion effects on foreign income or global prices of their imports.

Ros assumes what is usually called the '*AK*' specification of production, although we called it the '*aK*' model in Chapter 1 since we had used uppercase *A* for a different variable. As discussed in Chapters 1 and 2, this approach incorporates the proposition that output is proportional to the capital stock in the long run, which is common to both classical-Marxian heterodox growth theory (HGT) models and some species of NEGT models. Implicitly, this specification also assumes that the economy always operates at a normal utilization rate of unity $(u = u_n = 1)$, which implies that output is at the 'full-capacity' level $(Y = Y_K)$ and the capital–output ratio can be written as $a_1 = K/Y_K = K/Y$. Thus, the home and foreign countries' output is determined by their respective capital stocks:[34]

$$Y = K/a_1 \tag{10.31}$$

$$Y_f = K_f/a_1^f \tag{10.32}$$

In this specification, there is full utilization of capital, but not necessarily full employment of labour – which could be considered a realistic assumption for many emerging and developing nations in which there is some degree of surplus labour and capital is the main constraint on production. The quantity of exports is determined by foreign consumer demand (since the home good cannot be used for investment), and is given by the following equation:

$$X = \alpha_0(1 - s_f)\,(P/P_f)^{-\varepsilon_X} Y_f^{\eta_X} \tag{10.33}$$

where $\alpha_0$ is a positive constant, $s_f$ is the foreign marginal propensity to save $(0 < s_f < 1)$, $P$ and $P_f$ are home and foreign prices, respectively (both measured in the same currency), and the other variables and exponents (elasticities) are defined as previously. In this model, $P_f$ is definitely *not* taken as exogenously given; rather, the price ratio or 'terms of trade' $(P/P_f)$ will be endogenously determined as explained below.

On the (rather strong) simplifying assumption that all domestic consumption and investment goods are imported, the value of home imports must equal nominal home income:[35]

$$P_f M = PY \tag{10.34}$$

The BP equilibrium condition (assuming no capital flows) is that the value of exports must equal the value of imports:

$$PX = P_f M \tag{10.35}$$

Substituting (10.33) and (10.34) into (10.35) and using (10.31)–(10.32), we obtain the following solution for the short-run equilibrium terms of trade (relative price of home goods):

$$\frac{P}{P_f} = \left[\frac{\alpha_0(1 - s_f)(K_f/a_1^f)^{\eta_X}}{K/a_1}\right]^{\frac{1}{\varepsilon_X}} \tag{10.36}$$

Thus, equilibrium in the BP is maintained by the adjustment of relative prices or the terms of trade to a unique level, which is an increasing function of the foreign capital stock relative to the home capital stock.

The growth rate (rate of capital accumulation) of the foreign country is determined by its saving propensity divided by its capital–output ratio:[36]

$$g_f = I_f/K_f = s_f/a_1^f \tag{10.37}$$

For the home country, since all capital goods are imported, the feasible rate of accumulation also depends on the terms of trade, which determine the purchasing power of domestic savings over imported capital goods:

$$g = I/K = (s/a_1)(P/P_f) \tag{10.38}$$

These two growth rates (home and foreign) are depicted in panel (a) of Figure 10.2. The $g_f$ line is horizontal because foreign income growth is exogenous to the home country and independent of the terms of trade, while the $g$ line is an upward-sloping ray from the origin because home growth depends positively on the terms of trade per equation (10.38).

The long-run equilibrium does not occur where the two growth lines intersect, however, but rather at the point where the relative price converges to its long-run equilibrium level. To find this point, we take the logarithmic derivative of equation (10.36) with respect to time ($t$) to obtain the rate of change in relative prices. Assuming that the capital–output ratios, foreign saving propensity, and price and income elasticities all remain constant, the relative price changes at the rate

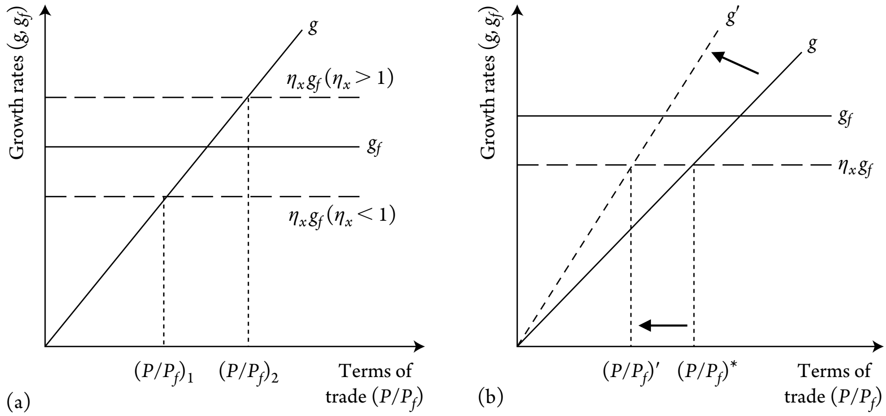



**Figure 10.2** Ros's model of a medium-large country with endogenous terms of trade (relative price): (a) cases of $\eta_x > 1$ and $\eta_x < 1$; (b) a rise in the home country saving rate (assuming $\eta_x < 1$)

$$\frac{d \ln(P/P_f)}{dt} = \frac{1}{\varepsilon_X}(\eta_X g_f - g) \tag{10.39}$$

where we use the fact that the growth rates can be written as $g = d \ln K/dt$ and $g_f = d \ln K_f/dt$. Long-run equilibrium occurs where the derivative (10.39) equals zero, or equivalently $g = \eta_X g_f$. This is seen in Figure 10.2, panel (a), as the point where the upward-sloping $g$ line intersects one of the dashed lines representing $\eta_X g_f$, which lie above or below the horizontal $g_f$ line depending on whether the income elasticity of export demand $(\eta_x)$ is greater or less than unity.[37] Holding all other factors constant, the equilibrium terms of trade $(P/P_f)$ are higher when $\eta_X > 1$ than when $\eta_X < 1$; this is shown as $(P/P_f)_2 > (P/P_f)_1$ in the diagram. It is easily seen that the adjustment to either equilibrium is stable because a rise in the terms of trade $P/P_f$ leads (via equation 10.38) to an increase in $g$, which then reduces the rate of increase in $P/P_f$ per equation (10.39).

A few comments on this equilibrium solution and its significance are in order. First, it can be seen that this solution is consistent with the strong version of Thirlwall's law, in which $y_B = (\eta_X/\eta_M)y_f$. Ros's solution corresponds to the special case in which $\eta_M = 1$, which is implicit in his simplifying assumption that all domestically consumed and invested goods are imported; Ros also uses the growth rate of capital $g$ instead of the growth rate of output $y$.[38] It is all the more striking that Thirlwall's law appears to hold, because growth in Ros's classical model is supply-driven (saving-led) rather than demand-driven (export-led). However, unlike in Thirlwall's model, in

Ros's model the attainment of long-run equilibrium growth requires adjustment of the relative price of exports (the home terms of trade, $P/P_f$) to a unique equilibrium level, which in turn means that the model only works for a medium-large country as defined earlier (since a small country would be a pure price-taker and hence could not affect its own terms of trade).[39] This result is what Ros calls 'Thirlwall's paradox': Thirlwall's law is only valid (when recast in terms of Ros's model) when relative prices adjust endogenously, contrary to the usual assumption of constant relative prices in most BPCG models, and this in turn is only possible if the country is large enough to affect those prices (but not so large as to have repercussion effects on world income).

Second, on the assumption that $\eta_M = 1$, whether the home country grows faster or slower than the foreign country depends solely on whether home exports are income elastic or inelastic. Ros (2013, pp. 241–2) considers only the case in which $\eta_X < 1$, which implies that the home country grows relatively slowly, but this would not seem to apply to all countries. In order to apply the model to countries that grow faster than the rest of the world, it is necessary also to allow for the case in which $\eta_X > 1$. If we consider a less developed home country that starts out with a lower level of per capita income than the (developed) foreign country, the former can be seen as a situation of divergence, while the latter would correspond to a situation of convergence.

Third, Ros's model shows explicitly what happens to a BP-constrained, medium-large country that attempts to raise its growth rate by increasing the propensity to save and invest at home. As shown in panel (b) of Figure 10.2, a rise in the home saving rate $s$ would rotate the $g$ line to become the steeper line $g'$ (for illustrative purposes, we have drawn the diagram for the case where $\eta_X < 1$, but the same logic would apply in the opposite case). In the short run, when the terms of trade are given, the home growth rate increases according to equation (10.38). However, in the long run the terms of trade adjust as shown in equation (10.39) until the home country growth rate returns to $g = \eta_X g_f$, at which point the equilibrium terms of trade have fallen from $(P/P_f)^*$ to $(P/P_f)'$ in Figure 10.2, panel (b). Essentially, temporarily faster home growth raises demand for imported capital goods, thus pushing up their relative price and lowering the terms of trade for home exports, which in turn makes capital goods more expensive and thereby slows accumulation (growth) until it returns to its long-run equilibrium rate. However, the home country still gains from the additional capital accumulated as a result of the faster growth achieved during the transition period, before the new long-run equilibrium is reached.

Ros's model thus implies that Thirlwall's law obtains only for countries that are large enough to affect their terms of trade, and in fact requires that those terms of trade adjust endogenously. This conclusion seems to undermine the standard BPCG analysis, in which relative prices are irrelevant to the long-run equilibrium growth rate. However, Ros's model in turn rests on certain very strong assumptions. Ros's model only works for a medium-large country that is able to affect its terms of trade, but is not big enough to affect foreign income levels. Moreover, the terms of trade that have to adjust in Ros's model are conceptually different from the RER that is featured in more standard BPCG models. Ros's terms of trade is the relative price of home exports of consumption goods compared with imports of capital goods, while the usual RER is the relative price of home goods compared with foreign products that are imperfect substitutes for domestic products. Finally, even if some of his more extreme simplifying assumptions are relaxed, Ros's model only applies in situations in which a country's terms of trade strongly affect the real value of domestic saving in terms of capital goods, which is most likely to occur when a country is heavily dependent on imported capital goods for domestic investment (even if not all capital goods are imported). Otherwise, a country's growth rate would not be so closely tied to its terms of trade as it appears in equation (10.38).

Nevertheless, Ros's model provides an alternative way of thinking about the way a BP constraint operates in such a medium-large country, just as Razmi's model provides an alternative framework for analysing such a constraint in a small, price-taking economy. Ros's model is especially reflective of the Latin American tradition of structuralist economics, in which the terms of trade play a key role in addition to income elasticities.[40] Contrary to the standard versions of Thirlwall's law, what Ros calls 'Thirlwall's paradox' applies: the law only holds for countries of sufficiently large size that one of the assumptions of the standard BPCG model – constant relative prices – cannot be maintained (rather, the terms of trade must adjust endogenously). Considered together, the Razmi and Ros models show the need to take a country's structural features into account in analysing BP-constrained growth. They reveal different ways in which relative price levels can matter to long-run growth: by encouraging more investment that enhances export capacity in small countries, and by adjusting to permit growth at the BP-equilibrium rate in medium-large countries.

## 10.6 A grand synthesis? Reconciling ELCC and BPCG

When we discussed incorporating Verdoorn's law into a Thirlwall-type BPCG model in Chapter 9 (section 9.4.2), we noted that the positive feedback effects of output growth on productivity growth would only matter in a time frame in which relative prices (the RER) could continuously change, so that a country could gain an increasing competitive advantage. The result was a hybrid ELCC–BPCG model, in which cumulative causation could positively affect growth within the limits set by a BP constraint but only for as long as relative prices could be plausibly assumed to continuously change. This leaves open the question, how does an economy transition from what we might call a medium-run period in which Kaldor–Verdoorn effects can create virtuous (or vicious) circles of cumulative causation to a long-run equilibrium in which such effects are absent? Or, to put it another way, how do relative prices eventually adjust so as to offset the positive feedbacks and blunt their impact in the long run?

Some previous literature has suggested possible adjustment mechanisms for the Thirlwall BPCG model, mostly focusing on adjustments of domestic wages and prices (for example, in response to a currency depreciation that boosts exports). Pugno (1998) used a model of labour supply and demand, in which a currency depreciation that leads to more rapid growth of exports and output causes labour demand to increase relative to (exogenously given) labour supply. The resulting upward pressure on wages and prices eventually fully offsets the initial competitive gains from the depreciation, so that Thirlwall's law (in which a depreciation has no effect) holds in the long run.[41] Porcile and Lima (2010) extended Pugno's analysis by considering a dual economy model in which the labour supply to the modern (internationally competitive) sector is endogenous and responds to the wage gap between that sector and the subsistence sector. In their model, the upward pressure on wages from rapid growth of exports and employment in the modern sector can be relieved by induced migration of labour from the subsistence sector; the degree to which this occurs depends crucially on the elasticity of intersectoral labour mobility in response to increases in the modern-subsistence wage gap.

Blecker (2013b) discussed (but did not formally model) how an economy could transition from an ELCC equilibrium in the medium run to a BPCG one in the long run. He hypothesized that adjustments in either nominal wages or exchange rates could offset the dynamic productivity gains from cumulative causation and move the economy towards the Thirlwall's law solution in the long run. Specifically, he argued that 'an export-led boom' in an economy characterized by Kaldor–Verdoorn effects 'could be expected to lead to either faster nominal wage growth . . . or currency appreciation' (higher $\hat{W}$ or lower $\hat{E}$ in our notation), which eventually could completely offset the productivity gains due to cumulative causation (Blecker, 2013b,

p. 405). Blecker's suggestion about wages has been pursued more formally by Ribeiro et al. (2017a), who specify a wage–price adjustment process for a model that combines ELCC and BPCG features by adapting the neo-Kaleckian model of markup pricing in an open economy (similar to what we covered in section 4.4.3 of Chapter 4). The model of Ribeiro et al. is, in fact, quite an ingenious and masterful synthesis of these various approaches, and also provides an important example of incorporating imports of intermediate goods into the analysis in a non-trivial way, so we will focus on their analysis here (with their notation modified so as to match the rest of this chapter and the previous one).[42]

Ribeiro et al. (2017a) assume a country with two kinds of imports: consumption goods and intermediate goods (capital goods are not explicitly modelled). Intermediate imports ($M_i$) are specified differently from how we treated them in Chapter 9 (section 9.3.2). Ribeiro et al. model them simply as a constant proportion $\delta$ ($0 < \delta < 1$) of output $Y$:

$$M_i = \delta Y \tag{10.40}$$

which implies a unitary income elasticity. Imports of consumption goods ($M_c$) are determined by a standard constant-elasticity demand function:

$$M_c = M_{c,0}\left(\frac{EP_f}{P}\right)^{-\varepsilon_c} Y^{\eta_c} \tag{10.41}$$

where, as previously, $E$ is the nominal exchange rate (domestic/foreign currency), $P_f$ is the foreign currency price of imported consumption goods and $P$ is the domestic price level. Converting these two equations to rate-of-change form, and assuming that $\delta$, $M_{c,0}$ and the elasticities $\varepsilon_c$, $\eta_c > 0$ are all constants, the two types of imports grow at the rates

$$m_i = y \tag{10.42}$$

$$m_c = -\varepsilon_c(\hat{E} + \hat{P}_f - \hat{P}) + \eta_c y \tag{10.43}$$

Exports (which consist entirely of final consumption goods) are determined by the same type of export demand function we used for the basic Thirlwall model in Chapter 9, which is equation (9.1) in levels or (9.5) in growth rate form; the latter is reproduced here for convenience:

$$x = \varepsilon_X(\hat{E} + \hat{P}_f - \hat{P}) + \eta_X y_f \tag{10.44}$$

where exports are assumed to be sold at the same prices as domestic goods.

Assuming no capital flows, BP equilibrium requires balanced trade or zero net exports, which can be written as

$$PX = E \cdot (P_f M_c + P_i M_i) \tag{10.45}$$

where $P_f$ is the foreign currency price of imported consumption goods and $P_i$ is the foreign currency price of imported intermediates. Converted into growth rate form, the equilibrium condition becomes

$$\hat{P} + x = \hat{E} + \hat{P}_f + (1 - \beta_i) m_c + \beta_i m_i \tag{10.46}$$

where (as in Chapter 9) $\beta_i$ is the share of intermediates in total imports and we assume for simplicity that the prices of all imports (consumption and intermediate goods) increase at the same rate (so $\hat{P}_i = \hat{P}_f$). Substituting (10.42)–(10.44) into (10.46) and rearranging, we obtain a general expression for BP equilibrium in this model:

$$[\beta_i + (1 - \beta_i)\eta_c]y = \eta_X y_f + [\varepsilon_X + (1 - \beta_i)\varepsilon_c - 1](\hat{E} + \hat{P}_f - \hat{P}) \tag{10.47}$$

where $\beta_i + (1 - \beta_i)\eta_c$ is the weighted average income elasticity of imports (recalling that the income elasticity for intermediate imports is unity per equation 10.40). Also note that the price elasticity of imports of consumption goods is weighted by the share of those goods in total imports, since imports of intermediate goods are price-inelastic; thus, the modified ML condition for this model is $\varepsilon_x + (1 - \beta_i)\varepsilon_c > 1$ (which is assumed to hold, so Ribeiro et al. do not rely on elasticity pessimism for any of their results).

Now the crucial innovation of Ribeiro et al. is the way they model wages, prices and markups. The domestic price is set by a markup ($\tau > 0$) on unit costs, including the costs of imported intermediate goods as well as nominal unit labour costs (NULC):

$$P = (1 + \tau)(Wa_0 + EP_f\delta) \tag{10.48}$$

where it is assumed for simplicity that all intermediate goods are imported. Converted to rate-of-change form, and still holding the imported intermediate coefficient $\delta$ constant, domestic prices increase at the rate

$$p = \tau' + \varphi(\hat{W} - q) + (1 - \varphi)(\hat{P}_f + \hat{E}) \tag{10.49}$$

where $\tau'$ is the proportional rate of increase in the markup factor or price–cost margin $(1 + \tau)$, $\hat{W}$ is the rate of growth of the nominal wage, $q = -\hat{a}_0$ is the growth rate of labour productivity and $\varphi = (Wa_0)/(Wa_0 + EP_f\delta)$ is the share of NULC in total unit costs (average variable costs). Ribeiro et al. (2017a) assume that when the bargaining power of labour strengthens, workers are able to obtain wage increases (relative to their productivity and to the costs of intermediate imports) that raise $\varphi$.

The markup is assumed to be an increasing function of the RER, as in Chapter 4, on the grounds that when home country products are more price-competitive internationally, firms are able to raise markups (while a reduction in competitiveness squeezes markups). Ribeiro et al. adopt a simplified version of the markup function we used in Chapter 4, where the markup factor is simply proportional to the RER by the constant factor $\mu$ (which must be positive *and* large enough so that $\tau > 0$):[43]

$$1 + \tau = \mu\left(\frac{EP_f}{P}\right) \tag{10.50}$$

Using equations (10.48) and (10.49) and the definition of $\varphi$, equation (10.50) can be expressed in proportional rate-of-change form as[44]

$$\tau' = -(\varphi/2)\,[(\hat{W} - q) - (\hat{P}_f + \hat{E})] \tag{10.51}$$

To model how wages and productivity adjust, we need to distinguish three different time periods. Ribeiro et al. (2017a) call these three periods the short run, medium-to-long run and long run (with the second of these emphasizing its transitional nature), but we will simplify by calling them short, medium and long.

*Short run*: Both the nominal wage and labour productivity grow at exogenously fixed rates, $\hat{W} = \overline{w}$ and $q = \overline{q}$. Using these two given rates along with equations (10.49) and (10.51), the short-run domestic inflation rate is given by

$$\hat{P} = (\varphi/2)\,(\overline{w} - \overline{q}) + (1 - \varphi/2)\,(\hat{P}_f + \hat{E}) \tag{10.52}$$

Then, substituting this equation for $\hat{P}$ into the equilibrium condition (10.47) and rearranging, we obtain the short-run solution for the BP-equilibrium growth rate as

$$y_{B\text{-}SR} = \frac{\eta_X y_f + [\varepsilon_X + (1 - \beta_i)\varepsilon_c - 1]\,(\varphi/2)\,(\hat{E} + \hat{P}_f - \overline{w} + \overline{q})}{\beta_i + (1 - \beta_i)\eta_c} \tag{10.53}$$

There is no cumulative causation (Verdoorn's law does not apply) in the short run because productivity is assumed to grow at a fixed rate. We will consider the short-run impact of a currency depreciation (a rise in $\hat{E}$) below.

*Medium run*: Wages adjust over time in response to domestic inflation, the growth of labour productivity and the gap between the workers' expected share of labour in total unit costs $\varphi^e$ and their actual share $\varphi$ (taken as a proxy for the gap between the expected and actual labour shares of income):

$$\hat{W} = \hat{P} + \alpha q \tag{10.54}$$

where $\alpha = 1 - (1/\kappa)(\varphi^e - \varphi)$ and $\kappa > 0$ is a constant adjustment parameter. Note that this is a form of the productivity bargaining discussed in section 5.2.2 in Chapter 5, but specified differently. Assuming that workers do not have absolute bargaining power, $\alpha < 1$ in the medium run. The assumption here is that workers implicitly bargain for a target or 'expected' labour cost share $\varphi^e$, as a proxy for how they target the wage share of national income $\psi$ (as discussed in Chapter 5).[45] Furthermore, productivity now increases endogenously according to Verdoorn's law (equation 8.21 or 9.42), which we rewrite here for convenience:

$$q = q_0 + \rho y \tag{10.55}$$

Now, using equations (10.54) and (10.51) in (10.49), the domestic inflation rate in the medium run is[46]

$$\hat{P} = \frac{(\varphi/2)(\alpha - 1)q}{1 - \varphi/2} + (\hat{P}_f + \hat{E}) \tag{10.56}$$

where the lower is $\alpha$, the more that productivity growth holds domestic inflation below the rate of increase in foreign prices measured in domestic currency, so that the RER depreciates and domestic competitiveness improves.

Then, substituting equation (10.55) into (10.56) and using the result in the equilibrium condition (10.47), the medium-run BP-equilibrium growth rate is

$$y_{B\text{-}MR} = \frac{(1 - \varphi/2)\eta_X y_f - [\varepsilon_X + (1 - \beta_i)\varepsilon_c - 1](\varphi/2)(\alpha - 1)q_0}{(1 - \varphi/2)[\beta_i + (1 - \beta_i)\eta_c] + [\varepsilon_X + (1 - \beta)\varepsilon_c - 1](\varphi/2)(\alpha - 1)\rho} \tag{10.57}$$

where a stronger Verdoorn feedback effect (a higher $\rho$) increases the medium-run BP-equilibrium growth rate provided that the modified ML condition holds and $\alpha < 1$ (so that nominal wages don't fully adjust to offset the resulting productivity gains). In this time frame, technology policies that could either boost autonomous technical progress (raise $q_0$) or increase the country's ability to absorb new technologies in response to faster growth (increase $\rho$) would have a payoff in enabling a country to grow faster without incurring BP deficits. Interestingly, the changes in the exchange rate and foreign prices drop out of the solution (10.57), because of the endogenous responses of domestic wages and prices, and hence the former have no direct impact on the medium-run equilibrium $y_{B\text{-}MR}$ (although they could have an indirect impact by altering the weighting factor $\varphi$, as discussed below for the short-run model).

*Long run*: The long run is defined as the period in which wages fully adjust so that the share of NULC in total unit costs converges to the workers' expected level: $\varphi \rightarrow \varphi^e$. In this case, $\alpha \rightarrow 1$, and the rate of domestic price inflation (10.56) becomes

$$\hat{P} = \hat{P}_f + \hat{E} \qquad (10.58)$$

signifying a constant RER. When $\alpha = 1$, equation (10.54) implies that the real wage increases at exactly the rate of productivity growth ($\hat{W} - \hat{P} = q$) in the long run, and the medium-run BP-equilibrium growth rate (10.57) transforms into the long-run rate

$$y_{B\text{-}LR} = \frac{\eta_X y_f}{\beta_i + (1 - \beta_i)\eta_c} \qquad (10.59)$$

which is Thirlwall's law for the Ribeiro et al. (2017a) model with intermediate goods. Thus, Verdoorn cumulative causation effects do not operate in the very long run, but any absolute gains in labour productivity and per capita income obtained during the medium-run period of adjustment would be kept – and the longer such a period could be prolonged (for example, via wage suppression, capital inflows or currency undervaluation), the more such long-term gains would ultimately accrue (but possibly at the cost of greater inequality along the way).

Ribeiro et al. (2016, 2017a, 2017b) emphasize, however, that the effects of currency depreciation are ambiguous in their model in all time frames. To show some of the ambiguities briefly, we will confine our discussion here to the short-run impact. The direct impact of a faster rate of depreciation (rise in $\hat{E}$) on the short-run BP-equilibrium growth rate $y_{B\text{-}SR}$ is positive, assuming that the modified ML condition $\varepsilon_X + (1 - \beta_i)\varepsilon_c > 1$ holds in equation (10.53). However, a faster depreciation also raises the cost of imported intermediate goods, which has an offsetting effect in raising domestic price inflation per equation (10.52). In addition, a rise in $\hat{E}$ reduces the share of NULC in total costs ($\varphi$), which is a weighting factor on the relative price term in the numerator of (10.53), and this can have an offsetting (indirect) negative effect on $y_{B\text{-}SR}$. Ribeiro et al. show that the total effect of a depreciation in the short run is given by

$$\frac{\partial y_{B\text{-}SR}}{\partial \hat{E}} = \frac{[\varepsilon_X + (1 - \beta_i)\varepsilon_c - 1][(\hat{E} + \hat{P}_f - \overline{w} + \overline{q})(\partial\varphi/\partial\hat{E}) - \varphi]}{2[\beta_i + (1 - \beta_i)\eta_c]} \tag{10.60}$$

where $\partial\varphi/\partial\hat{E} < 0$ by the definition of $\varphi$. From the derivative (10.60), it can be seen that a depreciation (rise in $\hat{E}$) will increase BP-equilibrium growth (in the short run) if and only if $(\hat{E} + \hat{P}_f - \overline{w} + \overline{q})(\partial\varphi/\partial\hat{E}) > \varphi$.

Further complications arise in modelling the medium- to long-term impact of a depreciation, for which the authors perform dynamic simulations that yield different outcomes under alternative parameter assumptions. Moreover, the effects of currency devaluation or undervaluation also become more complex (and uncertain) if one incorporates the impact on income distribution (the wage or profit share) and the effects of the latter on spending and saving (Ribeiro et al., 2017b)[47] or endogenous effects on technology and productivity (see Ribeiro et al., 2016). Although space constraints preclude us from covering these additional complexities in detail here, the authors' key point is that currency devaluation does not provide a panacea for sparking export-led growth or relieving BP constraints, once all indirect effects and feedbacks are taken into account, and the benefits of such a policy depend on various parameters and initial conditions.

In sum, the model of Ribeiro et al. (2017a) presents a neat analytical portrayal of the factors that affect the BP-equilibrium growth rate over different time horizons, defined by the exogeneity or endogeneity of productivity and the degree to which wages adjust. Relative prices or the RER may (subject to some qualifications) have their standard short-run effects (a real depreciation increasing BP-equilibrium growth) only in the short run, when productivity growth and wage increases are exogenously given by past conditions. In the medium run, when productivity is endogenous according to Verdoorn's law and wages only partly adjust, a country can achieve virtuous circles (or suffer vicious circles) of cumulative causation as envisioned in the ELCC model; in this framework, the endogenous responses of productivity affect the BP-equilibrium growth rate. In the long run, even though productivity growth is still endogenous, the cumulative causation mechanism ceases to function because (by assumption) wages fully adjust to productivity changes, thereby negating any cost advantages (or disadvantages) that arose prior to that adjustment (although Verdoorn effects can still affect the long-run equilibrium if they influence the income elasticities, as in the models of Ribeiro et al., 2016, Porcile and Spinola, 2018 and others discussed earlier).

If there is a policy message in this, it is perhaps that countries that are able to prolong the medium-run period can gain persistent competitive advantages and gains from cumulative causation. This may or may not be achievable through wage suppression or currency undervaluation that can delay the catch-up of real wages to productivity growth, depending on various parameters and feedbacks as discussed above, but even if it is, it would be achieved at the cost of worsened inequality (since the wage share would be depressed in such a process). Alternatively, the judicious use of industrial policies and efforts to augment innovative capabilities could also help to achieve competitive advantages during the medium-run time horizon by raising $q_0$ and/or $\rho$. It is also important to remember that the results of Ribeiro et al. (2017a) pertain only to the BP-equilibrium growth rate; they do not consider, for example, the possibility that rising inequality in the medium-run phase could create domestic demand-side stagnation, in which case actual output growth could fall below the BP-equilibrium rate. To avoid such a fate, a country would have to have sufficiently rapid growth of its exports to offset the depressed state of the internal market – but this would raise questions about the viability of such a strategy given the possibility of foreign markets becoming more closed (for example, through the imposition of tariffs or adoption of buy-domestic policies) or foreign countries engaging in counteracting competitive strategies (for example, competitive devaluations or 'race-to-the-bottom' wage cuts), assuming that any of these would be effective.

## 10.7 Conclusions

Although this chapter has covered what may seem like an overwhelming number of modifications of (or alternatives to) the standard BPCG model, in fact it has only covered the proverbial tip of an iceberg. The theoretical literature advancing this approach (or developing alternatives to it) has mushroomed in recent years, making it impossible to cover all of the new versions and debates in depth in a single chapter (and that is on top of the more well-established extensions already covered in Chapter 9). What we have tried to do here is to give the reader a 'flavour' of the new literature related to debates over Thirlwall's law and a selection of some of the main types of new modelling approaches. In these conclusions to the final chapter of this book, we will discuss a few of the most salient issues in contemporary debates about the BPCG approach and try to link them to broader issues about future directions for research in heterodox macro theory.

Perhaps the cutting-edge issue in BPCG modelling today is the role of relative prices (terms of trade, RER or RULC, depending on the model specification). In the standard formulation, relative prices enter the model in rate-of-change form, and since it seems safe to assume that relative prices don't continuously change in the same direction over long periods of time, relative price (change) effects can be ignored in a long-run analysis. This is the basis for the simpler (strong and weak) versions of Thirlwall's law, and for the policy implication that what matters to long-run growth is only non-price (qualitative) competitiveness (as reflected in the income elasticities of exports and imports), rather than cost competitiveness (as reflected in relative prices). However, the abundance of empirical evidence that levels of the RER or RULC do matter to long-run growth, even if these variables don't exhibit continuous changes over very long periods, has led to a variety of new modelling efforts aimed at incorporating levels of relative prices or RERs into a broader BPCG framework.

Several strands of this new work were covered in this chapter, of which we may distinguish at least three varieties. First, some models (for example, Cimoli and Porcile, 2014; Oreiro, 2016) allow the level of the RER to affect the range of goods produced in a country, which in turn influences the income elasticities of demand for traded goods: a greater range of goods produced at home tends to raise the income elasticity of exports and/or lower the income elasticity of imports. Second, the level of the RER is a crucial parameter for influencing capital accumulation and export supply, and thereby affects the BP-equilibrium growth rate in Razmi's (2016a) model of a small, price-taking economy. And third, several models imply that the terms of trade or RER has to adjust in order to reach a long-run equilibrium consistent with the BP constraint. This occurs, for example, in Oreiro's (2016) model of how the BP-equilibrium growth rate adjusts towards the natural rate of growth, Ros's (2013) model of how the growth of a medium-large country that depends on imports of capital goods adjusts to the BP-equilibrium rate, and the model of Ribeiro et al. (2017a) in which real wage increases eventually catch up with productivity growth so as to offset cumulative causation (positive feedbacks) in the long run.

Also, as noted in Chapter 9, Thirlwall's law is a specific type of supermultiplier model, and as a result, any version of it (however modified or qualified) depends on the validity of the supermultiplier logic for a given country in a given historical period. In Thirlwall's model, as in any supermultiplier approach, it must be assumed that investment and capital accumulation adjust endogenously to the growth trajectory predicated on the growth of some type of autonomous expenditure or external factor. For example, in the basic (weak) version of Thirlwall's law, investment and other domestic expenditures must ultimately (in the long run) adjust to the growth rate of output made possible by the ratio of the growth rate of exports to the income elasticity of import demand. Moreover, the BPCG approach is a very specific application of the supermultiplier principle, in which the external (international) constraint is *always* the binding one in the long run. Yet, surprisingly little effort has been devoted to explicitly modelling the *income* adjustment mechanisms that must occur (as opposed to, or in addition to, the relative price, RER or wage adjustments, which have been modelled more extensively) in order for growth to converge to the BP-equilibrium rate in the long run.

Furthermore, significant doubts about the general validity of the BPCG approach remain in many corners of heterodox economics. Is it really always true that the external constraint is the binding one in the long run in most countries, or only in certain types (for example, highly open countries that tend to run trade deficits and have net capital inflows)? Or only in countries of a certain size (not too big, not too small)? Even if the external constraint is binding, is it necessarily true that relative prices can only affect the long-run equilibrium growth rate if they operate via the (weighted-average) income elasticities of export and import demand, or could they also affect long-run growth through other channels? As one sceptic of the BPCG approach has expressed, there is a risk in some of the newer models that 'the income elasticities are becoming a container of all that is important for export and import growth' (and, we might add, for growth in general).[48] And, is it really true that domestic investment and capital accumulation are merely dragged along in response to the requisites of maintaining balanced trade (or sustainable net capital flows) in the long run, or can they be independent driving forces in long-run growth dynamics (which in turn would condition a country's export performance and import propensities)? These are the sorts of questions that current and future generations of heterodox macroeconomists will need to address in order to find the best path forward and to develop the most useful kinds of models for countries with various structural characteristics.

## ? STUDY QUESTIONS

1) What have been the main criticisms of the BPCG model and empirical tests of Thirlwall's law? How would defenders of this approach respond to those criticisms? Discuss and evaluate.
2) How does the analysis of BP-constrained growth differ in the case of a pure, price-taking small country? How does the analysis differ for a medium-large country, defined as one that is big enough to influence world prices of its export products, but not big enough to affect foreign income? Answer using appropriate models, carefully stating their assumptions and deriving their results.
3) Why is it a problem that the BP-equilibrium growth rate may not coincide with the 'natural rate of growth', and what mechanisms have been proposed to reconcile these two rates? What are the different implications of the alternative reconciliation mechanisms in regard to both theory and policy?
4) Why can the BPCG model be considered a type of 'supermultiplier' model, and how does it compare with the other supermultiplier models covered in Chapters 7 and 8? Discuss the similarities and differences.
5) Review the debate over the importance (or lack of importance) of relative price or RER effects in BPCG models. Does it matter if relative prices or the RER are considered in rates of change or levels? What are the parallels between this debate and the debate over 'Kaldor's paradox' as discussed in Chapter 8?
6) In what sense can the BPCG model be regarded as applying to the long run, while the ELCC model of Chapter 8 can be regarded as applying only in the 'medium run'? What kind of transition would need to occur for an economy to move from a medium-run growth path characterized by cumulative causation to a long-run equilibrium in which Thirlwall's law applies?
7) For any of the debates covered in this chapter (for example, about the role of relative prices or the causality between output growth and capital accumulation), propose an econometric strategy for testing the alternative hypotheses empirically.

## NOTES

1 See, for example, Rodrik (2008), Berg et al. (2012), Rapetti et al. (2012) and Boggio and Barbieri (2017). For contrary evidence, see Ribeiro et al. (2018), who find that the direct effects of the RER on long-run growth become insignificant once income distribution and technological capabilities are controlled for, while negative indirect effects remain.

2 Portions of this section draw heavily on Blecker (2016a). See also McCombie (2011) for much parallel discussion.

3 Razmi (2011) notes that conventional BPCG models have a tendency to overpredict long-run average growth rates, which he attributes to their failure to take non-traded goods into account. Razmi (2011) also shows that the standard BPCG results based (implicitly) on assuming a single domestically produced good do not easily generalize to a three-good framework including exportables, importables and non-tradables.

4 Authors' calculations based on data from US Bureau of Economic Analysis (2016). The years used here were chosen because the quantity indexes for real exports and imports in this source only go back to 1967, and hence the first year for which growth rates can be computed is 1968.

5 The view that the US and China are not BP constrained is supported by Razmi's (2016a) finding that the foreign (world) growth rate is insignificant in regressions for US and Chinese income growth, after controlling for the rate of domestic capital accumulation. Alonso and Garcimartín (1998–99) found that US national income does not adjust significantly in response to trade imbalances. They found that Japanese income does adjust, but earlier Thirlwall (1979) had found Japan to be an exceptional case in which growth was limited by capacity constraints at least up to the 1970s.

6 However, if supply curves for exports or imports are not infinitely price-elastic (horizontal), as discussed in section 10.2.4 below for the case of small countries, then conventional estimates of price elasticities

(which take prices as exogenous) are subject to simultaneity bias and are likely to be biased downward (in absolute value). Razmi (2005, pp. 681–2) reports finding a much larger (greater than unity, in absolute value) price elasticity for Indian imports using methods that control for simultaneity, compared with earlier studies that did not use such methods.

7 All translations from Spanish sources were done by the authors.

8 For the case of Mexico, Ibarra and Blecker (2016) found that after a prolonged period (1960–74) in which the economy grew persistently faster than its BP-equilibrium rate, the country subsequently experienced adjustments in both its national income (which grew more slowly over the next two decades) and the RER (which tended to depreciate), with both kinds of adjustment occurring not smoothly but through a series of BP and currency crises between the mid-1970s and mid-1990s. By 2001–12, Mexico's (very low) average annual growth rate of 2.0 per cent was very close to the authors' estimates of its BP-equilibrium rate for that period, indicating that the country's income had adjusted in the very long run.

9 As noted earlier (see section 9.3.2 in Chapter 9), the multisectoral BPCG model is an exception, because it emphasizes how the weighted average income elasticities at the aggregate level are likely to evolve over time as a result of structural change at the industry level (shifting shares of different goods). Other models that make the income elasticities endogenous, and hence subject to change, are discussed later in this chapter.

10 Since Razmi's data set (taken from Penn World Tables version 8.0) covers the years 1950–2011, the final period (2005–11) consists of seven years.

11 Although few BPCG theorists use this terminology, the condition that $\hat{E} + \hat{P}_f - \hat{P} \approx 0$ is the same as what is conventionally called relative 'purchasing power parity' (PPP, understood here as a long-run proposition). Razmi (2016a) claims that assuming relative PPP is inconsistent with Thirlwall's law, because the BPCG model requires that $\hat{P}_f = \hat{P} = 0$ (since prices of exported and imported goods are fixed in the seller's currency). Hence, in his view, any non-zero rate of nominal depreciation $\hat{E} \neq 0$ would cause the RER to change and hence would affect $y_B$ via equation (9.8). However, in the long run exogeneity of prices of imported and exported goods (in the seller's currency) requires only that $\hat{P}_f$ and $\hat{P}$ are exogenously given (the home and foreign inflation rates must be independent of the volumes traded), not that they are zero (that is, the price levels don't have to be constant), so relative PPP can still possibly hold as long as $\hat{E} = \hat{P} - \hat{P}_f$ in the long run.

12 However, the estimated price elasticities in Alonso and Garcimartín (1998–99) may be biased downward for econometric reasons. These authors used a dynamic specification of gradual convergence to equilibrium levels, which effectively imposes equal time lags on the quantity and price variables. Most empirical studies that don't impose this restriction find that lags are considerably longer for relative price effects than for income effects, and this could account for why Alonso and Garcimartín found fairly low price elasticities for most countries. Later, Garcimartín et al. (2010–11) found higher price elasticities that satisfy the ML condition for Spain and Portugal. For varying evidence on the US case, see Lawrence (1990), Blecker (1992) and Chinn (2004).

13 Of course, at a global level relative prices (especially the terms of trade for primary commodities) may depend on world demand conditions, but for a small country these effects are transmitted through those prices.

14 Krugman (1989, p. 1041; emphasis in original) notes that 'productivity gains can be represented as increases in the *effective* labour forces'.

15 For other efforts to incorporate technological gaps in BPCG models and to link the income elasticities to the RER, see Gabriel et al. (2016), Oreiro (2016), Marconi et al. (2016), Ribeiro et al. (2016) and Porcile and Spinola (2018). Oreiro's model will be covered in the next section, while the work of Ribeiro et al. will be covered in more depth later in this chapter. In a related but different vein, Santana and Oreiro (2018) incorporate structural change into a BPCG model with Kaldor–Verdoorn effects by making the Verdoorn coefficient (our $\rho$) a function of the share of manufacturing in GDP.

16 This outcome is not a mathematical necessity, because as some goods are shifted from North to South, the *average* income elasticities for the goods produced in *both* regions rise: while the South acquires some new goods that have higher income elasticities than the goods already produced there, under Cimoli and Porcile's assumptions it is the goods with relatively low income elasticities in the North (*ex ante*) that move, so the average income elasticity for the *remaining* goods in the North also ends up higher. Hence, Cimoli and Porcile implicitly assume that the effect on the South is stronger than the effect on the North.

**17** For an extensive review and analysis of alternative ways of reconciling the BP-equilibrium and natural rates of growth with a focus on the perspective of developing countries, see Porcile and Spinola (2018).

**18** Note the similarity between this special condition and equation (8.37) in Chapter 8, where the left-hand sides $[(q_0 + n)/(1 - \rho)]$ are the same but the right-hand sides differ. Both conditions relate to the difficulty in ensuring that growth occurs at the natural rate in the long run, except that Chapter 8 discussed this dilemma in relation to the ELCC model while the present chapter does so in relation to BPCG.

**19** Palley (2002c) called his variable '*E*' for 'excess capacity' and asserted that it would have a negative effect on $\eta_M$, which seems logical – but then he defined *E* as 'the ratio of current output to normal capacity output', which is the rate of capacity *utilization* – *not* the rate of *excess* capacity. Setterfield (2006b) cleared this up by defining *E* as the utilization rate and assuming that it has a positive effect on $\eta_M$; we do the same here, except we use *u* for the utilization rate.

**20** Oreiro (2016) claims that Palley's solution is still overdetermined, but this seems to be based on including an additional variable (the growth rate of full-capacity output as determined by capital accumulation) that was not included in Palley's original model. See section 8.5 in Chapter 8 as well as Palley (2009), Setterfield (2011) and Porcile and Spinola (2018) for further discussion of Palley's approach and alternatives.

**21** For further discussion of Setterfield (2006b) see section 8.5 in Chapter 8. Sceptics might question why higher utilization does not induce the adoption of more capital-saving as well as more labour-saving innovations, and whether the labour-saving variety in particular might not be more responsive to tighter labour market conditions (rising real wages or a higher wage share) rather than a high utilization rate. Both the rate of employment and the rate of capacity utilization are responsive to demand conditions, of course, but the two variables do not necessarily move in tandem.

**22** Oreiro (2016) also introduces a third growth rate, the growth rate of productive capacity, which he identifies with the rate of capital accumulation. He calls the latter the 'warranted rate of growth', although it is not defined in the same way as the warranted rate of Harrod discussed in Chapters 3 and 6: Harrod's warranted rate of growth equals the ratio of the saving rate to the incremental capital–output ratio, while Oreiro's warranted rate is determined by an independent investment function in which investment depends on the utilization rate and the RER. However, the introduction of this additional variable does not alter the long-run solution from what is discussed in the text here. For further extensions of this modeling approach, see Gabriel et al. (2016).

**23** Porcile and Spinola (2018) propose a BPCG model that incorporates a North–South technological gap along with Verdoorn's law, in which the BP-equilibrium and natural rates of growth are both endogenous in the long run (although in some cases, the equilibrium is not stable).

**24** In Oreiro's model, once the long-run equilibrium natural rate of growth is determined, then one can deduce the RER that reconciles the BP-equilibrium growth rate with the natural rate and the utilization rate that reconciles the rate of capital accumulation with these two other rates. The solution for the long-run utilization rate is based on the following investment demand function (where $g = I/K$):

$$g = g_0 + g_1 u + g_2(EP_f/P)$$

Assuming that $g = y_N = y_D$ must hold in a steady-state equilibrium, and using the other equations specified above, the equilibrium utilization rate is

$$u^* = \left(\frac{1}{g_1}\right)\left\{\left[\left(\frac{q_0 + \bar{n}}{1 - \rho}\right) - g_0\right] - g_2\left(\frac{EP_f}{P}\right)^*\right\}$$

where $(EP_f/P)^*$ is given by equation (10.22).

**25** The need to avoid currency overvaluation is emphasized in the 'new developmentalist' approach, for example by Bresser-Periera et al. (2015), who especially stress how overvaluation can lead to deindustrialization or 'Dutch disease'. But as discussed earlier, the work of Ribeiro et al. (2016, 2017b, 2018) suggests that the effects of currency under- or overvaluation are more complex and ambiguous than they are usually thought to be.

**26** One might think that it should be the capital invested in the export sector that matters to export supply, but Razmi uses the total capital stock presumably on the grounds that there is only one aggregated domestic industrial sector in the model. Also, even some of the capital not explicitly invested in export production could contribute to potential output of exports, for example via the provision of necessary inputs or infrastructure.

**27** See Oreiro's investment function as shown in note 24 to this chapter.

**28** Note the implicit assumption here of openness to foreign direct investment and participation in global supply chains, but not necessarily openness to capital flows in the financial sense as discussed in section 9.3.1. For empirical evidence that investment responds positively to RER depreciation in a major emerging country, see Ibarra (2015b) and Ibarra and Ros (2018). See also Blecker (2007) on the real value of the US dollar and investment in US manufacturing.

**29** The model shown here is a simplified version of Razmi's, designed to highlight only the small country case. In fact, Razmi (2016a) develops a more general model that nests the two cases of traditional BPCG (Thirlwall's law, assuming a Keynesian small open economy) and his small country alternative as special cases. Aside from notational differences, however, the solution of the model for a small country given here is the same as in Razmi (2016a).

**30** It is not clear why $P_f$ could not increase at an exogenously given positive rate in a small open economy. An exogenous variable need not remain constant.

**31** Students of international finance will recognize here the standard implication of the small country model that the domestic inflation rate equals the nominal depreciation rate of the home currency (holding foreign prices constant). Given that the import price remains unchanged ($\hat{P}_f = 0$), relative PPP necessarily holds in the small country model (since $\hat{E} + \hat{P}_f - \hat{P} = \hat{E} - \hat{E} = 0$), which again is as expected. However, absolute PPP does not hold because the *level* of the RER ($EP_f/P$) can vary.

**32** Razmi's solution does imply that the rate of nominal depreciation $\hat{E}$ could also play a role in affecting the BP-equilibrium growth rate for a small country, $y_S$, because the domestic price level must also increase at this rate (since $\hat{P} = \hat{E}$) and export supply is price-elastic per equation (10.23) or (10.24). Razmi does not pay much attention to this implication in drawing his policy conclusions, however.

**33** Ros borrows most directly from Dutt (2002), but also acknowledges earlier contributions such as Dutt (1990) and Taylor (1981, 1983).

**34** As remarked in note 36 in Chapter 1, $a_1$ in our notation is the *reciprocal* of $A$ in the standard $AK$ model, or $a$ in our $aK$ version. Aside from making the notation more consistent with the rest of this book, we have also streamlined the presentation of Ros's (2013) model and corrected some typographical errors in the original. An extended version of this model is presented (in Spanish) by Clavijo and Ros (2015a); for critical discussion and debate, see Ibarra (2015a), Pérez Caldentey (2015), Vernengo (2015), Clavijo and Ros (2015b) and Blecker (2016a).

**35** Taken literally, this assumption implies that the home country is an extreme sort of enclave economy, which produces only for export and imports all necessary goods for domestic consumption and investment. However, Ros makes this assumption only for simplicity, and similar qualitative conclusions would follow under less restrictive assumptions. This strong assumption is relaxed in Clavijo and Ros (2015a), who reach similar conclusions.

**36** This is essentially Harrod's warranted rate of growth, but without the Harrodian instability discussed in Chapters 3 and 6.

**37** In the borderline case where $\eta_X = 1$, the $\eta_X g_f$ and $g_f$ lines would coincide, and the home and foreign growth rates would be equal in long-run equilibrium.

**38** Of course, $y = g$ should hold in a long-run, steady-state equilibrium as long as the capital–output ratio $a_1$ remains constant and the utilization rate $u$ stabilizes at an equilibrium or normal rate.

**39** Ros (2013) does not construct a small country model, but suggests that growth would be independent of the income elasticity of demand for exports in a small, price-taking country, similar to what we observed in the model of Razmi (2016a) covered in the previous subsection. See also Clavijo and Ros (2015a, 2015b) for an extension of Ros's framework to the small country case.

**40** Indeed, Thirlwall (2011) recognizes that one of the founders of Latin American structuralism (Prebisch, 1950) had anticipated what later became known as the strong form of the BPCG model.

**41** Garcimartín et al. (2010–11) propose a relative price adjustment mechanism that restores (relative) PPP in the context of deviations due to non-traded goods, but they do not investigate wage responses.

**42** Note especially that because we define the price elasticities as positive while they define them as negative, some of the signs in the equations below may be different from how they appear in the original article.

43 In effect, this assumes that the elasticity of the markup factor with respect to the RER is set to unity, that is, $\eta = 1$ in equation (4.40).
44 See Appendix 10.1 for details on this derivation (based on Appendix 4 in Ribeiro et al. 2017a, in our notation).
45 To see the connection to the labour share of (gross) income, Ribeiro et al. define the labour share as $\psi_L = Wa_0/P$ and the intermediate input share as $\psi_i = EP_f\delta/P$, in which case $\varphi = (Wa_0)/(Wa_0 + EP_f\delta) = \psi_L/(\psi_L + \psi_i)$ and $\psi_L$ and $\varphi$ are positively related.
46 This equation and the next include extra $(\varphi/2)$ terms that were inadvertently omitted from the published version of Ribeiro et al. (2017a), as confirmed in an email from Rafael Ribeiro on 3 August 2018. The qualitative results are not affected.
47 See also Blecker (1999, 2002a) and section 4.4.3 of Chapter 4.
48 Email correspondence to the authors from Jaime Ros, 1 October 2018. Cited with permission.

# Appendix 10.1 Deriving the rate of change in the markup factor

To derive equation (10.51), note that (using equation 10.48 and the definition of $\varphi$) equation (10.50) can be written as

$$1 + \tau = \left[\frac{\mu}{\delta}(1 - \varphi)\right]^{1/2}$$

Assuming that $\mu$ and $\delta$ are constants while $\varphi$ is a function of time ($t$),

$$\frac{d(1 + \tau)}{dt} = -\frac{1}{2}\left(\frac{\mu/\delta}{1 - \varphi}\right)^{1/2}\frac{d\varphi}{dt}$$

Then, using the above expression for $(1 + t)$ and $1 = \varphi/\varphi$,

$$\tau' = \frac{d(1 + \tau)/dt}{1 + \tau} = -\left(\frac{\varphi/2}{1 - \varphi}\right)\frac{d\varphi/dt}{\varphi}$$

Taking the log derivative of the definition of $\varphi$ with respect to time, we can find the approximation

$$\frac{d\varphi/dt}{\varphi} = \frac{d\ln\varphi}{dt} = (1 - \varphi)[\hat{W} - q - \hat{E} - \hat{P}_f]$$

Finally, substituting this derivative into the solution for $\tau'$ yields (after some cancellation) equation (10.51).

# References


Aghion, P. and P. Howitt (1992), 'A model of growth through creative destruction', *Econometrica*, 60 (2), 323–51.

Aghion, P. and P. Howitt (2009), *The Economics of Growth*, Cambridge, MA, USA: MIT Press.

Allain, O. (2015), 'Tackling the instability of growth: a Kaleckian–Harrodian model with an autonomous expenditure component', *Cambridge Journal of Economics*, 39 (5), 1351–71.

Allain, O. (2019), 'Demographic growth, Harrodian (in)stability and the supermultiplier', *Cambridge Journal of Economics*, 43 (1), 85–106.

Alleyne, D. and E.E. Francis (2008), 'Balance of payments growth in developing countries: a theoretical perspective', *Metroeconomica*, 59 (2), 189–202.

Alonso, J. and C. Garcimartín (1998–99), 'A new approach to the balance-of-payments constraint: some empirical evidence', *Journal of Post Keynesian Economics*, 21 (2), 259–82.

Amable, B., J. Henry, F. Lordon and R. Topol (1993), 'Unit root in the wage-price spiral is not hysteresis in unemployment', *Journal of Economic Studies*, 20 (1/2), 123–35.

Amable, B., J. Henry, F. Lordon and R. Topol (1994), 'Strong hysteresis versus zero-root dynamics', *Economics Letters*, 44 (1–2), 43–7.

Amable, B., J. Henry, F. Lordon and R. Topol (1995), 'Hysteresis revisited: a methodological approach', in R. Cross (ed.), *The Natural Rate of Unemployment: Reflections on 25 Years of the Hypothesis*, Cambridge, UK: Cambridge University Press, pp. 203–30.

Amable, B. and B. Verspagen (1995), 'The role of technology in market shares dynamics', *Applied Economics*, 27 (2), 197–204.

Amadeo, E.U. (1986), 'Notes on capacity utilization, distribution and accumulation', *Contributions to Political Economy*, 5, 83–94.

Amendola, G., G. Dosi and E. Papagni (1993), 'The dynamics of international competitiveness', *Review of World Economics (Weltwirtschaftliches Archiv)*, 129 (3), 451–71.

Amsden, A.H. (1989), *Asia's Next Giant: South Korea and Late Industrialization*, New York, NY, USA: Oxford University Press.

Andrews, P.W.S. (1949), *Manufacturing Business*, London, UK: Macmillan.

Anzoategui, D., D. Comin, M. Gertler and J. Martinez (2016), 'Endogenous technology adoption and R&D as sources of business cycle persistence', NBER Working Paper No. 22005, February.

Araujo, R.A. and J.F. Costa Santos (2018), 'Using non-linear estimation strategies to test an extended version of the Goodwin model to the US economy', unpublished manuscript, University of Brasilia and Federal University of Uberlandia.

Araujo, R.A. and G.T. Lima (2007), 'A structural economic dynamics approach to balance-of-payments-constrained growth', *Cambridge Journal of Economics*, 31 (5), 755–74.

Arestis, P. and W. Milberg (1993–94), 'Degree of monopoly, pricing, and flexible exchange rates', *Journal of Post Keynesian Economics*, 16 (2), 167–88.

Arestis, P. and M. Sawyer (eds) (2007), *A Handbook of Alternative Monetary Economics*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Arrow, K.J. (1962), 'The economic implications of learning by doing', *Review of Economic Studies*, 29 (3), 155–73.

Artis, M. and T. Bayoumi (1990), 'Saving, investment, financial integration, and the balance of payments', in International Monetary Fund (ed.), *Staff Studies for the World Economic Outlook 1990*, Washington, DC, USA: International Monetary Fund, pp. 19–34.

Asimakopulos, A. (1975), 'A Kaleckian theory of income distribution', *Canadian Journal of Economics*, 8 (3), 313–33.

Asimakopulos, A. (1991), *Keynes' General Theory and Accumulation*, Cambridge, UK: Cambridge University Press.

Autor, D., D. Dorn, L.F. Katz, C. Patterson and J. Van Reenen (2017), 'The fall of the labor share and the rise of superstar firms', NBER Working Paper No. 23396, May.

Backhouse, R.E. and M. Boianovsky (2016), 'Theories of stagnation in historical perspective', *European Journal of Economics and Economic Policy: Intervention*, 13 (2), 147–59.

Ball, L. (2014), 'Long-term damage from the Great Recession in OECD countries', *European Journal of Economics and Economic Policies: Intervention*, 11 (2), 149–60.

Ball, L. and R. Moffitt (2001), 'Productivity growth and the Phillips curve', in Alan B. Krueger and Robert M. Solow (eds), *The Roaring Nineties: Can Full Employment Be Sustained?* New York, NY, USA: Russell Sage Foundation, pp. 61–90.

Ballinger, R.M. (2013), 'Exchange rates, financial constraints, and investment in the U.S. manufacturing sector', unpublished PhD dissertation, American University.

Baran, P.A. and P.M. Sweezy (1966), *Monopoly Capital: An Essay on the American Economic and Social Order*, New York, NY, USA: Monthly Review Press.

Barba, A. and M. Pivetti (2009), 'Rising household debt: its causes and macroeconomic implications – a long-period analysis', *Cambridge Journal of Economics*, 33 (1), 113–37.

Barbosa-Filho, N. (2000), 'A note on the theory of demand-led growth', *Contributions to Political Economy*, 19, 19–32.

Barbosa-Filho, N.H. and L. Taylor (2006), 'Distributive and demand cycles in the US economy: a structuralist Goodwin model', *Metroeconomica*, 57 (3), 389–411.

Barrales, J. and R. von Arnim (2017), 'Longer run distributive cycles: wavelet decompositions for the US, 1948–2011', *Review of Keynesian Economics*, 5 (2), 196–217.

Basu, D. and R. Vasudevan (2013), 'Technology, distribution and the rate of profit in the US economy: understanding the current crisis', *Cambridge Journal of Economics*, 37 (1), 57–89.

Baumol, W.J., S.A.B. Blackman and E.N. Wolff (1989), *Productivity and American Leadership: The Long View*, Cambridge, MA, USA: MIT Press.

Beckerman, W. (1962), 'Projecting Europe's growth', *Economic Journal*, 72 (288), 912–25.

Benería, L., G. Berik and M.S. Floro (2015), *Gender, Development, and Globalization: Economics as if All People Mattered*, New York, NY, USA: Routledge.

Benigno, G. and L. Fornaro (2018), 'Stagnation traps', *Review of Economic Studies*, 85 (3), 1425–70.

Berg, A., J.D. Ostry and J. Zettelmeyer (2012), 'What makes growth sustained?', *Journal of Development Economics*, 98 (2), 149–66.

Bhaduri, A. (2006), 'The dynamics of profit- and wage-led expansion: a note', in E. Hein, A. Heise and A. Truger (eds), *Wages, Employment, Distribution and Growth*, Basingstoke, UK: Palgrave Macmillan, pp. 247–53.

Bhaduri, A. (2008), 'On the dynamics of profit-led and wage-led growth', *Cambridge Journal of Economics*, 32 (1), 147–60.

Bhaduri, A. and S.A. Marglin (1990), 'Unemployment and the real wage: the economic basis for contesting political ideologies', *Cambridge Journal of Economics*, 14 (4), 375–93.

Bharadwaj, K. and B. Schefold (eds) (1990), *Essays in Honour of Piero Sraffa: Critical Perspectives on New Developments of Classical Theory*, London, UK: Unwin & Hyman.

Bivens, J. (2006), 'Wages, profits, and rent-sharing in an open economy', *International Review of Applied Economics*, 20 (1), 69–83.

Bivens, J. (2017), *Inequality Is Slowing US Economic Growth*, EPI Report, Washington, DC, USA: Economic Policy Institute, 12 December.

Blackburn, K. (1999), 'Can stabilisation policy reduce long-run growth?', *Economic Journal*, 109 (452), 67–77.

Blackburn, K. and R. Galindev (2003), 'Growth, volatility and learning', *Economics Letters*, 79 (3), 417–21.

Blanchard, O. (2017), *Macroeconomics*, 7th edition, Boston, MA, USA: Pearson.

Blanchflower, D.G., A.J. Oswald and M.D. Garrett (1990), 'Insider power in wage determination', *Economica*, 57 (226), 143–70.

Blecker, R.A. (1986), 'International competition and economic growth: a model of capital accumulation in an open economy and a case study of the American steel industry, 1948–83', unpublished PhD dissertation, Stanford University, August.

Blecker, R.A. (1989a), 'International competition, income distribution and economic growth', *Cambridge Journal of Economics*, 13 (3), 395–412.

Blecker, R.A. (1989b), 'Markup pricing, import competition, and the decline of the American steel industry', *Journal of Post Keynesian Economics*, 12 (1), 70–87.

Blecker, R.A. (1991), 'Import competition, investment finance, and cumulative decline in the U.S. steel industry, 1962–1981', *International Review of Applied Economics*, 5 (2), 171–95.

Blecker, R.A. (1992), 'Structural roots of U.S. trade problems: income elasticities, secular trends, and hysteresis', *Journal of Post Keynesian Economics*, 14 (3), 321–46.

Blecker, R.A. (1996), 'The new economic integration: structuralist models of North-South trade and investment liberalization', *Structural Change and Economic Dynamics*, 7 (3), 321–45.

Blecker, R.A. (1997a), 'Policy implications of the international saving-investment correlation', in R. Pollin (ed.), *The Macroeconomics of Finance, Saving, and Investment*, Ann Arbor, MI: University of Michigan Press, pp. 173–229.

Blecker, R.A. (1997b), 'The "unnatural and retrograde order": Adam Smith's theories of trade and development reconsidered', *Economica*, 64 (255), 527–37.

Blecker, R.A. (1998), 'International competition, relative wages, and the balance-of-payments constraint', *Journal of Post Keynesian Economics*, 20 (4), 495–526.

Blecker, R.A. (1999), 'Kaleckian macro models for open economies', in J. Deprez and J.T. Harvey (eds), *Foundations of International Economics: Post Keynesian Perspectives*, London, UK: Routledge, pp. 116–49.

Blecker, R.A. (2002a), 'Demand, distribution, and growth in neo-Kaleckian macro models', in M. Setterfield (ed.), *The Economics of Demand-Led Growth: Challenging the Supply-Side Vision of the Long Run*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 129–52.

Blecker, R.A. (2002b), 'The balance-of-payments-constrained growth model and the limits to export-led growth', in P. Davidson (ed.), *A Post Keynesian Perspective on Twenty-First Century Economic Problems*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 69–88.

Blecker, R.A. (2007), 'The economic consequences of dollar appreciation for U.S. manufacturing profits and investment: a time-series analysis', *International Review of Applied Economics*, 21 (4), 491–517.

Blecker, R.A. (2008), 'U.S. steel import tariffs: the politics of global markets', in R.G. Carter (ed.), *Contemporary Cases in U.S. Foreign Policy: From Terrorism to Trade*, 3rd edition, Washington, DC, USA: Congressional Quarterly Press, pp. 249–80.

Blecker, R.A. (2009), 'External shocks, structural change, and economic growth in Mexico, 1979–2007', *World Development*, 37 (7), 1274–84.

Blecker, R.A. (2011), 'Open economy models of growth and distribution', in E. Hein and

E. Stockhammer (eds), *A Modern Guide to Keynesian Macroeconomics and Economic Policies*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 215–39.

Blecker, R.A. (2013a), 'Global imbalances and the U.S. trade deficit', in B.Z. Cynamon, S.M. Fazzari and M. Setterfield (eds), *After the Great Recession: The Struggle for Economic Recovery and Growth*, New York, NY, USA: Cambridge University Press, pp. 187–215.

Blecker, R.A. (2013b), 'Long-run growth in open economies: export-led cumulative causation or a balance-of-payments constraint?', in G.C. Harcourt and P. Kriesler (eds), *The Oxford Handbook of Post-Keynesian Economics*, vol. I, *Theory and Origins*, Oxford, UK: Oxford University Press, pp. 390–414.

Blecker, R.A. (2016a), 'The debate over "Thirlwall's law": balance-of-payments-constrained growth reconsidered', *European Journal of Economics and Economic Policy: Intervention*, 13 (3), 275–90.

Blecker, R.A. (2016b), 'The US economy since the crisis: slow recovery and secular stagnation', *European Journal of Economics and Economic Policy: Intervention*, 13 (2), 203–14.

Blecker, R.A. (2016c), 'Wage-led versus profit-led demand regimes: the long and the short of it', *Review of Keynesian Economics*, 4 (4), 373–90.

Blecker, R.A. (2020), 'Long-term shifts in demand and distribution in neo-Kaleckian and neo-Goodwinian models', in H. Bougrine and L.-P. Rochon (eds), *Money and Crises in Post-Keynesian Economics: Essays in Honour of Marc Lavoie and Mario Seccareccia*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, forthcoming.

Blecker, R.A. and C.A. Ibarra (2013), 'Trade liberalization and the balance of payments constraint with intermediate imports: the case of Mexico revisited', *Structural Change and Economic Dynamics*, 25 (June), 33–47.

Blecker, R.A. and A. Razmi (2008), 'The fallacy of composition and contractionary devaluations: output effects of real exchange rate shocks in semi-industrialised countries', *Cambridge Journal of Economics*, 32 (1), 83–109.

Blecker, R.A. and A. Razmi (2010), 'Export-led growth, real exchange rates, and the fallacy of composition', in M. Setterfield (ed.), *The Handbook of Alternative Theories of Economic Growth*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 379–96.

Blecker, R. and S. Seguino (2002), 'Macroeconomic effects of reducing gender wage inequality in an export-oriented, semi-industrialized economy', *Review of Development Economics*, 6 (1), 103–19.

Blecker, R. and S. Seguino (2007), 'Macroeconomic effects of reducing gender wage inequality in an export-oriented, semi-industrialized economy', in I. van Stavern, D. Elson, C. Grown and N. Çağatay (eds), *The Feminist Economics of Trade*, Abingdon, UK: Routledge, pp. 91–114.

Boddy, R. and J. Crotty (1975), 'Class conflict and macro policy: the political business cycle', *Review of Radical Political Economics*, 7 (1), 1–19.

Boggio, L. and L. Barbieri (2017), 'International competitiveness in post-Keynesian growth theory: controversies and empirical evidence', *Cambridge Journal of Economics*, 41 (1), 25–47.

Bortis, H. (1997), *Institutions, Behavior, and Economic Theory: A Contribution to Classical-Keynesian Political Economy*, Cambridge, UK: Cambridge University Press.

Botte, F. (2019), 'Endogenous business cycles and Harrodian instability in an agent-based model', *Journal of Post Keynesian Economics*, 42 (2), 232–54.

Bowles, S. and R. Boyer (1988), 'Labor discipline and aggregate demand: a macroeconomic model', *American Economic Review*, 78 (2), 395–400.

Bowles, S. and R. Boyer (1990), 'A wage-led employment regime: income distribution, labour discipline, and aggregate demand in welfare capitalism', in S.A. Marglin and J.B. Schor (eds), *The Golden Age of Capitalism: Reinterpreting the Postwar Experience*, Oxford, UK: Clarendon Press, pp. 187–217.

Branson, W.H. (1983), 'Economic structure and policy for external balance', in A.W. Hooke (ed.), *Exchange Rate Regimes and Policy Interdependence*, Washington, DC, USA: International Monetary Fund, pp. 39–66.

Braunstein, E., I. van Staveren and D. Tavani (2011), 'Embedding care and unpaid work in macro-economic modeling: a structuralist approach', *Feminist Economics*, 17 (4), 5–31.

Bresser-Pereira, L.C., J.L. Oreiro and N. Marconi (2015), *Developmental Macroeconomics: New Developmentalism as a Growth Strategy*, Abingdon, UK: Routledge.

Brochier, L. and A.C. Macedo e Silva (2019), 'A supermultiplier Stock-Flow Consistent model: the "return" of the paradoxes of thrift and costs in the long run?' *Cambridge Journal of Economics*, 43 (2), 413–42.

Brouwer, E. and A. Kleinknecht (1999), 'Keynes-plus? Effective demand and changes in firm-level R&D: an empirical note', *Cambridge Journal of Economics*, 23 (3), 385–91.

Brown, C. (2008), *Inequality, Consumer Credit and the Saving Puzzle*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Bruno, O. (1999), 'Long-run positions and short-run dynamics in a classical growth model', *Metroeconomica*, 50 (1), 119–37.

Burdekin, R.C.K. and P. Burkett (1988), 'Monetary accommodation of income claims and the expectations-augmented Phillips curve: in search of a stable policy rule', *Weltwirtschäftliches Archiv*, 124, 169–77.

Burgstaller, A. (1985), 'North-South trade and capital flows in a Ricardian model of accumulation', *Journal of International Economics*, 18 (3–4), 241–60.

Campbell, J.Y. and G.N. Mankiw (1989), 'Consumption, income and interest rates: reinterpreting the time series evidence', in O.J. Blanchard and S. Fischer (eds), *NBER Macro-economics Annual 1989*, vol. 4, Cambridge, MA, USA: MIT Press, pp. 185–216.

Canova, F. (1998), 'Detrending and business cycle facts', *Journal of Monetary Economics*, 41 (3), 475–512.

Carlin, W., A. Glyn and J. Van Reenan (2001), 'Export market performance of OECD countries: an empirical examination of the role of cost competitiveness', *Economic Journal*, 111 (468), 128–62.

Carlin, W. and D. Soskice (2005), *Contributions to Macroeconomics*, 5 (1), 1–36.

Carlin, W. and D. Soskice (2009), 'Teaching intermediate macroeconomics using the 3-equation model', in G. Fontana and M. Setterfield (eds), *Macroeconomic Theory and Macroeconomic Pedagogy*, Basingstoke, UK: Palgrave Macmillan, pp. 13–35.

Carlin, W. and D. Soskice (2015), *Macroeconomics: Institutions, Instability, and the Financial System*, Oxford, UK: Oxford University Press.

Carroll, C.D. and L.H. Summers (1991), 'Consumption growth parallels income growth: some new evidence', in B.D. Bernheim and J.B. Shoven (eds), *National Saving and Economic Performance*, Chicago, IL, USA: University of Chicago Press, pp. 305–47.

Carvalho, L. and A. Rezai (2016), 'Personal income inequality and aggregate demand', *Cambridge Journal of Economics*, 40 (2), 491–505.

Casarosa, C. (1978), 'A new formulation of the Ricardian system', *Oxford Economic Papers*, 30 (1), 38–63.

Cauvel, M. (2018), 'Three essays on the empirical estimation of wage-led and profit-led demand regimes', unpublished PhD dissertation, American University, July.

Caverzasi, E. and A. Godin (2015), 'Post-Keynesian stock-flow-consistent modelling: a survey', *Cambridge Journal of Economics*, 39 (1), 157–87.

Cerra, V., U. Panizza and S.C. Saxena (2013), 'International evidence on recovery from recessions', *Contemporary Economic Policy*, 31 (2), 424–39.

Cerra, V. and S.C. Saxena (2008), 'Growth dynamics: the myth of economic recovery', *American Economic Review*, 98 (1), 429–57.

Cerra, V. and S.C. Saxena (2017), 'Booms, crises, and recoveries: a new paradigm of the business cycle and its policy implications', IMF Working Paper, WP/17/250, November.

Cesaratto, S. (2015), 'Neo-Kaleckian and Sraffian controversies on the theory of accumulation', *Review of Political Economy*, 27 (2), 154–82.

Charles, S. (2008), 'Corporate debt, variable retention rate and the appearance of financial fragility', *Cambridge Journal of Economics*, 32 (5), 781–95.

Charpe, M., S. Bridji and P. McAdam (2018), 'Labor share and growth in the long run', manuscript, International Labour Organization and DG Research, European Central Bank, 11 November.

Chenery, H.B. and M. Bruno (1962), 'Development alternatives in an open economy: the case of Israel', *Economic Journal*, 72 (285), 79–103.

Chenery, H.B. and A.M. Strout (1966), 'Foreign assistance and economic development', *American Economic Review*, 56 (4), 679–733.

Chick, V. and M. Caserta (1997), 'Provisional equilibrium and macroeconomic theory', in P. Arestis, G. Palma and M. Sawyer (eds), *Markets, Unemployment and Economic Policy: Essays in Honour of Geoff Harcourt*, vol. 2, Abingdon UK: Routledge, pp. 223–37.

Chinn, M.D. (2004), 'Incomes, exchange rates, and the U.S. trade deficit, once again', *International Finance*, 7 (3), 451–69.

Chirinko, R.S., S.M. Fazzari and A.P. Meyer (1999), 'How responsive is business capital formation to its user cost? An exploration with micro data', *Journal of Public Economics*, 74 (1), 53–80.

Chirinko, R.S., S.M. Fazzari and A.P. Meyer (2011), 'A new approach to estimating production function parameters: the elusive capital–labor substitution elasticity', *Journal of Business and Economic Statistics*, 29 (4), 587–94.

Cimoli, M. (1988), 'Technological gaps and institutional asymmetries in a North-South model with a continuum of goods', *Metroeconomica*, 39 (3), 245–74.

Cimoli, M. and G. Porcile (2010), 'Specialization, wage bargaining and technology in a multi-goods growth model', *Metroeconomica*, 61 (1), 219–38.

Cimoli, M. and G. Porcile (2014), 'Technology, structural change and BOP-constrained growth: a structuralist toolbox', *Cambridge Journal of Economics*, 38 (1), 215–37.

Clavijo, P.H. and J. Ros (2015a), 'La ley de Thirlwall: una lectura crítica', *Investigación Económica*, 74 (292), 11–40.

Clavijo, P.H. and J. Ros (2015b), 'Respuesta a Ibarra, Pérez y Vernengo', *Investigación Económica*, 74 (292), 81–90.

Cline, W.R. (1989), *United States External Adjustment and the World Economy*, Washington, DC, USA: Institute for International Economics.

Cogley, T. and J.M. Nason (1995), 'Effects of the Hodrick-Prescott filter on trend and difference stationary time series: implications for business cycle research', *Journal of Economic Dynamics and Control*, 19 (1–2), 253–78.

Colander, D. (1999), 'A Post-Walrasian explanation of wage and price inflexibility and a Keynesian unemployment equilibrium system', in M. Setterfield (ed.), *Growth, Employment and Inflation: Essays in Honour of John Cornwall*, Basingstoke, UK: Macmillan, pp. 211–25.

Comin, D. and M. Gertler (2006), 'Medium-term business cycles', *American Economic Review*, 96 (3), 523–51.

Cordero, J.A. (2002), 'A model of growth and conflict inflation for a small open economy', *Metroeconomica*, 53 (3), 261–89.

Cornwall, J. (1972), *Growth and Stability in a Mature Economy*, London: Martin Robertson.

Cornwall, J. (1977), *Modern Capitalism: Its Growth and Transformation*, London, UK: Martin Robertson.

Cornwall, J. (1991), 'Prospects for unemployment in the 1990s with hysteresis', in J. Cornwall (ed.), *The Capitalist Economies: Prospects for the 1990s*, Aldershot, UK and Brookfield, VT, USA: Edward Elgar Publishing.

Cornwall, J. and W. Cornwall (2001), *Capitalist Development in the Twentieth Century: An Evolutionary Keynesian Analysis*, Cambridge, UK: Cambridge University Press.

Cross, R. (1995), 'Is the natural rate hypothesis consistent with hysteresis?', in R. Cross (ed.), *The Natural Rate of Unemployment: Reflections on 25 Years of the Hypothesis*, Cambridge, UK: Cambridge University Press, pp. 181–200.

Cynamon, B.Z. and S. Fazzari (2008), 'Household debt in the consumer age: source of growth – risk of collapse', *Capitalism and Society*, 3 (2), Article 3.

Cynamon, B.Z. and S.M. Fazzari (2017a), 'Household income, demand, and saving: deriving macro data with micro data concepts', *Review of Income and Wealth*, 63 (1), 53–69.

Cynamon, B.Z. and S.M. Fazzari (2017b), 'Secular demand stagnation in the 21st century U.S. economy', paper presented at INET conference, 21 October, https://www.ineteconomics.org/research/research-papers/secular-demand-stagnation-in-the-21st-century-u-s-economy-1, accessed 13 June 2019.

Dallery, T. (2007), 'Kaleckian models of growth and distribution revisited: evaluating their relevance through simulations', paper presented at the 11th Workshop of the Research Network *Macroeconomic Policies*, Berlin, October.

Dallery, T. and T. van Treeck (2011), 'Conflicting claims and equilibrium adjustment processes in a stock-flow consistent macro model', *Review of Political Economy*, 23 (2), 189–211.

Davidson, P. (1990–91), 'A Post Keynesian positive contribution to theory', *Journal of Post Keynesian Economics*, 13 (2), 298–303.

Davidson, P. (1992), *International Money and the Real World*, New York, NY, USA: St. Martin's Press.

De Loecker, J. and J. Eeckhout (2017), 'The rise of market power and the macroeconomic implications', NBER Working Paper No. 23687, August.

Del Monte, A. (1975), 'Grado di monopolio e sviluppo economico', *Rivista Internazionale di Scienze Sociale*, 83 (3), 231–63.

Dixon, R. and A.P. Thirlwall (1975), 'A model of regional growth rate differences along Kaldorian lines', *Oxford Economic Papers*, 27 (2), 201–14.

Domar, E.D. (1946), 'Capital expansion, rate of growth, and employment', *Econometrica*, 14 (2), 137–47.

Dornbusch, R. (1980), *Open Economy Macroeconomics*, New York, NY, USA: Basic Books.

Dornbusch, R. (1987), 'Exchange rates and prices', *American Economic Review*, 77 (1), 93–106.

Dornbusch, R., S. Fischer and P. Samuelson (1977), 'Comparative advantage, trade, and payments in a Ricardian model with a continuum of goods', *American Economic Review*, 67 (5), 823–39.

Dosi, G., K. Pavitt and L. Soete (1990), *The Economics of Technical Change and International Trade*, Brighton, UK: Wheatsheaf.

Duflo, E. (2012), 'Women, empowerment, and economic development', *Journal of Economic Literature*, 50 (4), 1051–79.

Duménil, G. and D. Lévy (1993), *The Economics of the Profit Rate: Competition, Crises, and Historical Tendencies in Capitalism*, Aldershot, UK and Brookfield, VT, USA: Edward Elgar Publishing.

Duménil, G. and D. Lévy (1995), 'A stochastic model of technical change: an application to the US economy (1869–1989)', *Metroeconomica*, 46 (3), 213–45.

Duménil, G. and D. Lévy (1999), 'Being Keynesian in the short term and classical in the long term: the traverse to classical long-term equilibrium', *The Manchester School*, 67 (6), 684–716.

Duménil, G. and D. Lévy (2003), 'Technology and distribution: historical trajectories à la Marx', *Journal of Economic Behavior and Organization*, 52 (2), 201–33.

Dutt, A.K. (1984), 'Stagnation, income distribution and monopoly power', *Cambridge Journal of Economics*, 8 (1), 25–40.

Dutt, A.K. (1987), 'Alternative closures again: a comment on "Growth, distribution and inflation"', *Cambridge Journal of Economics*, 11 (1), 75–82.

Dutt, A.K. (1988), 'Monopoly power and uneven development: Baran revisited', *Journal of Development Studies*, 24 (2), 161–76.

Dutt, A.K. (1990), *Growth, Distribution and Uneven Development*, Cambridge, UK: Cambridge University Press.

Dutt, A.K. (1997), 'Equilibrium, path dependence and hysteresis in post-Keynesian models', in P. Arestis, G. Palma and M. Sawyer (eds), *Markets, Employment and Economic Policy: Essays in Honour of Geoff Harcourt*, vol. 2, Abingdon, UK: Routledge, pp. 238–53.

Dutt, A.K. (2002), 'Thirlwall's law and uneven development', *Journal of Post Keynesian Economics*, 24 (3), 367–90.

Dutt, A.K. (2005), 'Consumption, debt and growth', in M. Setterfield (ed.), *Interactions in Analytical Political Economy*, Armonk, NY, USA: M.E. Sharpe.

Dutt, A.K. (2006a), 'Aggregate demand, aggregate supply and economic growth', *International Review of Applied Economics*, 20 (3), 319–36.

Dutt, A.K. (2006b), 'Maturity, stagnation and consumer debt: a Steindlian approach', *Metroeconomica*, 57 (3), 339–64.

Dutt, A.K. (2008), 'The dependence effect, consumption and happiness: Galbraith revisited', *Review of Political Economy*, 20 (4), 527–50.

Dutt, A.K. (2009), 'Path dependence, equilibrium and economic growth', in P. Arestis and M. Sawyer (eds), *Path Dependency in Macroeconomics: International Papers in Political Economy*, Basingstoke, UK: Palgrave Macmillan.

Dutt, A.K. (2010a), 'Equilibrium, stability and path dependence in post-Keynesian models of economic growth', in A. Birolo, D. Foley, H.D. Kurz, B. Schefold and I. Steedman (eds), *Production, Distribution and Trade: Alternative Perspectives. Essays in Honour of Sergio Parrinello*, Abingdon, UK: Routledge, pp. 233–53.

Dutt, A.K. (2010b), 'Reconciling the growth of aggregate demand and aggregate supply', in M. Setterfield (ed.), *Handbook of Alternative Theories of Economic Growth*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 220–40.

Dutt, A.K. (2018), 'Some observations on models of growth and distribution with autonomous demand growth', *Metroeconomica*, published online, 5 December, https://doi.org/10.1111/meca.12234.

Dutt, A.K. and J. Ros (2007), 'Aggregate demand shocks and economic growth', *Structural Change and Economic Dynamics*, 18 (1), 75–99.

Eatwell, J. and M. Milgate (eds) (1983), *Keynes's Economics and the Theory of Value and Distribution*, London, UK: Duckworth.

Eichner, A. (1976), *The Megacorp and Oligopoly: Micro Foundations of Macro Dynamics*, Cambridge, UK: Cambridge University Press.

Elmslie, B. (1994), 'Positive feedback mechanisms in Adam Smith's theories of international trade', *European Journal of the History of Economic Thought*, 1 (2), 253–71.

Epstein, G. and H. Gintis (1992), 'International capital markets and the limits of national economic policy', in T. Banuri and J.B. Schor (eds), *Financial Openness and National Autonomy: Opportunities and Constraints*, Oxford, UK: Oxford University Press, pp. 167–97.

Fazzari, S.M. (1985), 'Keynes, Harrod, and the rational expectations revolution', *Journal of Post Keynesian Economics*, 8 (1), 66–80.

Fazzari, S.M. (1993), 'Monetary policy, financial structure, and investment', in G.A. Dymski, G. Epstein and R. Pollin (eds), *Transforming the U.S. Financial System: Equity and Efficiency for the 21st Century*, Armonk, NY: M.E. Sharpe, pp. 35–63.

Fazzari, S.M., P.E. Ferri, E.G. Greenberg and A.M. Variato (2013), 'Aggregate demand, instability, and growth', *Review of Keynesian Economics*, 1 (1), 1–21.

Fazzari, S.M. and E. Greenberg (2015), 'Are macroeconomic models with ceilings and floors useful in understanding the Great Recession?', in A. Cristini, S.M. Fazzari, E. Greenberg and R. Leoni (eds), *Cycles, Growth and the Great Recession*, Abingdon, UK: Routledge, pp. 45–65.

Fazzari, S.M., R.G. Hubbard and B.C. Peterson (1988), 'Financing constraints and corporate investment', *Brookings Papers on Economic Activity*, 1988 (1), 141–206.

Feenstra, R.C. (1989), 'Symmetric pass-through of tariffs and exchange rates under imperfect competition: an empirical test', *Journal of International Economics*, 27 (1), 25–45.

Fernandez, J. (2005), 'Income distribution and aggregate demand in the U.S. economy: an empirical investigation', unpublished PhD dissertation, American University.

Ferri, P., S. Fazzari, E. Greenberg and A. Variato (2011), 'Aggregate demand, Harrod's instability and fluctuations', *Computational Economics*, 38 (3), 209–20.

Ferri, P. and H. Minsky (1992), 'Market processes and thwarting systems', *Structural Change and Economic Dynamics*, 3 (1), 79–91.

Fiebiger, B. (2018), 'Semi-autonomous household expenditures as the *causa causans* of postwar US business cycles: the stability and instability of Luxemburg-type external markets', *Cambridge Journal of Economics*, 42 (1), 155–75.

Fisher, F.M. (1983), *Disequilibrium Foundations of Equilibrium Economics*, Cambridge, UK: Cambridge University Press.

Fisher, I. (1930), *The Theory of Interest as Determined by Impatience to Spend Income and Opportunity to Invest It*, New York, NY, USA: Macmillan.

Flaschel, P., R. Franke and W. Semmler (1997), *Dynamic Macroeconomics: Instability, Fluctuations and Growth in Monetary Economics*, Cambridge, MA, USA: MIT Press.

Floro, M.S. (1995), 'Economic restructuring, gender and the allocation of time', *World Development*, 23 (11), 1913–29.

Foley, D.K. (1986), *Understanding Capital: Marx's Economic Theory*, Cambridge, MA, USA: Harvard University Press.

Foley, D.K. and T.R. Michl (1999), *Growth and Distribution*, Cambridge, MA, USA: Harvard University Press.

Foley, D.K. and T.R. Michl (2010), 'The classical theory of growth and distribution', in M. Setterfield (ed.), *Handbook of Alternative Theories of Economic Growth*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 49–63.

Foley, D.K., T.R. Michl and D. Tavani (2019), *Growth and Distribution*, 2nd edition, Cambridge, MA, USA: Harvard University Press.

Foster, J.B. (2014), *The Theory of Monopoly Capitalism*, new edition, New York, NY, USA: Monthly Review Press.

Franke, R. (2015), 'An examination of Professor Shaikh's proposal to tame Harrodian instability', *European Journal of Economics and Economic Policies: Intervention*, 12 (1), 7–19.

Franke, R. (2017), 'A simple approach to overcome the problems arising from the Keynesian stability condition', *European Journal of Economics and Economic Policies: Intervention*, 14 (1), 48–69.

Franke, R. (2018), 'Can monetary policy tame Harrodian instability?', *Metroeconomica*, 69 (3), 593–618.

Franke, R. (2019), 'On Harrodian instability: two stabilizing mechanisms may be jointly destabilizing', *Review of Keynesian Economics*, 7 (1), 43–56.

Freitas, F. and F. Serrano (2015), 'Growth rate and level effects, the stability of the adjustment of capacity to demand, and the Sraffian supermultiplier', *Review of Political Economy*, 27 (3), 258–81.

Friedman, M. (1970), 'Counter-revolution in monetary theory', First Wincott Memorial Lecture, University of London, Institute of Economic Affairs, Occasional paper 33.

Furtado, C. (1963), *The Economic Growth of Brazil*, trans. Ricardo W. de Aguiar and Eric Charles Drysdale, Berkeley, CA, USA: University of California Press.

Gabriel, L.F., F.G. Jayme and J.L. Oreiro (2016), 'A North-South model of economic growth, technological gap, structural change and real exchange rate', *Structural Change and Economic Dynamics*, 38 (September), 83–94.

Gaggl, P. and S. Steindl (2008), 'On business cycle models and their extensions to endogenous growth theory', Austrian Institute of Economic Research (WIFO), mimeo.

Galindev, R. (2009), 'Note on "Growth, cycles and stabilization policy"', *Oxford Economic Papers*, 61 (1), 201–6.

Galor, O. (2011), *Unified Growth Theory*, Princeton, NJ, USA: Princeton University Press.

Garcimartín, C., L.A. Rivas and P. García Martínez (2010–11), 'On the role of relative prices and capital flows in balance-of-payment-constrained growth: the experiences of Portugal and Spain in the euro area', *Journal of Post Keynesian Economics*, 33 (2), 281–306.

Godley, W. (1999), 'Open economy macroeconomics using models of closed systems', Working Paper No. 281, Jerome Levy Institute of Bard College.

Godley, W. and A. Izurieta (2002), 'The case for a severe recession', *Challenge*, 45 (2), 27–51.

Godley, W. and M. Lavoie (2007), *Monetary Economics: An Integrated Approach to Credit, Money, Income, Production and Wealth*, Basingstoke, UK: Palgrave Macmillan.

Goodhart, C.A.E. (2005), 'Safeguarding good policy practice', *Federal Reserve Bank of St. Louis Review*, 87, 298–302.

Goodwin, R.M. (1967), 'A growth cycle', in Feinstein, C. (ed.), *Socialism, Capitalism and Economic Growth*, Cambridge, UK: Cambridge University Press, pp. 54–8.

Gordon, D.M. (1991), 'Kaldor's macro system: too much cumulation, too few contradictions', in E.J. Nell and W. Semmler (eds), *Nicholas Kaldor and Mainstream Economics*, New York, NY, USA: St. Martin's Press, pp. 518–48.

Gordon, D.M., R. Edwards and M. Reich (1982), *Segmented Work, Divided Workers*, Cambridge, UK: Cambridge University Press.

Gordon, R.J. and R. Krenn (2010), 'The end of the Great Depression 1939–41: policy contributions and fiscal multipliers', NBER Working Paper No. 16380, September.

Gouri Suresh, S. and M. Setterfield (2015), 'Firm performance, macroeconomic performance, and "animal spirits" in a post-Keynesian model of aggregate fluctuations', *Journal of Post Keynesian Economics*, 38 (1), 38–63.

Gouvêa, R.R. and G.T. Lima (2010), 'Structural change, balance-of-payments constraint, and economic growth: evidence from the multisectoral Thirlwall's law', *Journal of Post Keynesian Economics*, 33 (1), 169–204.

Gouvêa, R.R. and G.T. Lima (2013), 'Balance-of-payments-constrained growth in a multisectoral framework: a panel data investigation', *Journal of Economic Studies*, 40 (2), 240–54.

Grasselli, M.R. and A. Maheshwari (2017), 'A comment on "Testing Goodwin: growth cycles in ten OECD countries"', *Cambridge Journal of Economics*, 41 (6), 1761–6.

Grossman, G.M. and E. Helpman (1991), *Innovation and Growth in the Global Economy*, Cambridge, MA, USA: MIT Press.

Hahnel, R. and H. Sherman (1982), 'The rate of profit over the business cycle', *Cambridge Journal of Economics*, 6 (2), 185–94.

Halevi, J. and P. Kriesler (1992), 'An introduction to the traverse in economic theory', in J. Halevi,

D. Laibman and E.J. Nell (eds), *Beyond the Steady State: A Revival of Growth Theory*, New York, NY, USA: St. Martin's Press, pp. 225–34.

Hall, R.L. and C.J. Hitch (1939), 'Price theory and business behaviour', *Oxford Economic Papers*, 2 (1), 12–45, reprinted in M.C. Sawyer (ed.) (1988), *Post-Keynesian Economics*, Aldershot, UK and Brookfield, VT, USA: Edward Elgar Publishing, pp. 205–26.

Halsmayer, V. and K.D. Hoover (2016), 'Solow's Harrod: transforming macroeconomic dynamics into a model of long-run growth', *European Journal of the History of Economic Thought*, 23 (4), 561–96.

Halter, D., M. Oechslin and J. Zweimüller (2014), 'Inequality and growth: the neglected time dimension', *Journal of Economic Growth*, 19 (1), 81–104.

Hamilton, J.D. (2018), 'Why you should never use the Hodrick-Prescott filter', *Review of Economics and Statistics*, 100 (5), 831–43.

Harcourt, G.C. (1972), *Some Cambridge Controversies in the Theory of Capital*, Cambridge, UK: Cambridge University Press.

Harcourt, G.C. (1981), 'Marshall, Sraffa and Keynes: incompatible bedfellows?', *Eastern Economic Journal*, 7(1), 39–50, reprinted in G.C. Harcourt (1982), *The Social Science Imperialists: Selected Essays of G.C. Harcourt* (ed.), London, UK: Routledge and Kegan Paul, pp. 205–22.

Harcourt, G.C. (1992), 'Markets, madness and a middle way', in *Capitalism, Socialism and Post-Keynesianism: Selected Essays of G.C. Harcourt*, Aldershot, UK and Brookfield, VT, USA: Edward Elgar Publishing, pp. 9–24.

Harcourt, G.C. (2001), 'The Kaldor legacy', in *50 Years a Keynesian and Other Essays*, Basingstoke, UK: Palgrave, pp. 238–59.

Harcourt, G.C. (2007), 'The theoretical and political importance of the economics of Keynes. Or, what would Marx and Keynes have made of the happenings of the past 30 years and more?', in M. Forstater, G. Mongiovi and S. Pressman (eds), *Post Keynesian Macroeconomics: Essays in Honour of Ingrid Rima*, Abingdon, UK: Routledge, pp. 56–69.

Harris, D.J. (1974), 'The price policy of firms, the level of employment and distribution of income in the short run', *Australian Economic Papers*, 13 (22), 144–51.

Harris, D.J. (1978), *Capital Accumulation and Income Distribution*, Palo Alto, CA, USA: Stanford University Press.

Harris, D.J. (1983), 'Accumulation of capital and the rate of profit in Marxian theory', *Cambridge Journal of Economics*, 7 (3–4), 311–30.

Harris, D.J. (1986), 'Are there macroeconomic laws? The "law" of the falling rate of profit reconsidered', in H.-J. Wagener and H.W. Drukker (eds), *The Economic Law of Motion of Modern Society: A Marx-Keynes-Schumpeter Centennial*, Cambridge, UK: Cambridge University Press, pp. 49–70.

Harris, D.J. (1991), 'Equilibrium and stability in classical theory', in E.J. Nell and W. Semmler (eds), *Nicholas Kaldor and Mainstream Economics*, London, UK: Macmillan, pp. 88–100.

Harris, D.J. (2005), 'Joan Robinson on "history versus equilibrium"', in B. Gibson (ed.), *Joan Robinson's Economics*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 81–108.

Harrod, R.F. (1933), *International Economics*, Cambridge, UK: Cambridge University Press.

Harrod, R.F. (1936), *The Trade Cycle*, Oxford, UK: Clarendon Press.

Harrod, R.F. (1939), 'An essay in dynamic theory', *Economic Journal*, 49 (193), 14–33.

Harrod, R.F. (1948), *Towards a Dynamic Economics*, London, UK: Macmillan.

Harrod, R.F. (1970), 'Harrod after twenty-one years: a comment', *Economic Journal*, 80 (319), 737–41.

Harrod, R.F. (1973), *Economic Dynamics*, London, UK: Macmillan.

Harvie, D. (2000), 'Testing Goodwin: growth cycles in ten OECD countries', *Cambridge Journal of Economics*, 24 (3), 349–76.

Heilbroner, R. and W. Milberg (1996), *The Crisis of Vision in Modern Economic Thought*, Cambridge, UK: Cambridge University Press.

Hein, E. (2006), 'Wage bargaining and monetary policy in a Kaleckian monetary distribution and growth model: trying to make sense of the NAIRU', *European Journal of Economics and Economic Policy: Intervention*, 3 (2), 305–29.

Hein, E. (2008), *Money, Distribution Conflict and Capital Accumulation: Contributions to 'Monetary Analysis'*, Basingstoke, UK: Palgrave Macmillan.

Hein, E. (2012a), 'Financialization, distribution, capital accumulation, and productivity growth in a post-Kaleckian model', *Journal of Post Keynesian Economics*, 34 (3), 475–96.

Hein, E. (2012b), *The Macroeconomics of Finance-Dominated Capitalism – And Its Crisis*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Hein, E. (2014), *Distribution and Growth after Keynes: A Post-Keynesian Guide*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Hein, E. (2016), 'Secular stagnation or stagnation policy? Steindl after Summers', *PSL Quarterly Review*, 69 (276), 3–47.

Hein, E. (2017), 'The Bhaduri–Marglin post-Kaleckian model in the history of distribution and growth theories: an assessment by means of model closures', *Review of Keynesian Economics*, 5 (2), 218–38.

Hein, E., M. Lavoie and T. van Treeck (2011), 'Some instability puzzles in Kaleckian models of growth and distribution: a critical survey', *Cambridge Journal of Economics*, 35 (3), 587–612.

Hein, E., M. Lavoie and T. van Treeck (2012), 'Harrodian instability and the "normal rate" of capacity utilization in Kaleckian models of distribution and growth – a survey', *Metroeconomica*, 63 (1), 139–69.

Hein, E. and M. Mundt (2013), 'Financialization, the financial and economic crisis, and the requirements and potentials for wage-led recovery', in M. Lavoie and E. Stockhammer (eds), *Wage-Led Growth: An Equitable Strategy for Economic Recovery*, Basingstoke, UK: Palgrave Macmillan, pp. 153–86.

Hein, E. and E. Stockhammer (2011a), 'A post-Keynesian macroeconomic model of inflation, distribution and employment', in E. Hein and E. Stockhammer (eds), *A Modern Guide to Keynesian Macroeconomics and Economic Policies*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 112–36.

Hein, E. and E. Stockhammer (eds) (2011b), *A Modern Guide to Keynesian Macroeconomics and Economic Policies*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Hein, E. and L. Vogel (2008), 'Distribution and growth reconsidered: empirical results for six OECD countries', *Cambridge Journal of Economics*, 32 (3), 479–511.

Hicks, J.R. (1950), *A Contribution to the Theory of the Trade Cycle*, Oxford, UK: Clarendon Press.

Hicks, J.R. (1965), *Capital and Growth*, Oxford, UK: Clarendon Press.

Holling, C.S. (1973), 'Resilience and stability of ecological systems', *Annual Review of Ecology and Systematics*, 4 (1), 1–23.

Houthakker, H.S. and S.P. Magee (1969), 'Income and price elasticities in world trade', *Review of Economics and Statistics*, 51 (2), 111–25.

Ibarra, C.A. (2011), 'A note on intermediate imports and the BPCG model in Mexico', *Economic Change and Restructuring*, 44 (4), 357–68.

Ibarra, C.A. (2015a), 'Comentario a "La ley de Thirlwall: una lectura crítica"', *Investigación Económica*, 74 (292), 41–5.

Ibarra, C.A. (2015b), 'Investment and the real exchange rate's profitability channel in Mexico', *International Review of Applied Economics*, 29 (5), 716–39.

Ibarra, C.A. and R.A. Blecker (2016), 'Structural change, the real exchange rate and the balance of payments in Mexico, 1960–2012', *Cambridge Journal of Economics*, 40 (2), 507–39.

Ibarra, C.A. and J. Ros (2018), 'Profitability and capital accumulation in Mexico: a first look at tradables and non-tradables based on KLEMS', *International Review of Applied Economics*, published online, 17 October, https://doi.org/10.1080/02692171.2018.1511691.

Isaac, A.G. (1990), 'Monetary accommodation of income claims: a comment', *Weltwirtshäftliches Archiv*, 126 (1), 173–7.

Isaac, A.G. (1991), 'Economic stabilization and money supply endogeneity in a conflicting claims model', *Journal of Post Keynesian Economics*, 14 (1), 93–110.

Isaac, A.G. (2009), 'Monetary and fiscal interactions: short-run and long-run implications', *Metroeconomica*, 60 (1), 197–223.

Isaac, A.G. and Y.K. Kim (2013), 'Consumer and corporate debt: a neo-Kaleckian synthesis', *Metroeconomica*, 24 (2), 244–71.

Johnson, H. (1958), *International Trade and Economic Growth: Studies in Pure Theory*, London, UK: Allen and Unwin.

Jones, C.I. (1995), 'R&D-based models of economic growth', *Journal of Political Economy*, 103 (4), 759–84.

Jones, C.I. (2002), *Introduction to Economic Growth*, 2nd edition, New York, NY, USA: W.W. Norton.

Jones, C.I. (2018), *Macroeconomics*, New York, NY, USA: W.W. Norton.

Kaldor, N. (1956), 'Alternative theories of distribution', *Review of Economic Studies*, 23 (2), 83–100.

Kaldor, N. (1957), 'A model of economic growth', *Economic Journal*, 67 (268), 591–624.

Kaldor, N. (1961), 'Capital accumulation and economic growth', in F. Lutz (ed.), *The Theory of Capital*, London, UK: Macmillan, pp. 177–222, reprinted in F. Targetti and A.P. Thirlwall (eds) (1989), *The Essential Kaldor*, New York, NY, USA: Holmes & Meier, pp. 229–81.

Kaldor, N. (1966a), *Causes of the Slow Rate of Economic Growth in the United Kingdom*, Cambridge, UK: Cambridge University Press, reprinted in F. Targetti and A.P. Thirlwall (eds) (1989), *The Essential Kaldor*, New York: Holmes & Meier, pp. 282–310.

Kaldor, N. (1966b), 'Marginal productivity and the macroeconomic theories of distribution', *Review of Economic Studies*, 33 (4), 309–19.

Kaldor, N. (1970), 'The case for regional policies', *Scottish Journal of Political Economy*, 17 (3), 337–48.

Kaldor, N. (1971), 'Conflicts in national economic objectives', *Economic Journal*, 81 (321), 1–16.

Kaldor, N. (1972), 'The irrelevance of equilibrium economics', *Economic Journal*, 82 (328), 1237–55.

Kaldor, N. (1978), 'The effect of devaluations on trade in manufactures', in *Further Essays on Applied Economics*, London, UK: Duckworth, pp. 99–118.

Kaldor, N. (1981), 'The role of increasing returns, technical progress and cumulative causation in the theory of international trade and economic growth', *Economie Appliquée*, 34 (4), 593–617.

Kaldor, N. (1985), *Economics without Equilibrium*, Cardiff, UK: University College Cardiff Press.

Kaldor, N. (1996), *Causes of Growth and Stagnation in the World Economy*, Cambridge, UK: Cambridge University Press.

Kalecki, M. (1937), 'The principle of increasing risk', *Economica*, 4 (16), 440–47.

Kalecki, M. (1943), 'Political aspects of full employment', *Political Quarterly*, 14 (4), 322–31.

Kalecki, M. (1954), *Theory of Economic Dynamics*, London, UK: Unwin, reprinted New York, NY, USA: Monthly Review Press, 1968.

Kalecki, M. (1968), 'Trend and business cycle reconsidered', *Economic Journal*, 78 (310), 263–76.

Kalecki, M. (1971a), 'Class struggle and the distribution of national income', *Kyklos*, 24 (1), 1–9.

Kalecki, M. (1971b), *Selected Essays on the Dynamics of the Capitalist Economy, 1933–1970*, Cambridge, UK: Cambridge University Press.

Kalecki, M. (1990), *Collected Works of Michal Kalecki*, vol. I, *Capitalism: Business Cycles and Full Employment*, ed. J. Osiatynski, trans. C.A. Kisiel, Oxford, UK: Oxford University Press.

Kalecki, M. (1991), *Collected Works of Michal Kalecki*, vol. II, *Capitalism: Economic Dynamics*, ed. J. Osiatynski, trans. C.A. Kisiel, Oxford, UK: Oxford University Press.

Kapeller, J. and B. Schütz (2015), 'Conspicuous consumption, inequality and debt: the nature of consumption-driven profit-led regimes', *Metroeconomica*, 66 (1), 51–70.

Karabarbounis, L. and B. Neiman (2014), 'The global decline of the labor share', *Quarterly Journal of Economics*, 129 (1), 61–103.

Katz, L.F. and L.H. Summers (1989), 'Industry rents: evidence and implications', *Brookings Papers: Microeconomics*, 1989, 209–75.

Keynes, J.M. (1930), *A Treatise on Money*, London, UK: Macmillan.

Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money*, London, UK: Macmillan.

Kiefer, D. and C. Rada (2015), 'Profit maximising goes global: the race to the bottom', *Cambridge Journal of Economics*, 39 (5), 1333–50.

Kim, Y.K. (2013), 'Household debt, financialization, and macroeconomic performance in the United States, 1951–2009', *Journal of Post Keynesian Economics*, 35 (4), 675–94.

King, J. (2010), 'Kaldor and the Kaldorians', in M. Setterfield (ed.), *Handbook of Alternative Theories of Economic Growth*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 157–72.

Klein, M.W. (2002), *Mathematical Methods for Economics*, 2nd edition, Boston, MA, USA: Pearson/Addison-Wesley.

Kregel, J.A. (1980), 'Economic dynamics in the theory of steady growth: an historical essay on Harrod's knife edge', *History of Political Economy*, 12 (1), 97–123.

Kregel, J.A. (1985), 'Hamlet without the prince: Cambridge macroeconomics without money', *American Economic Review*, 75 (2), 133–139.

Kriesler, P. (1988), 'Kalecki's pricing theory revisited', *Journal of Post Keynesian Economics*, 11 (1), 108–30.

Krueger, A.B. and L.H. Summers (1988), 'Efficiency wages and the inter-industry wage structure', *Econometrica*, 56 (2), 259–93.

Krugman, P.R. (1979), 'Increasing returns, monopolistic competition, and international trade', *Journal of International Economics*, 9 (4), 469–79.

Krugman, P.R. (1980), 'Scale economies, product differentiation, and the pattern of trade', *American Economic Review*, 70 (5), 950–59.

Krugman, P.R. (1989), 'Differences in income elasticities and trends in real exchange rates', *European Economic Review*, 33 (5), 1031–54.

Krugman, P.R. and L. Taylor (1978), 'Contractionary effects of devaluation', *Journal of International Economics*, 8 (3), 445–56.

Kurz, H.D. (1990), 'Effective demand, employment and capital utilisation in the short run', *Cambridge Journal of Economics*, 14 (2), 205–17.

Kurz, H.D. (1994), 'Growth and distribution', *Review of Political Economy*, 6 (4), 393–420.

Kurz, H.D. (ed.) (2008), *Critical Essays on Piero Sraffa's Legacy in Economics*, Cambridge, UK: Cambridge University Press.

Kurz, H.D. and N. Salvadori (eds) (2003), *The Legacy of Piero Sraffa*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Kuznets, S. (1955), 'Economic growth and income inequality', *American Economic Review*, 45 (1), 1–28.

Lamfalussy, A. (1961), *Investment and Growth in Mature Economies*, New York, NY, USA: Macmillan.

Lamfalussy, A. (1963), *The United Kingdom and the Six: An Essay on Economic Growth in Western Europe*, Homewood, IL, USA: Irwin.

Lavoie, M. (1992), *Foundations of Post-Keynesian Economic Analysis*, Aldershot, UK and Brookfield, VT, USA: Edward Elgar Publishing.

Lavoie, M. (1995a), 'Interest rates in post-Keynesian models of growth and distribution', *Metroeconomica*, 46 (2), 146–77.

Lavoie, M. (1995b), 'The Kaleckian model of growth and distribution and its neo-Ricardian and neo-Marxian critiques', *Cambridge Journal of Economics*, 19 (6), 789–818.

Lavoie, M. (1996), 'Traverse, hysteresis and normal rates of capacity utilization in Kaleckian models of growth and distribution', *Review of Radical Political Economics*, 28 (4), 113–47.

Lavoie, M. (2002), 'The Kaleckian growth model with target return pricing and conflict inflation', in M. Setterfield (ed.), *The Economics of Demand-Led Growth*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 172–88.

Lavoie, M. (2003), 'Kaleckian effective demand and Sraffian normal prices: towards a reconciliation', *Review of Political Economy*, 15 (1), 53–74.

Lavoie, M. (2010), 'Surveying short run and long run stability issues with the Kaleckian model of growth', in M. Setterfield (ed.), *Handbook of Alternative Theories of Economic Growth*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 132–56.

Lavoie, M. (2014), *Post-Keynesian Economics: New Foundations*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Lavoie, M. (2016), 'Convergence towards the normal rate of capacity utilization in neo-Kaleckian models: the role of non-capacity creating autonomous expenditures', *Metroeconomica*, 67 (1), 172–201.

Lavoie, M. (2017), 'The origins and evolution of the debate on wage-led and profit-led regimes', *European Journal of Economics and Economic Policies: Intervention*, 14 (2), 200–221.

Lavoie, M. and W. Godley (2001), 'Kaleckian models of growth in a coherent stock-flow monetary framework', *Journal of Post Keynesian Economics*, 24 (2), 277–311.

Lavoie, M. and P. Kriesler (2007), 'Capacity utilization, inflation and monetary policy: the Duménil and Lévy macro model and the new Keynesian consensus', *Review of Radical Political Economics*, 39 (4), 586–98.

Lavoie, M., G. Rodríguez and M. Seccareccia (2004), 'Similitudes and discrepancies in post-Keynesian and Marxist theories of investment: a theoretical and empirical investigation', *International Review of Applied Economics*, 18 (2), 127–49.

Lavoie, M. and E. Stockhammer (eds) (2013), *Wage-Led Growth: An Equitable Strategy for Economic Recovery*, Basingstoke, UK: Palgrave Macmillan.

Lawrence, R.Z. (1990), 'US current account adjustment: an appraisal', *Brookings Papers on Economic Activity*, 1990 (2), 343–89.

Lee, F.S. (1998), *Post Keynesian Price Theory*, Cambridge, UK: Cambridge University Press.

Lee, K. (2013), *Schumpeterian Analysis of Economic Catch-Up: Knowledge, Path-Creation and the Middle Income Trap*, New York, NY, USA: Cambridge University Press.

León-Ledesma, M. (2002), 'Accumulation, innovation and catching-up: an extended cumulative growth model', *Cambridge Journal of Economics*, 26 (2), 201–18.

León-Ledesma, M. and M. Lanzafame (2010), 'The endogenous nature of the "natural" rate of growth', M. Setterfield (ed.), *Handbook of Alternative Theories of Economic Growth*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 208–19.

León-Ledesma, M. and A.P. Thirlwall (2000), 'Is the natural rate of growth exogenous', *Banca Nazionale Del Lavoro Quarterly Review*, 53 (215), 433–45.

León-Ledesma, M. and A.P. Thirlwall, (2002), 'The endogeneity of the natural rate of growth', *Cambridge Journal of Economics*, 26 (4), 441–59.

Lerner, A.P. (1943), 'Functional finance and the federal debt', *Social Research*, 10 (1), 38–51.

Lewis, W.A. (1954), 'Economic development with unlimited supplies of labour', *The Manchester School*, 22 (2), 139–91.

Lima, G.T. (2010), 'Profit sharing, capacity utilization and growth in a Post Keynesian macromodel', in M. Setterfield (ed.), *Handbook of Alternative Theories of Economic Growth*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 344–59.

Lima, G.T. and V.R. Carvalho (2008), 'Macrodinâmica do produto sob restrição externa: a experiência brasilera no período 1930–2004', *Economia Aplicada*, 12 (1), 55–77.

Lima, G.T. and A.J.A. Meirelles (2007), 'Macrodynamics of debt regimes, financial instability and growth', *Cambridge Journal of Economics*, 31 (4), 563–80.

Lima, G.T. and M. Setterfield (2014), 'The cost channel of monetary transmission and stabilization policy in a Post-Keynesian macrodynamic model', *Review of Political Economy*, 26 (2), 258–81.

Lima, G.T. and M. Setterfield (2016), 'Expectations and stability in the Kaleckian growth model', *Brazilian Keynesian Review*, 2 (1), 11–25.

López, J., A. Sanchez and A. Spanos (2011), 'Macroeconomic linkages in Mexico', *Metroeconomica*, 62 (2), 356–85.

Lucas, R.E. (1988), 'On the mechanics of economic development', *Journal of Monetary Economics*, 22 (1), 3–42.

Maddison, A. (1991), *Dynamic Forces in Capitalist Development*, Oxford, UK: Oxford University Press.

Maddison, A. (2008), 'Historical statistics of the world economy: 1–2008 AD', accessed 1 August 2017 at www.ggdc.net/maddison/Maddison.htm.

Magacho, G.R. and J.S.L. McCombie (2017), 'Verdoorn's law and productivity dynamics: an empirical investigation into the demand and supply approaches', *Journal of Post Keynesian Economics*, 40 (4), 600–621.

Malthus, T.R. (1798), *An Essay on the Principle of Population*, reprinted in G. Gilbert (ed.) (1993), Oxford, UK: Oxford World's Classics.

Malthus, T.R. (1820), *Principles of Political Economy with a View to Their Practical Application*, London, UK: Murray, reprinted in P. Sraffa (ed.) (1951), *The Works and Correspondence of David Ricardo*, vol. II, *Notes on Malthus's Principles of Political Economy*, Cambridge, UK: Cambridge University Press.

Maneschi, A. (1983), 'Dynamic aspects of Ricardo's international trade theory', *Oxford Economic Papers*, 35 (1), 67–80.

Mankiw, G., D. Romer and D. Weil (1992), 'A contribution to the empirics of economic growth', *Quarterly Journal of Economics*, 107 (2), 407–37.

Marconi, N., E.C. Araújo and J.L. Oreiro (2016), 'The exchange rate, income elasticities, and structural change: theoretical foundations and empirical evidence', Instituto de Economia, Universidad Federal de Rio de Janeiro (UFRJ), Discussion Paper 012/2016.

Marglin, S.A. (1984a), 'Growth, distribution and inflation: a centennial synthesis', *Cambridge Journal of Economics*, 8 (2), 115–44.

Marglin, S.A. (1984b), *Growth, Distribution, and Prices*, Cambridge, MA, USA: Harvard University Press.

Marglin, S.A. (2017), 'Wages, prices, and employment in a Keynesian long run', *Review of Keynesian Economics*, 5 (3), 360–425.

Marglin, S.A. and A. Bhaduri (1990), 'Profit squeeze and Keynesian theory', in S.A. Marglin and J.B. Schor (eds), *The Golden Age of Capitalism*, Oxford, UK: Oxford University Press, pp. 153–86.

Marglin, S.A. and J.B. Schor (eds) (1990), *The Golden Age of Capitalism*, New York, NY, USA: Oxford University Press.

Martin, P. and C.A. Rogers (1997), 'Stabilization policy, learning by doing and economic growth', *Oxford Review of Economic Policy*, 49 (2), 152–66.

Marx, K. (1867), *Capital: A Critique of Political Economy, vol. I,* trans. B. Fowkes (1976), London, UK: Penguin/Vintage.

Marx, K. (1894), *Capital: A Critique of Political Economy, vol. III,* trans. David Fernbach (1981), London, UK: Pelican/Penguin.

Marx, K. and F. Engels (1968), *Selected Works in One Volume*, New York, NY, USA: International Publishers.

McCombie, J.S.L. (1981), 'Are international growth rates constrained by the balance of payments? A comment on Professor Thirlwall', *Banca Nazionale del Lavoro Quarterly Review*, 34 (139), 455–8.

McCombie, J.S.L. (1985), 'Economic growth, the Harrod foreign trade multiplier and the Hicks super-multiplier', *Applied Economics*, 17 (1), 55–72.

McCombie, J.S.L. (1989), '"Thirlwall's law" and balance of payments constrained growth – a comment on the debate', *Applied Economics*, 21 (5), 611–29.

McCombie, J.S.L. (1993), 'Economic growth, trade interlinkages, and the balance of payments constraint', *Journal of Post Keynesian Economics*, 15 (4), 471–505.

McCombie, J.S.L. (2011), 'Criticisms and defences of the balance-of-payments constrained growth model: some old, some new', *PSL Quarterly Review*, 64 (259), 353–92.

McCombie, J.S.L., M. Pugno and B. Soro (eds) (2003), *Productivity Growth and Economic Performance: Essays on Verdoorn's Law*, Basingstoke, UK: Palgrave Macmillan.

McCombie, J.S.L. and A.P. Thirlwall (1994), *Economic Growth and the Balance-of-Payments Constraint*, Basingstoke, UK: Macmillan.

McCombie, J.S.L. and A.P. Thirlwall (1997), 'Economic growth and the balance-of-payments constraint revisited', in P. Arestis, G. Palma and M. Sawyer (eds), *Markets, Unemployment and Economic Policy: Essays in Honour of Geoff Harcourt*, vol. 2, Abingdon, UK: Routledge, pp. 498–511.

McCombie, J.S.L. and A.P. Thirlwall (1999), 'Growth in an international context: a post Keynesian view', in J. Deprez and J.T. Harvey (eds), *Foundations of International Economics: Post Keynesian Perspectives*, New York, NY, USA: Routledge, pp. 35–90.

McCombie, J.S.L. and A.P. Thirlwall (eds) (2004), *Essays on Balance of Payments Constrained Growth: Theory and Evidence*, Abingdon, UK: Routledge.

McGregor, P.G. and J.K. Swales (1985), 'Professor Thirlwall and balance of payments constrained growth', *Applied Economics*, 17 (1), 17–32.

Melitz, M.J. (2003), 'The impact of trade on intra-industry reallocations and aggregate industry productivity', *Econometrica*, 71 (6), 1695–725.

Minsky, H.P. (1959), 'A linear model of cyclical growth', *Review of Economics and Statistics*, 41 (2), 133–45.

Minsky, H.P. (1975), *John Maynard Keynes*, New York, NY, USA: Columbia University Press.

Minsky, H.P. (1982), *Can 'It' Happen Again?* Armonk, NY, USA: M.E. Sharpe.

Minsky, H.P. (1986), *Stabilizing an Unstable Economy*, New Haven, CT, USA: Yale University Press.

Mishel, L. and J. Schieder (2018), *CEO Compensation Surged in 2017*, Washington, DC, USA: Economic Policy Institute, 16 August.

Mohun, S. (2006), 'Distributive shares in the US economy, 1964–2001', *Cambridge Journal of Economics*, 30 (3), 347–70.

Moore, B.J. (1988), *Horizontalists and Verticalists: The Macroeconomics of Credit Money*, Cambridge, UK: Cambridge University Press.

Moreno-Brid, J.C. (1998), 'Balance-of-payments constrained economic growth: the case of Mexico', *Banca Nazionale del Lavoro Quarterly Review*, 51 (207), 413–33.

Moreno-Brid, J.C. (1998–99), 'On capital flows and the balance-of-payments-constrained growth model', *Journal of Post Keynesian Economics*, 21 (2), 283–98.

Moreno-Brid, J.C. (1999), 'Mexico's economic growth and the balance of payments constraint: a cointegration analysis', *International Review of Applied Economics*, 13 (2), 149–59.

Morishima, M. (1973), *Marx's Economics: A Dual Theory of Value and Growth*, Cambridge, UK: Cambridge University Press.

Mott, T. and E. Slattery (1994a), 'Tax incidence and macroeconomic effects in a Kaleckian model when profits finance affects investment and prices may respond to taxes', *Journal of Post Keynesian Economics*, 16 (3), 391–409.

Mott, T. and E. Slattery (1994b), 'The influence of changes in income distribution on aggregate demand in a Kaleckian model: stagnation vs. exhilaration reconsidered', in P. Davidson and J. Kregel (eds), *Employment, Growth, and Finance*, Aldershot, UK and Brookfield, VT, USA: Edward Elgar Publishing, pp. 69–82.

Myrdal, G. (1957), *Economic Theory and Underdeveloped Regions*, London, UK: Duckworth.

Naastepad, C.W.M. and S. Storm (2006), 'OECD demand regimes, 1960–2000', *Journal of Post Keynesian Economics*, 29 (2), 213–48.

Nell, E.J. (1985), 'Jean Baptiste Marglin: a comment on "Growth, distribution and inflation"', *Cambridge Journal of Economics*, 9 (2), 173–8.

Nikiforos, M. (2013), 'The (normal) rate of capacity utilization at the firm level', *Metroeconomica*, 64 (3), 513–38.

Nikiforos, M. (2016a), 'Distribution-led growth in the long run', *Review of Keynesian Economics*, 4 (4), 391–408.

Nikiforos, M. (2016b), 'On the 'utilisation controversy': a theoretical and empirical discussion of the Kaleckian model of growth and distribution', *Cambridge Journal of Economics*, 40 (2), 437–67.

Nikiforos, M. (2017), 'Uncertainty and contradiction: an essay on the business cycle', *Review of Radical Political Economics*, 49 (2), 247–64.

Nikiforos, M. (2018), 'Some comments on the Sraffian supermultiplier approach to growth and distribution', *Journal of Post Keynesian Economics*, 41 (4), 659–75.

Nikiforos, M. and D.K. Foley (2012), 'Distribution and capacity utilization: conceptual issues and empirical evidence', *Metroeconomica*, 63 (1), 200–229.

Nikiforos, M. and G. Zezza (2017), 'Stock-flow consistent macroeconomic models: a survey', *Journal of Economic Surveys*, 31 (5), 1204–39.

Nikolaidi, M. and E. Stockhammer (2017), 'Minsky models: a structured survey', *Journal of Economic Surveys*, 31 (5), 1304–31.

Noyola Vázquez, J. (1956), 'El desarrollo económico y la inflación en México y otros países latinoamericanos', *Investigación Económica*, 16 (4), 603–48.

Ohno, T. (2014), 'Entry deterrence in the neo-Kaleckian model: optimizing foundations', *Economic Review*, 65 (4), 289–302.

Okishio, N. (1961), 'Technical changes and the rate of profit', *Kobe University Economic Review*, 7, 85–99.

Onaran, Ö., L. Andersen, G. Cozzi, S. Dahl, T. Nissen, T. Obst and D. Tori (2017), 'An investment and equality-led sustainable development strategy for Europe', Foundation for European Progressive Studies (FEPS), February.

Onaran, Ö. and G. Galanis (2012), 'Is aggregate demand wage-led or profit-led? National and global effects', Conditions of Work and Employment Series No. 40, Geneva: International Labour Office.

Onaran, Ö. and G. Galanis (2013), 'Is aggregate demand wage-led or profit-led? A global model', in M. Lavoie and E. Stockhammer (eds), *Wage-Led Growth: An Equitable Strategy for Economic Recovery*, Basingstoke, UK: Palgrave Macmillan, pp. 71–99.

Onaran, Ö. and T. Obst (2016), 'Wage-led growth in the EU15 member states: the effects of income distribution on growth, investment, trade balance, and inflation', *Cambridge Journal of Economics*, 40 (6), 1517–51.

Onaran, Ö., E. Stockhammer and L. Grafl (2011), 'The finance-dominated growth regime, distribution, and aggregate demand in the US', *Cambridge Journal of Economics*, 35 (4), 637–61.

Oreiro, J.L. (2016), 'Inconsistency and over-determination in balance-of-payments-constrained growth models: a note', *Review of Keynesian Economics*, 4 (2), 193–200.

Pacheco-López, P. (2005), 'The impact of trade liberalization on exports, imports, the balance of trade, and growth: the case of Mexico', *Journal of Post Keynesian Economics*, 27 (4), 595–619.

Pacheco-López, P. and A.P. Thirlwall (2004), 'Trade liberalisation in Mexico: rhetoric and reality', *Banca Nazionale del Lavoro Quarterly Review*, 57 (229), 141–67.

Palley, T.I. (1994), 'Debt, aggregated demand, and the business cycle: a model in the spirit of Kaldor and Minsky', *Journal of Post Keynesian Economics*, 16 (3), 371–90.

Palley, T.I. (1996a), 'Aggregate demand in a reconstruction of growth theory: the macro foundations of economic growth', *Review of Political Economy*, 8 (1), 23–36.

Palley, T.I. (1996b), 'Growth theory in a Keynesian mode: some Keynesian foundations for the theory of economic growth', *Journal of Post Keynesian Economics*, 19 (1), 113–35.

Palley, T.I. (1996c), *Post Keynesian Economics: Debt, Distribution and the Macroeconomy*, New York, NY, USA: St. Martin's Press.

Palley, T.I. (1997), 'Aggregate demand and endogenous growth: a generalized Keynes-Kaldor model of economic growth', *Metroeconomica*, 48 (2), 161–76.

Palley, T.I. (2002a), 'Economic contradictions coming home to roost? Does the US economy face a long-term aggregate demand generation problem?', *Journal of Post Keynesian Economics*, 25 (1), 9–32.

Palley, T.I. (2002b), 'Keynesian macroeconomics and the theory of economic growth: putting aggregate demand back in the picture', in M. Setterfield (ed.), *The Economics of Demand-Led Growth: Challenging the Supply-Side Vision of the Long Run*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 19–40.

Palley, T.I. (2002c), 'Pitfalls in the theory of growth: an application to the balance-of-payments-constrained growth model', in M. Setterfield (ed.), *The Economics of Demand-Led Growth: Challenging the Supply-Side Vision of the Long Run*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 115–25.

Palley, T.I. (2009), 'Imports and the income-expenditure model: implications for fiscal policy and recession fighting', *Journal of Post Keynesian Economics*, 32 (2), 311–22.

Palley, T.I. (2013a), 'Cambridge and neo-Kaleckian growth and distribution theory: comparison with an application to fiscal policy', *Review of Keynesian Economics*, 1 (1), 79–104.

Palley, T.I. (2013b), 'Horizontalists, verticalists, and structuralists: the theory of endogenous money reassessed', *Review of Keynesian Economics*, 1 (4), 406–24.

Palley, T.I. (2014), 'Enriching the neo-Kaleckian growth model: nonlinearities, political economy, and *q* theory', in S. Dullien, E. Hein and A. Truger (eds), *Macroeconomics, Development and Economic Policies: Festschrift for Jan Priewe*, Marburg, Germany: Metropolis-Verlag, pp. 69–82.

Palley, T.I. (2017), 'Wage- vs. profit-led growth: the role of the distribution of wages in determining regime character', *Cambridge Journal of Economics*, 41 (1), 49–61.

Palumbo, A. (2009), 'Adjusting theory to reality: the role of aggregate demand in Kaldor's late contributions on economic growth', *Review of Political Economy*, 21 (3), 341–68.

Panico, C., A. Pinto and M.P. Anyul (2012), 'Income distribution and the size of the financial sector: a Sraffian analysis', *Cambridge Journal of Economics*, 36 (6), 1455–77.

Pariboni, R. (2016), 'Household consumer debt, endogenous money and growth: a supermultiplier-based analysis', *PSL Quarterly Review*, 69 (278), 211–33.

Park, M.S. (2000), 'Autonomous demand and warranted rate of growth', *Contributions to Political Economy*, 19, 1–18.

Pasinetti, L.L. (1960), 'A mathematical formulation of the Ricardian system', *Review of Economic Studies*, 27 (2), 78–98.

Pasinetti, L.L. (1962), 'Rate of profit and income distribution in relation to the rate of economic growth', *Review of Economic Studies*, 29 (4), 267–79.

Pasinetti, L.L. (1974), *Growth and Income Distribution: Essays in Economic Theory*, Cambridge, UK: Cambridge University Press.

Pasinetti, L.L. (1977), *Lectures in the Theory of Production*, New York, NY, USA: Columbia University Press.

Pasinetti, L.L. (1981), *Structural Change and Economic Growth: A Theoretical Essay on the Dynamics of the Wealth of Nations*, Cambridge, UK: Cambridge University Press.

Pasinetti, L.L. (1993), *Structural Economic Dynamics: A Theory of the Economic Consequences of Human Learning*, Cambridge, UK: Cambridge University Press.

Pérez Caldentey, E. (2015), 'Una lectura crítica de la "lectura crítica" de la Ley de Thirlwall', *Investigación Económica*, 74 (292), 47–65.

Pérez Caldentey, E. (2018), 'Porqué importa el enfoque estructural de la inflación', manuscript, UN Economic Commission for Latin America and the Caribbean (CEPAL), Santiago, Chile.

Pérez Caldentey, E. and J.C. Moreno-Brid (1999), 'Terms of trade, exports and economic growth in Central America: a long-term view', *Banca Nazionale del Lavoro Quarterly Review*, 52 (211), 431–49.

Perraton, J. (2003), 'Balance of payments constrained growth and developing countries: an examination of Thirlwall's hypothesis', *International Review of Applied Economics*, 17 (1), 1–22.

Perrotini-Hernández, I. and J.A. Vázquez-Muñoz (2017), 'Endogenous growth and economic capacity: theory and empirical evidence for the NAFTA countries', *PSL Quarterly Review*, 70 (282), 247–82.

Phillips, A.W. (1958), 'The relation between unemployment and the rate of change of money wage rates in the United Kingdom, 1861–1957', *Economica*, 25 (100), 283–99.

Pigou, A.C. (1933), *Theory of Unemployment*, London, UK: Macmillan.

Piketty, T. (2014), *Capital in the Twenty-First Century*, trans. A. Goldhammer, Cambridge, MA, USA: Harvard University Press.

Pollin, R. (2005), *Contours of Descent: U.S. Economic Fractures and the Landscape of Global Austerity*, revised and updated edition, London, UK: Verso.

Porcile, G. and G.T. Lima (2010), 'Real exchange rate and the elasticity of labour supply in a balance-of-payments constrained macrodynamics', *Cambridge Journal of Economics*, 34 (6), 1019–39.

Porcile, G. and D. Spinola (2018), 'Natural, effective and BOP-constrained rates of growth: adjustment mechanisms and closure equations', *PSL Quarterly Review*, 71 (285), 139–60.

Prebisch, R. (1950), *The Economic Development of Latin America and Its Principal Problems*, Lake Success, NY, USA: United Nations, Department of Economic Affairs.

Pugno, M. (1998), 'The stability of Thirlwall's model of economic growth and the balance-of-payments constraint', *Journal of Post Keynesian Economics*, 20 (4), 559–81.

Rada, C. (2007), 'Stagnation or transformation of a dual economy through endogenous productivity growth', *Cambridge Journal of Economics*, 31 (5), 711–40.

Ramsey, F.P. (1928), 'A mathematical theory of saving', *Economic Journal*, 38 (152), 543–59.

Ranis, G. and J.C.H. Fei (1961), 'A theory of economic development', *American Economic Review*, 51 (4), 533–65.

Rapetti, M., P. Skott and A. Razmi (2012), 'The real exchange rate and economic growth: are developing countries different?', *International Review of Applied Economics*, 26 (6), 735–53.

Razmi, A. (2005), 'Balance-of-payments-constrained growth model: the case of India', *Journal of Post Keynesian Economics*, 27 (4), 655–87.

Razmi, A. (2011), 'Exploring the robustness of the balance of payments-constrained growth idea in a multiple good framework', *Cambridge Journal of Economics*, 35 (3), 545–67.

Razmi, A. (2016a), 'Correctly analysing the balance-of-payments constraint on growth', *Cambridge Journal of Economics*, 40 (6), 1581–1608.

Razmi, A. (2016b), 'Growth and distribution in low-income economies: modifying post-Keynesian analysis in light of theory and history', *Review of Keynesian Economics*, 4 (4), 429–49.

Razmi, A. (2018), 'Is planet Earth as a whole likely to be wage-led?', *Review of Keynesian Economics*, 6 (3), 289–306.

Reifschneider, D., W.L. Wascher and D. Wilcox (2015), 'Aggregate supply in the United States: recent developments and implications for the conduct of monetary policy', *IMF Economic Review*, 63 (1), 71–109.

Riach, P.A. (1971), 'Kalecki's "degree of monopoly" reconsidered', *Australian Economic Papers*, 10 (16), 50–60.

Ribeiro, R.S.M., J.S.L. McCombie and G.T. Lima (2016), 'Exchange rate, income distribution and technical change in a balance-of-payments constrained growth model', *Review of Political Economy*, 28 (4), 545–65.

Ribeiro, R.S.M., J.S.L. McCombie and G.T. Lima (2017a), 'A reconciliation proposal of demand-driven growth models in open economies', *Journal of Economic Studies*, 44 (2), 226–44.

Ribeiro, R.S.M., J.S.L. McCombie and G.T. Lima (2017b), 'Some unpleasant currency-devaluation arithmetic in a post Keynesian macromodel', *Journal of Post Keynesian Economics*, 40 (2), 145–67.

Ribeiro, R.S.M., J.S.L. McCombie and G.T. Lima (2018), 'Does real exchange rate undervaluation really promote economic growth', working paper, Universidade Federal de Minas Gerais (Brazil), Faculdade de Ciências Econômicas, Centro de Desenvolvimento e Planejamento Regional.

Ricardo, D. (1821), *Principles of Political Economy and Taxation*, 3rd edition, reprinted in P. Sraffa (ed.) (1951), *The Works and Correspondence of David Ricardo*, vol. I, Cambridge, UK: Cambridge University Press.

Ricoy, C.J. (1987), 'Cumulative causation', in J. Eatwell, M. Milgate and P. Newman (eds), *The New Palgrave: A Dictionary of Economics*, London, UK: Macmillan.

Roberts, M. (2007), 'The conditional convergence properties of simple Kaldorian growth models', *International Review of Applied Economics*, 21 (5), 619–32.

Roberts, M. and M. Setterfield (2007), 'What is endogenous growth theory?', in P. Arestis, M. Baddeley and J.S.L. McCombie (eds), *Economic Growth: New Directions in Theory and Policy*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 14–31.

Robinson, J. (1946), 'The pure theory of international trade', 14 (2), 98–122, reprinted in J. Robinson (1951), *Collected Economic Papers of Joan Robinson*, vol. I, Oxford, UK: Basil Blackwell, pp. 182–205.

Robinson, J. (1947), *Essays in the Theory of Employment*, 2nd edition, Oxford, UK: Basil Blackwell.

Robinson, J. (1956), *The Accumulation of Capital*, London, UK: Macmillan.

Robinson, J. (1962), *Essays in the Theory of Economic Growth*, London, UK: Macmillan.

Rochon, L.-P. (1999), *Credit, Money and Production: An Alternative Post-Keynesian Approach*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Rochon, L.-P. (2005), 'Robinson on credit, money and finance', in B. Gibson (ed.), *Joan Robinson's Economics: A Centennial Celebration*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 267–81.

Rodríguez, O. (1977), 'On the conception of the centre–periphery system', *CEPAL Review*, No. 3, 195–239.

Rodrik, D. (1997), *Has Globalization Gone Too Far?*, Washington, DC, USA: Institute for International Economics.

Rodrik, D. (2008), 'The real exchange rate and economic growth', *Brookings Papers on Economic Activity*, 2008 (2), 365–412.

Rodrik, D. (2013), 'Unconditional convergence in manufacturing', *Quarterly Journal of Economics*, 128 (1), 165–204.

Rodrik, D. (2014), 'The past, present, and future of economic growth', in F. Allen, J.R. Behrman, N. Birdsall, S. Fardoust, D. Rodrik, A. Steer and A. Subramanian (eds), *Towards a Better Global Economy: Policy Implications for Citizens Worldwide in the 21st Century*, Oxford, UK: Oxford University Press, pp. 70–119.

Rodrik, D. (2016), 'Premature deindustrialization', *Journal of Economic Growth*, 21 (1), 1–33.

Rodrik, D., M. McMillan and C. Sepúlveda (2016), 'Structural change, fundamentals, and growth', in M. McMillan, D. Rodrik, and C. Sepúlveda (eds), *Structural Change, Fundamentals, and Growth: A Framework and Case Studies*, Washington, DC, USA: International Food Policy Research Institute, pp. 1–38.

Roemer, J. (1981), *Analytical Foundations of Marxian Economic Theory*, Cambridge, UK: Cambridge University Press.

Rogers, C. (2018a), 'The new macroeconomics has no clothes', *Review of Keynesian Economics*, 6 (1), 22–33.

Rogers, C. (2018b), 'The conceptual flaw in the microeconomic foundations of dynamic stochastic general equilibrium models', *Review of Political Economy*, 30 (1), 72–83.

Romer, P.M. (1986), 'Increasing returns and long run growth', *Journal of Political Economy*, 94 (5), 1002–37.

Romer, P.M. (1990), 'Endogenous technological change', *Journal of Political Economy*, 98 (5), S71–S102.

Roncaglia, A. (2009), *Piero Sraffa*, Basingstoke, UK: Palgrave Macmillan.

Ros, J. (2013), *Rethinking Economic Development, Growth, and Institutions*, Oxford, UK: Oxford University Press.

Ros, J. (2016), 'Can growth be wage-led in small open developing economies?', *Review of Keynesian Economics*, 4 (4), 450–57.

Rosenberg, N. (1976), *Perspectives on Technology*, Cambridge, UK: Cambridge University Press.

Rowthorn, R.E. (1977), 'Conflict, inflation and money', *Cambridge Journal of Economics*, 1 (3), 213–39.

Rowthorn, R.E. (1981), 'Demand, real wages and economic growth', *Thames Papers in Political Economy*, Autumn, 1–39.

Ryoo, S. (2013), 'Bank profitability, leverage and financial instability: a Minsky–Harrod model', *Cambridge Journal of Economics*, 37 (5), 1127–60.

Ryoo, S. (2016), 'Demand-driven inequality, endogenous saving rate and macroeconomic instability', *Cambridge Journal of Economics*, 40 (1), 201–25.

Ryoo, S. and P. Skott (2017), 'Fiscal and monetary policy rules in an unstable economy', *Metroeconomica*, 68 (3), 500–548.

Saez, E. and G. Zucman (2016), 'Wealth inequality in the United States since 1913: evidence from capitalized income tax data', *Quarterly Journal of Economics*, 131 (2), 519–78.

Samuelson, P.A. (1939a), 'A synthesis of the principle of acceleration and the multiplier', *Journal of Political Economy*, 47 (6), 786–97.

Samuelson, P.A. (1939b), 'Interactions between the multiplier analysis and the principle of acceleration', *Review of Economics and Statistics*, 21 (2), 75–8.

Samuelson, P.A. and F. Modigliani (1966), 'The Pasinetti paradox in neoclassical and more general models', *Review of Economic Studies*, 33 (4), 269–301.
Santana, B.M. and J.L. Oreiro (2018), 'Real exchange rate and structural change in a Kaldorian balance of payments constrained growth model', *Revista de Economia Política*, 38 (1), 48–69.
Santos-Paulino, A. and A.P. Thirlwall (2004), 'The impact of trade liberalisation on exports, imports and the balance of payments of developing countries', *Economic Journal*, 114 (493), F50–F72.
Say, J.B. (1803), *A Treatise on Political Economy*, reprinted New York, NY, USA: Augustus M. Kelley Publishers, 1971.
Schmookler, J. (1966), *Invention and Economic Growth*, Cambridge, MA, USA: Harvard University Press.
Schoder, C. (2012), 'Hysteresis in the Kaleckian growth model: a Bayesian analysis for the US manufacturing sector from 1984 to 2007', *Metroeconomica*, 63 (3), 542–68.
Schoder, C. (2013), 'Credit vs. demand constraints: the determinants of US firm-level investment over the business cycles from 1977 to 2011', *North American Journal of Economics and Finance*, 26 (December), 1–27.
Schoder, C. (2014), 'Effective demand, exogenous normal utilization and endogenous capacity in the long run: evidence from a cointegrated vector autoregression analysis for the USA', *Metroeconomica*, 65 (2), 298–320.
Scitovsky, T. (1956), 'Economies of scale, competition, and European integration', *American Economic Review*, 46 (1), 71–91.
Scott, M.F.G. (1989), *A New View of Economic Growth*, Oxford, UK: Clarendon.
Seguino, S. (2019), 'Engendering macroeconomic theory and policy', *Feminist Economics*, published online, 28 May, https://www.tandfonline.com/doi/full/10.1080/13545701.2019.1609691.
Seguino, S. and M. Setterfield (2010), 'Gender equality and the sustainability of steady state growth paths', in M. Setterfield (ed.), *Handbook of Alternative Theories of Economic Growth*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 360–75.
Sen, A.K. (1963), 'Neo-classical and neo-Keynesian theories of distribution', *Economic Record*, 39 (85), 53–64.
Sen, A.K. (1970), *Growth Economics*, Harmondsworth, UK: Penguin.
Serrano, F. (1995), 'Long period effective demand and the Sraffian supermultiplier', *Contributions to Political Economy*, 14, 67–90.
Serrano, F. and F. Freitas (2017), 'The Sraffian supermultiplier as an alternative closure for heterodox growth theory', *European Journal of Economics and Economic Policies: Intervention*, 14 (1), 70–91.
Setterfield, M. (1993), 'Towards a long-run theory of effective demand: modeling macroeconomic systems with hysteresis', *Journal of Post Keynesian Economics*, 15 (3), 347–64.
Setterfield, M. (1997a), 'History versus equilibrium and the theory of economic growth', *Cambridge Journal of Economics*, 21 (3), 365–78.
Setterfield, M. (1997b), *Rapid Growth and Relative Decline: Modelling Macroeconomic Dynamics with Hysteresis*, Basingstoke, UK: Macmillan.
Setterfield, M. (1997c), 'Should economists dispense with the notion of equilibrium?', *Journal of Post Keynesian Economics*, 20 (1), 47–78.
Setterfield, M. (2002), 'A model of Kaldorian traverse: cumulative causation, structural change and evolutionary hysteresis', in M. Setterfield (ed.), *The Economics of Demand-Led Growth: Challenging the Supply-Side Vision of the Long Run*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 215–33.
Setterfield, M. (2006a), 'Balancing the macroeconomic books on the backs of workers: a simple

analytical political economy model of contemporary US capitalism', *International Journal of Political Economy*, 35 (3), 46–63. (Erratum, *International Journal of Political Economy*, 37 (4), 2008–09, 104.)

Setterfield, M. (2006b), 'Thirlwall's law and Palley's pitfalls: a reconsideration', in P. Arestis, J. McCombie and R. Vickerman (eds), *Growth and Economic Development: Essays in Honour of A.P. Thirlwall*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishers, pp. 47–59.

Setterfield, M. (2007), 'The rise, decline and rise of incomes policies in the US during the post-war era: an institutional-analytical explanation of inflation and the functional distribution of income', *Journal of Institutional Economics*, 3 (2), 127–46.

Setterfield, M. (2009), 'Path dependency, hysteresis and macrodynamics', in P. Arestis and M. Sawyer (eds), *Path Dependency and Macroeconomics*, International Papers in Political Economy, Basingstoke, UK: Palgrave Macmillan, pp. 37–79.

Setterfield, M. (ed.) (2010), *Handbook of Alternative Theories of Economic Growth*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Setterfield, M. (2011), 'The remarkable durability of Thirlwall's law', *PSL Quarterly Review*, 64 (259), 393–427.

Setterfield, M. (2013a), 'Endogenous growth: a Kaldorian approach', in G.C. Harcourt and P. Kriesler (eds), *The Oxford Handbook of Post-Keynesian Economics*, vol. I, *Theory and Origins*, Oxford, UK: Oxford University Press, pp. 231–56.

Setterfield, M. (2013b), 'Wages, demand and US macroeconomic travails: diagnosis and prognosis', in S.M. Fazzari, B.Z. Cynamon and M. Setterfield (eds), *After the Great Recession: The Struggle for Economic Recovery and Growth*, Cambridge, UK: Cambridge University Press, pp. 158–84.

Setterfield, M. (2014), 'Neoclassical growth theory and heterodox growth theory: opportunities for (and obstacles to) greater engagement', *Eastern Economic Journal*, 40 (3), 365–86.

Setterfield, M. (2018), 'Long-run variation in capacity utilization in the presence of a fixed normal rate', *Cambridge Journal of Economics*, published online, 18 July, https://doi.org/10.1093/cje/bey016.

Setterfield, M. and J.D. Avritzer (2019), 'Hysteresis in the normal rate of capacity utilization: a behavioural explanation', New School for Social Research, Working Paper 1907.

Setterfield, M. and A. Budd (2011), 'A Keynes–Kalecki model of cyclical growth with agent-based features', in P. Arestis (ed.), *Microeconomics, Macroeconomics, and Economic Policy: Essays in Honour of Malcolm Sawyer*, Basingstoke, UK: Palgrave Macmillan, pp. 228–50.

Setterfield, M. and J. Cornwall (2002), 'A neo-Kaldorian perspective on the rise and decline of the golden age', in M. Setterfield (ed.), *The Economics of Demand-Led Growth: Challenging the Supply-Side Vision of the Long Run*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 67–82.

Setterfield, M. and Y.K. Kim (2016), 'Debt servicing, aggregate consumption, and growth', *Structural Change and Economic Dynamics*, 36 (March), 22–33.

Setterfield, M. and Y.K. Kim (2017), 'Household borrowing and the possibility of "consumption-driven, profit-led growth"', *Review of Keynesian Economics*, 5 (1), 43–60.

Setterfield, M., Y.K. Kim and J. Rees (2016), 'Inequality, debt servicing, and the sustainability of steady-state growth', *Review of Political Economy*, 28 (1), 45–63.

Setterfield, M. and T. Lovejoy (2006), 'Aspirations, bargaining power, and macroeconomic performance', *Journal of Post Keynesian Economics*, 29 (1), 117–48.

Setterfield, M. and S. Ozcelik (2018), 'Is the balance of payments constrained growth rate time-varying? Exchange rate over valuation, policy-induced recessions, deindustrialization, and long run growth', in P. Arestis (ed.), *Alternative Approaches in Macroeconomics: Essays in Honour of John McCombie*, Cham, Switzerland: Palgrave Macmillan, pp. 331–53.

Shaikh, A. (2007), 'A proposed synthesis of classical and Keynesian growth', SCEPA Working Paper WP 2007-1, The New School.

Shaikh, A. (2009), 'Economic policy in a growth context: a classical synthesis of Keynes and Harrod', *Metroeconomica*, 60 (3), 455–94.

Shaikh, A. (2016), *Capitalism: Competition, Conflict, Crises*, Oxford, UK: Oxford University Press.

Sherman, H.J. and G.R. Evans (1984), *Macroeconomics: Keynesian, Monetarist, and Marxist Views*, New York, NY, USA: Harper & Row.

Silva de Jesus, C., R.A. Araujo and C.E. Drumond (2018), 'An empirical test of the Post-Keynesian growth model applied to functional income distribution and the growth regime in Brazil', *International Review of Applied Economics*, 32 (4), 428–49.

Sinha, A. (2016), *A Revolution in Economic Theory: The Economics of Piero Sraffa*, Cham, Switzerland: Palgrave Macmillan.

Skolka, J. (1980), 'Außenhandel und Wirtschaftswachstum', *Wirtschaft und Gesellschaft*, 80 (1), 9–16.

Skott, P. (1989), *Conflict and Effective Demand in Economic Growth*, Cambridge, UK: Cambridge University Press.

Skott, P. (1999), 'Growth and stagnation in a two-sector model: Kaldor's Mattioli lectures', *Cambridge Journal of Economics*, 23 (3), 353–70.

Skott, P. (2010), 'Growth, instability, and cycles: Harrodian and Kaleckian models of accumulation and income distribution', in M. Setterfield (ed.), *Handbook of Alternative Theories of Economic Growth*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 108–31.

Skott, P. (2012), 'Theoretical and empirical shortcomings of the Kaleckian investment function', *Metroeconomica*, 63 (1), 109–38.

Skott, P. (2017a), 'Autonomous demand and the Harrodian criticisms of the Kaleckian model', *Metroeconomica*, 68 (1), 185–93.

Skott, P. (2017b), 'Autonomous demand, Harrodian instability and the supply side', *Metroeconomica*, published online, 14 July, https://onlinelibrary.wiley.com/doi/10.1111/meca.12181.

Skott, P. (2017c), 'Weaknesses of "wage-led growth"', *Review of Keynesian Economics*, 5 (3), 336–59.

Skott, P. and S. Ryoo (2008), 'Macroeconomic implications of financialisation', *Cambridge Journal of Economics*, 32 (6), 827–62.

Skott, P. and B. Zipperer (2012), 'An empirical evaluation of three post-Keynesian models', *European Journal of Economics and Economic Policies: Intervention*, 9 (2), 277–307.

Smith, A. (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations*, reprinted, Chicago, IL, USA: University of Chicago Press, 1976.

Snowdon, B. and H. Vane, (2005), *Modern Macroeconomics: Its Origins, Development and Current State*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Solow, R.M. (1956), 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, 70 (1), 65–94.

Solow, R.M. (1994), 'Perspectives on growth theory', *Journal of Economic Perspectives*, 8 (1), 45–54.

Solow, R.M. (2007), 'The last 50 years in growth theory and the next 10', *Oxford Review of Economic Policy*, 23 (1), 3–14.

Sraffa, P. (1960), *Production of Commodities by Means of Commodities: Prelude to a Critique of Economic Theory*, Cambridge, UK: Cambridge University Press.

Steedman, I. (1977), *Marx After Sraffa*, London, UK: New Left Books.

Steindl, J. (1952), *Maturity and Stagnation in American Capitalism*, Oxford, UK: Blackwell, reprinted New York, NY, USA: Monthly Review Press, 1976.

Steindl, J. (1979), 'Stagnation theory and stagnation policy', *Cambridge Journal of Economics*, 3 (1), 1–14.

Stockhammer, E. (2011), 'The macroeconomics of unemployment', in E. Hein and E. Stockhammer (eds), *A Modern Guide to Keynesian Macroeconomics and Economic Policies*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 137–64.

Stockhammer, E. (2013), 'Why have wage shares fallen? An analysis of the determinants of functional income distribution', in M. Lavoie and E. Stockhammer (eds), *Wage-Led Growth: An Equitable Strategy for Economic Recovery*, Basingstoke, UK: Palgrave Macmillan, pp. 40–70.

Stockhammer, E. (2017), 'Wage-led versus profit-led demand: what have we learned? A Kaleckian–Minskyan view', *Review of Keynesian Economics*, 5 (1), 25–42.

Stockhammer, E., E. Hein and L. Grafl (2011), 'Globalization and the effects of changes in the functional income distribution on aggregate demand in Germany', *International Review of Applied Economics*, 25 (1), 1–24.

Stockhammer, E. and J. Michell (2017), 'Pseudo-Goodwin cycles in a Minsky model', *Cambridge Journal of Economics*, 41 (1), 105–25.

Stockhammer, E., Ö. Onaran and S. Ederer (2009), 'Functional income distribution and aggregate demand in the Euro area', *Cambridge Journal of Economics*, 33 (1), 139–59.

Stockhammer, E., J. Rabinovich and N. Reddy (2017), 'Distribution, wealth and demand regimes in historical perspective: USA, UK, France and Germany, 1855–2010', Kingston University London, Economics Discussion Papers 2017-5, November.

Stockhammer, E. and R. Stehrer (2011), 'Goodwin or Kalecki in demand? Functional income distribution and aggregate demand in the short run', *Review of Radical Political Economics*, 43 (4), 506–22.

Stockhammer, E. and R. Wildauer (2016), 'Debt-driven growth? Wealth, distribution and demand in OECD countries', *Cambridge Journal of Economics*, 40 (6), 1609–34.

Storm, S. and C.W.M. Naastepad (2012), *Macroeconomics Beyond the NAIRU*, Cambridge, MA, USA: Harvard University Press.

Storm, S. and C.W.M. Naastepad (2017), 'Bhaduri–Marglin meet Kaldor–Marx: wages, productivity and investment', *Review of Keynesian Economics*, 5 (1), 4–24.

Summers, L.H. (1988), 'Tax policy and international competitiveness', in J.A. Frenkel (ed.), *International Aspects of Fiscal Policies*, Chicago, IL, USA: University of Chicago Press.

Summers, L.H. (2014), 'U.S. economic prospects: secular stagnation, hysteresis, and the zero lower bound', *Business Economics*, 49 (2), 65–73.

Sunkel, O. (1960), 'Inflation in Chile: an unorthodox approach', *International Economic Papers*, 10, 107–31.

Swan, T.W. (1956), 'Economic growth and capital accumulation', *Economic Record*, 32 (2), 334–61.

Sweezy, P.M. (1942), *The Theory of Capitalist Development: Principles of Marxian Political Economy*, New York, NY, USA: Oxford University Press.

Sweezy, P.M. (1981), *Four Lectures on Marxism*, New York, NY, USA: Monthly Review Press.

Szirmai, A. (2012), 'Industrialisation as an engine of growth in developing countries, 1950–2005', *Structural Change and Economic Dynamics*, 23 (4), 406–20.

Tavani, D., P. Flaschel and L. Taylor (2011), 'Estimated non-linearities and multiple equilibria in a model of distributive-demand cycles', *International Review of Applied Economics*, 25 (5), 519–38.

Tavani, D. and R. Vasudevan (2014), 'Capitalists, workers, and managers: wage inequality and effective demand', *Structural Change and Economic Dynamics*, 30 (September), 120–31.

Tavani, D. and L. Zamparelli (2017), 'Government spending composition, aggregate demand, growth, and distribution', *Review of Keynesian Economics*, 5 (2), 239–58.

Taylor, A.M. and M.P. Taylor (2004), 'The purchasing power parity debate', *Journal of Economic Perspectives*, 18 (4), 135–58.

Taylor, L. (1979), *Macro Models for Developing Countries*, New York, NY, USA: McGraw-Hill.

Taylor, L. (1981), 'South-north trade and southern growth: bleak prospects from a structuralist point of view', *Journal of International Economics*, 11 (4), 589–602.

Taylor, L. (1983), *Structuralist Macroeconomics: Applicable Models for the Third World*, New York, NY, USA: Basic Books.

Taylor, L. (1985), 'A stagnationist model of economic growth', *Cambridge Journal of Economics*, 9 (4), 383–403.

Taylor, L. (1990), 'Real and money wages, output and inflation in the semi-industrialized world', *Economica*, 57 (227), 329–53.

Taylor, L. (1991), *Income Distribution, Inflation, and Growth: Lectures on Structuralist Macroeconomic Theory*, Cambridge, MA, USA: MIT Press.

Taylor, L. (2004), *Reconstructing Macroeconomics: Structuralist Proposals and Critiques of the Mainstream*, Cambridge, MA, USA: Harvard University Press.

Thirlwall, A.P. (1979), 'The balance of payments constraint as an explanation of international growth rate differences', *Banca Nazionale del Lavoro Quarterly Review*, 32 (128), 45–53.

Thirlwall, A.P. (1981), 'A reply to Mr. McCombie', *Banca Nazionale del Lavoro Quarterly Review*, 34 (139), 458–9.

Thirlwall, A.P. (1983), 'A plain man's guide to Kaldor's growth laws', *Journal of Post Keynesian Economics*, 5 (3), 345–58.

Thirlwall, A.P. (1991), 'Professor Krugman's 45-degree rule', *Journal of Post Keynesian Economics*, 14 (1), 23–8.

Thirlwall, A.P. (2002), *The Nature of Economic Growth*: *An Alternative Framework for Understanding the Performance of Nations*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Thirlwall, A.P. (2011), 'Balance of payments constrained growth models: history and overview', *PSL Quarterly Review*, 64 (259), 307–51.

Thirlwall, A.P. (2013), *Economic Growth in an Open Developing Economy: The Role of Structure and Demand*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Thirlwall, A.P. and R.J. Dixon (1979), 'A model of export-led growth with a balance of payments constraint', in J.K. Bowers (ed.), *Inflation, Development and Integration: Essays in Honour of A. J. Brown*, Leeds, UK: Leeds University Press.

Thirlwall, A.P. and M.N. Hussain (1982), 'The balance of payments constraint, capital flows and growth rate differences between developing countries', *Oxford Economic Papers*, 34 (3), 498–510.

Todaro, M.P. and S.C. Smith (2012), *Economic Development*, 11th edition, Boston, MA, USA: Addison-Wesley/Pearson.

Toner, P. (1999), *Main Currents in Cumulative Causation*, Basingstoke, UK: Macmillan.

Trezzini, A. (1995), 'Capacity utilisation in the long run and the autonomous components of aggregate demand', *Contributions to Political Economy*, 14, 33–66.

Trezzini, A. (1998), 'Capacity utilisation in the long run: some further considerations', *Contributions to Political Economy*, 17, 53–67.

Tu, P.N.V. (1988), 'Towards an optimal wildlife management', Discussion paper series No. 111, University of Calgary, cited in Barbosa-Filho and Taylor (2006).

US Bureau of Economic Analysis (2016), *National Income and Product Accounts (NIPAs), Table 4.2.3*, accessed 1 July 2016 at www.bea.gov.

Vargas Sánchez, G. and A. Luna (2014), 'Slow growth in the Mexican economy', *Journal of Post Keynesian Economics*, 37 (1), 115–33.

Vasudevan, R. (2017), 'Finance and distribution', *Review of Keynesian Economics*, 5 (1), 78–93.

Veblen, T. (1912), *The Theory of the Leisure Class: An Economic Study of Institutions*, London, UK: Macmillan.

Veneziani, R. and S. Mohun (2006), 'Structural stability and Goodwin's growth cycle', *Structural Change and Economic Dynamics*, 17 (4), 437–51.

Verdoorn, P.J. (1949), 'Fattori che regolano lo sviluppo della produttività del lavoro', *L'Industria*, 1 (March), 3–10.

Vernengo, M. (2006), 'Money and inflation', in P. Arestis and M. Sawyer (eds), *A Handbook of Alternative Monetary Economics*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing, pp. 471–89.

Vernengo, M. (2015), 'Una lectura crítica de la crítica al modelo de Thirlwall', *Investigación Económica*, 74 (292), 67–80.

Verspagen, B. (1993), *Uneven Growth between Interdependent Economies: An Evolutionary View on Technology Gaps, Trade and Growth*, Aldershot, UK and Brookfield, VT, USA: Edward Elgar Publishing.

Verspagen, B. and K. Wakelin (1997), 'Technology, employment and trade: perspectives on European integration', in J. Fagerberg, P. Hansson, L. Lundberg and A. Melchior (eds), *Technology and International Trade*, Aldershot, UK and Brookfield, VT, USA: Edward Elgar Publishing, pp. 56–74.

von Arnim, R. (2011), 'Wage policy in an open-economy Kalecki–Kaldor model: a simulation study', *Metroeconomica*, 62 (2), 235–64.

von Arnim, R., D. Tavani and L. Carvalho (2014), 'Redistribution in a neo-Kaleckian two-country model', *Metroeconomica*, 65 (3), 430–59.

Wade, R. (1990), *Governing the Market: Economic Theory and the Role of Government in East Asian Industrialization*, Princeton, NJ, USA: Princeton University Press.

Weisskopf, T. (1979), 'Marxian crisis theory and the rate of profit in the postwar U.S. economy', *Cambridge Journal of Economics*, 3 (4), 341–78.

Wicksell, K. (1898), *Interest and Prices: A Study of the Causes Regulating the Value of Money*, trans. R.F. Kahn (1936), London, UK: Macmillan.

Wisman, J.D. (2013), 'Wage stagnation, rising inequality and the financial crisis of 2008', *Cambridge Journal of Economics*, 37 (4), 921–45.

Wolff, E.N. (2003), 'What's behind the rise in profitability in the US in the 1980s and 1990s?', *Cambridge Journal of Economics*, 27 (4), 479–99.

Wray, L.R. (1990), *Money and Credit in Capitalist Economies: The Endogenous Money Approach*, Aldershot, UK and Brookfield, VT, USA: Edward Elgar Publishing.

Wray, L.R. (1998), *Understanding Modern Money: The Key to Full Employment and Price Stability*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar Publishing.

Wray, L.R. (2012), 'Imbalances? What imbalances? A dissenting view', Levy Economics Institute of Bard College, Working Paper No. 704.

Young, A.A. (1928), 'Increasing returns and economic progress', *Economic Journal*, 38 (152), 527–42.

# Index