ROUTLEDGE EXPLORATIONS IN ECONOMIC HISTORY

Large Databases in Economic History

Research methods and case studies

Edited by
Mark Casson and
Nigar Hashimzade

ROUTLEDGE

Large Databases in Economic History

'Big data' is now readily available to economic historians, thanks to the digitisation of primary sources, collaborative research linking different data sets and the publication of databases on the internet. Key economic indicators, such as the consumer price index, can be tracked over long periods, and qualitative information, such as land use, can be converted to a quantitative form. In order to fully exploit these innovations it is necessary to use sophisticated statistical techniques to reveal the patterns hidden in data sets, and this book shows how this can be done.

A distinguished group of economic historians have teamed up with younger researchers to pilot the application of new techniques to 'big data'. Topics addressed in this volume include prices and the standard of living, money supply, credit markets, land values and land use, transport, technological innovation and business networks. The research spans the medieval, early modern and modern periods. Research methods include simultaneous equation systems, stochastic trends and discrete choice modelling.

This book is essential reading for doctoral and post-doctoral researchers in business, economic and social history. The case studies will also appeal to historical geographers and applied econometricians.

Mark Casson is Professor of Economics at the University of Reading, UK. His research focuses on economic history, business history and econometrics.

Nigar Hashimzade is Professor of Economics at the University of Durham, UK. Her research focuses on time series econometrics and public economics.

Routledge explorations in economic history

Edited by Lars Magnusson

Uppsala University, Sweden

1 **Economic Ideas and Government Policy**
Contributions to contemporary economic history
Sir Alec Cairncross

2 **The Organization of Labour Markets**
Modernity, culture and governance in Germany, Sweden, Britain and Japan
Bo Stråth

3 **Currency Convertibility**
The gold standard and beyond
Edited by Jorge Braga de Macedo, Barry Eichengreen and Jaime Reis

4 **Britain's Place in the World**
A historical enquiry into import controls 1945–1960
Alan S. Milward and George Brennan

5 **France and the International Economy**
From Vichy to the Treaty of Rome
Frances M.B. Lynch

6 **Monetary Standards and Exchange Rates**
M.C. Marcuzzo, L. Officer and A. Rosselli

7 **Production Efficiency in Domesday England, 1086**
John McDonald

8 **Free Trade and its Reception 1815–1960**
Freedom and trade: volume I
Edited by Andrew Marrison

9 **Conceiving Companies**
Joint-stock politics in Victorian England
Timothy L. Alborn

10 **The British Industrial Decline Reconsidered**
Edited by Jean-Pierre Dormois and Michael Dintenfass

11 **The Conservatives and Industrial Efficiency, 1951–1964**
Thirteen wasted years?
Nick Tiratsoo and Jim Tomlinson

12 **Pacific Centuries**
Pacific and Pacific Rim economic history since the sixteenth century
Edited by Dennis O. Flynn, Lionel Frost and A.J.H. Latham

13 **The Premodern Chinese Economy**
Structural equilibrium and capitalist sterility
Gang Deng

14 The Role of Banks in Monitoring Firms
The case of the Crédit Mobilier
Elisabeth Paulet

15 Management of the National Debt in the United Kingdom, 1900–1932
Jeremy Wormell

16 An Economic History of Sweden
Lars Magnusson

17 Freedom and Growth
The rise of states and markets in Europe, 1300–1750
S.R. Epstein

18 The Mediterranean Response to Globalization before 1950
Sevket Pamuk and Jeffrey G. Williamson

19 Production and Consumption in English Households 1600–1750
Mark Overton, Jane Whittle, Darron Dean and Andrew Hann

20 Governance, the State, Regulation and Industrial Relations
Ian Clark

21 Early Modern Capitalism
Economic and social change in Europe 1400–1800
Edited by Maarten Prak

22 An Economic History of London, 1800–1914
Michael Ball and David Sunderland

23 The Origins of National Financial Systems
Alexander Gerschenkron reconsidered
Edited by Douglas J. Forsyth and Daniel Verdier

24 The Russian Revolutionary Economy, 1890–1940
Ideas, debates and alternatives
Vincent Barnett

25 Land Rights, Ethno Nationality and Sovereignty in History
Edited by Stanley L. Engerman and Jacob Metzer

26 An Economic History of Film
Edited by John Sedgwick and Mike Pokorny

27 The Foreign Exchange Market of London
Development since 1900
John Atkin

28 Rethinking Economic Change in India
Labour and livelihood
Tirthankar Roy

29 The Mechanics of Modernity in Europe and East Asia
The institutional origins of social change and stagnation
Erik Ringmar

30 International Economic Integration in Historical Perspective
Dennis M.P. McCarthy

31 Theories of International Trade
Adam Klug
Edited by Warren Young and Michael Bordo

32 Classical Trade Protectionism 1815–1914
Edited by Jean Pierre Dormois and Pedro Lains

33 **Economy and Economics of Ancient Greece**
Takeshi Amemiya

34 **Social Capital, Trust and the Industrial Revolution**
1780–1880
David Sunderland

35 **Pricing Theory, Financing of International Organisations and Monetary History**
Lawrence H. Officer

36 **Political Competition and Economic Regulation**
Edited by Peter Bernholz and Roland Vaubel

37 **Industrial Development in Postwar Japan**
Hirohisa Kohama

38 **Reflections on the Cliometrics Revolution**
Conversations with economic historians
Edited by John S. Lyons, Louis P. Cain, and Samuel H. Williamson

39 **Agriculture and Economic Development in Europe since 1870**
Edited by Pedro Lains and Vicente Pinilla

40 **Quantitative Economic History**
The good of counting
Edited by Joshua Rosenbloom

41 **A History of Macroeconomic Policy in the United States**
John H. Wood

42 **An Economic History of the American Steel Industry**
Robert P. Rogers

43 **Ireland and the Industrial Revolution**
The impact of the Industrial Revolution on Irish industry and society, 1801–1922
Andy Bielenberg

44 **Intra-Asian Trade and Industrialization**
Essays in memory of Yasukichi Yasuba
Edited by A.J.H. Latham and Heita Kawakatsu

45 **Nation, State and the Industrial Revolution**
The visible hand
Lars Magnusson

46 **A Cultural History of Finance**
Irene Finel-Honigman

47 **Managing Crises and De-globalisation**
Nordic foreign trade and exchange 1919–1939
Edited by Sven-Olof Olsson

48 **The International Tin Cartel**
John Hillman

49 **The South Sea Bubble**
Helen J. Paul

50 **Ideas and Economic Crises in Britain from Attlee to Blair (1945–2005)**
Matthias Matthijs

51 **Bengal Industries and the British Industrial Revolution (1757–1857)**
Indrajit Ray

52 **The Evolving Structure of the East Asian Economic System since 1700**
Edited by A.J.H. Latham and Heita Kawakatsu

53 **German Immigration and Servitude in America, 1709–1920**
Farley Grubb

54 **The Rise of Planning in Industrial America, 1865–1914**
Richard Adelstein

55 **An Economic History of Modern Sweden**
Lennart Schön

56 **The Standard of Living and Revolutions in Russia, 1700–1917**
Boris Mironov

57 **Europe's Green Revolution and Others Since**
The rise and fall of peasant-friendly plant breeding
Jonathan Harwood

58 **Economic Analysis of Institutional Change in Ancient Greece**
Carl Hampus-Lyttkens

59 **Labour-Intensive Industrialization in Global History**
Edited by Gareth Austin and Kaoru Sugihara

60 **The History of Bankruptcy**
Economic, social and cultural implications in early modern Europe
Edited by Thomas Max Safley

61 **The Political Economy of Disaster and Underdevelopment**
Destitution, plunder and earthquake in Haiti
Mats Lundahl

62 **Nationalism and Economic Development in Modern Eurasia**
Carl Mosk

63 **Agricultural Transformation in a Global History Perspective**
Edited by Ellen Hillbom and Patrick Svensson

64 **Colonial Exploitation and Economic Development**
The Belgian Congo and the Netherlands Indies compared
Edited by Ewout Frankema and Frans Buelens

65 **The State and Business in the Major Powers**
An economic history 1815–1939
Robert Millward

66 **Privatization and Transition in Russia in the Early 1990s**
Carol Scott Leonard and David Pitt-Watson

67 **Large Databases in Economic History**
Research methods and case studies
Edited by Mark Casson and Nigar Hashimzade

This page intentionally left blank

Large Databases in Economic History

Research methods and case studies

Edited by Mark Casson and Nigar Hashimzade

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2013
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

and by Routledge
711 Third Avenue, New York, NY 10017

Routledge is an imprint of the Taylor & Francis Group, an informa business

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging in Publication Data
Large databases in economic history: research methods and case studies / edited by Mark Casson, Nigar Hashimzade.
pages cm. – (Routledge explorations in economic history; 67)
Includes bibliographical references and index.
1. Economic indicators–Statistical methods. 2. Big data–Statistical methods. I. Casson, Mark, 1945- II. Hashimzade, Nigar.
HD75.L367 2013
330.01′171–dc23
2013019994

ISBN: 978-0-415-82068-4 (hbk)
ISBN: 978-1-315-86742-7 (ebk)

Typeset in Times New Roman
by Wearset Ltd, Boldon, Tyne and Wear

Contents

Figures

Tables

Contributors

Editors

Mark Casson is Professor of Economics at the University of Reading. His previous books include *Enterprise and Competitiveness* (1990), *Economics of Business Culture* (1991), *Information and Organization* (1997), *Oxford Handbook of Entrepreneurship* (co-editor, 2006) and *The World's First Railway System* (2009). He has published in a range of journals relevant to this book, including *Econometrica*, *Review of Economic Studies*, *Economic Journal*, *Business History*, *Business History Review* and *Scandinavian Economic History Review*. He was Leverhulme Major Research Fellow in the Economics of Social Networks, 2006–9 and President of the Association of Business Historians, 2007–9.

Nigar Hashimzade is Professor of Economics at the University of Durham. She researches in econometrics and public economics, and has published in *Econometric Theory*, *Journal of Time Series Analysis*, *Computational Statistics and Data Analysis*, *Macroeconomic Dynamics* and other leading journals. She is co-editor of the *Handbook of Research Methods and Applications in Empirical Macroeconomics* (2013).

Contributors

Adrian R. Bell is Professor of the History of Finance and Director of the International Capital Markets Association Centre, University of Reading. His research focuses on medieval finance, the life of the medieval soldier and the economics of professional football. He has directed several funded projects involving the construction and use of large databases; his most recent book is the *English Wool Market,* c.*1230–1327* (with Chris Brooks and Paul Dryburgh, 2007). He is currently joint principal investigator on a Leverhulme Trust-funded project on medieval foreign exchange.

Chris Brooks is Professor of Finance and Director of Research at the International Capital Markets Association Centre, University of Reading. His research interests include statistical issues in risk management, econometric modelling in finance and the historical evolution of financial markets. He has published in the *Economic Journal*, *Journal of Banking and Finance*, *Journal of Empirical Finance* and many other journals. He is Associate Editor of the *International Journal of Forecasting* and was a member of the RAE 2008 Accounting and Finance sub-panel. He is currently joint principal investigator on a Leverhulme Trust-funded project on medieval foreign exchange.

Anna Campbell is an Associate of the Graduate Centre for Medieval Studies at the University of Reading. With degrees in history and ecclesiastical history from Oxford and a doctorate from Reading, she is currently researching French medieval history, with special reference to St Colette of Corbie. Her forthcoming publications include chapters in A. Van Oosterwijk (ed.), *Staging the Court of Burgundy* and Joan Mueller and Nancy Bradley Warren (eds), *A Companion to Colette of Corbie*.

Catherine Casson is Research Fellow at the Winton Institute of Monetary History, Ashmolean Museum, University of Oxford and Teaching Fellow in the School of History and Cultures, University of Birmingham. She is co-author of *The Entrepreneur in History: From Medieval Merchant to Modern Business Leader* (2013) and co-editor of *The History of Entrepreneurship: Innovation and Risk-taking, 1200–2000* (2013). She is currently researching property rents and civic reputation in medieval English towns. Her work on long-run prices has developed from her doctoral research on market regulation, published in *Urban History* (2012).

Janet Casson is a DPhil student at the University of Oxford, researching women's ownership of land in Victorian England under the supervision of Jane Humphries. Following a career as a biology teacher and economic consultant, she obtained an MSc in English Local History in 2007. She is Honorary Treasurer of the Business Enterprise Heritage Trust.

John Haggerty is Lecturer in Information Systems Security at the University of Salford. He has published in a wide range of journals, including *Explorations in Economic History*, *Enterprise and Society*, *International Journal of Digital Crime and Forensics* and *Computer Law and Security Review*.

Sheryllynne Haggerty is Associate Professor of History at the University of Nottingham, specialising in eighteenth-century trade. She is the author of *The British-Atlantic Trading Community, 1760–1810: Men, Women, and the Distribution of Goods* (2006) and *'Merely for Money?': Business Culture in the British Atlantic, 1750–1815* (2012). She has published in *Explorations in Economic History*, *Business History*, *Atlantic Studies*, *International Journal of Maritime History* and other leading journals.

Sara Horrell is Senior Lecturer in the Faculty of Economics and a Fellow of Murray Edwards College, University of Cambridge. She has published on work, poverty, the standard of living and family history in the *Economic*

History Review, *Explorations in Economic History*, *Journal of Development Studies* and other leading journals.

Jane Humphries is Professor of Economic History at the University of Oxford and a Fellow of All Souls College. Her recent book, *Childhood and Child Labour in the British Industrial Revolution* (2011) won the Giorgy Ranki Prize of the Economic History Association, and a BBC TV film documentary based on the book won the History Makers International Prize. She has published extensively on the interface between the family and the economy, in the *Economic History Review*, *Journal of Economic History*, *Explorations in Economic History* and other leading journals. She is President of the Economic History Society 2010–13.

Jane McCutchan is a PhD student at the University of Reading, supervised by Andrew Godley and Mark Casson. She holds a scholarship funded through the Museum of English Rural Life, and uses their collections in her research. She was previously project manager for *Steam: The Museum of the Great Western Railway*, based in Swindon.

Nick Mayhew is Professor of Numismatics and Monetary History at the University of Oxford, Deputy-Director (Collections) at the Ashmolean Museum and Director of the Winton Institute of Monetary History there. His publications include *Changing Values in Medieval Scotland: A Study of Prices, Money, and Weights and Measures* (with Elizabeth Gemmill, 1995) and *Sterling: The Rise and Fall of a Currency* (1999). His current research interests include the history of UK banking policy and the impact of monetary factors on long-run movements in property prices.

Tony K. Moore is Research Associate at the International Capital Markets Association Centre, University of Reading, and joint principal investigator on a Leverhulme Trust-funded project on medieval foreign exchange. He was previously a research associate on an ESRC-funded project on medieval finance and the wool trade. He has published in the *English Historical Review* and other leading journals.

A.E.M. Satchell is Research Associate at the Department of Geography and member of the Cambridge Group for the History of Population and Social Structure at the University of Cambridge. He is currently working on the *Occupational Structure of Britain,* c.*1379–1911* project, where his principal focus over recent years has been the creation of a long-run historical GIS of England's evolving transport infrastructure. He is a specialist in occupational structure, transport history, GIS and digital cartography.

Leigh Shaw-Taylor is Lecturer in Economic and Social History at the University of Cambridge and Deputy Director of the Cambridge Group for the History of Population and Social Structure. He is currently Director of the *Occupational Structure of Britain,* c.*1379–1911* project, which has received substantial funding from the ESRC, the Leverhulme Trust and the British

Academy. His publications include articles in the *Economic History Review*, *Journal of Economic History* and *Past and Present*.

Ken Sneath is Associate Director of Studies for Economics at Peterhouse College, Cambridge. He teaches both British economic history and local history at the University of Cambridge and his main research interest is the history of consumption. He is also a historical advisor to the Cambridge Community Heritage Project.

Professor Sir E.A. Wrigley is co-founder of the Cambridge Group for the History of Population and Social Structure and a Past-President of the British Academy. He is currently Deputy-Director of the *Occupational Structure of Britain,* c.*1379–1911* project. His recent publications include *English Population History from Family Reconstitution, 1580–1837* (1997), *Poverty, Progress and Population* (2004), *Energy and the English Industrial Revolution* (2010) and *The Early English Censuses* (2011).

Margaret Yates was Director of the Centre for Economic History at the University of Reading, 2011–13, and Senior Lecturer in Medieval History. Her recent publications include *Town and Countryside in West Berkshire, 1327–1600* (2007) and papers in the *Economic History Review* and *Agricultural History Review*. She has also contributed to the *Victorian County History of Oxfordshire*.

1 Introduction

Research methods for large databases

Mark Casson and Nigar Hashimzade

1.1 The research agenda

Three recent developments have opened up new opportunities for research in economic history: digitisation of primary sources, collaborative research linking different data sets and the publication of databases on the internet. Systematic exploitation of source materials now makes it possible to generate large representative samples, such as plots of land in a country or region, or comprehensive population data, such as the stock of ploughing engines in use at any given time. Collaborative research grants have funded the development of new long-run annual time series on prices, outputs, money supply, etc., for up to 750 years (1250–2000), and publication on the internet has widened access to such material. Panel data sets can be constructed that track the same sample over time, for example, parish population from Census data or attendance at meetings by members of an organisation. Linking data from different sources, such as railways, geology and population, widens the range of research questions that can be addressed.

New statistical methods have facilitated the discovery of hidden patterns in long-run data, involving the analysis of autocorrelation, regression to the mean, stochastic and deterministic trends, co-integrating relationships between co-evolving series, and structural breaks. In addition, panel estimation techniques have facilitated the synthesis of time-series and cross-section data. In analytical work, large data sets make asymptotic theory more relevant. Greater computing power and modern software packages make estimation of complex models using large data sets very quick.

These developments have made it possible to analyse very long-run processes that are important agents of economic change, including technological progress, market integration, political integration and institution-building. They can be analysed in a systematic way using structural models. Economic historians can now move from simple questions answered by descriptive statistics to more complex questions answered by estimated models, and from questions about a single variable (e.g. standard of living) to questions about the co-movement of variables (e.g. prices, population and money supply). Causation can be considered more explicitly using dynamic models, and multiple causation, as well as the direction of causation, can be analysed more fully.

The main challenges in utilising the advantages offered by these new developments are technical ones. For numerical analysis the historical data, especially the qualitative data, must be codified in an appropriate form. Historical data are prone to missing observations and data input errors, requiring considerable care to avoid mistakes in estimation. It is important to specify the right model and to interpret statistical results correctly, keeping in mind that the results may only be valid under certain explicit or implicit assumptions. The analysis should aim at finding all the patterns concealed within the data, ensuring that what is unexplained is, in some sense, truly random.

Instead of relying on generalisations from a range of specific studies, often carried out using different methodologies, it is possible to construct a single coherent account based on evidence of a consistent standard. The book reports the results of new research of this type.

The book is intended for use by doctoral and post-doctoral researchers in business history, economic history and social history. The case studies will also appeal to historical geographers and applied econometricians, and the techniques explained in the book are potentially useful to government policy-makers too. This book demonstrates how to create 'big data', and, above all, how to exploit it to the full. Many historians only 'scratch the surface' of the data they collect. There are often significant patterns hidden in their data that they fail to discover. This book shows how to unlock hidden patterns, and hence get more information out of the data.

Unlike conventional statistical texts, this book demonstrates how to put principles into practice with the aid of practical historical studies. This agenda leads, in some cases, to a reappraisal of conventional wisdom on such important issues as the development of the land market, the pricing of commodities, monetary instability, the economic impact of railways, the diffusion of steam technology and the role of women in the economy.

Many readers will be familiar with general statistical texts such as Wooldridge (2006). They will also be aware of 'cliometrics' literature, as summarised recently in Greasley and Oxley (2011). The case studies in this book build upon previous research in cliometrics. However, despite some similarities, the new research differs from earlier research in important respects.

Most cliometrics relies on single equation models and makes limited use of simultaneous equation models, stochastic trends and other concepts featured in the book. There is now more emphasis on qualitative evidence. Early cliometric research emphasised quantification (e.g. using heights as indicators of health and welfare) whereas much of the evidence in modern databases is qualitative. Recent research has tended to combine quantitative and qualitative evidence by using binary variables; this is particularly useful for testing institutional theories of economic change.

The book emphasises the importance of testing alternative theories rather than fitting models based on one specific theory. Cliometric research in the 1970s and 1980s tended to react against the Marxist turn in economic history during the 1960s by emphasising the ubiquitous and providential role of market forces.

Recent research has paid closer attention to the speed of market adjustment, and has suggested that market adjustment is often a rather sluggish process, and certainly not an instantaneous one. The research reported in this book takes no ideological position on the role of markets whatsoever. The philosophy is simply to 'let the data speak'; in practice, this means comparing the effectiveness of alternative models in explaining the patterns that are revealed by statistical analysis. Models therefore need to allow for variable speeds of adjustment in market processes.

1.2 Structure and content of the book

The structure of the book is as follows. Chapter 2 examines the covariation between the prices of eight widely traded commodities in England, 1250–1914. Chapter 3 presents new annual estimates for stocks of gold and silver coin, 1220–1750. These estimates are combined with price and output data to test the Quantity Theory of Money. Chapter 4 reviews the evidence on medieval international financial transactions, and shows how econometric studies of medieval finance can be used to identify structural breaks in economic behaviour. Chapter 5 analyses time series data on land and property values from feets of fines for two English counties, Essex and Warwickshire, 1300–1500. It demonstrates the changing uses of land and the differential movements in the values of various types of property, including agricultural land, mills and manorial rights. It also highlights the decline in smallholdings and the build-up of complex estates at the end of the fifteenth century. Chapter 6 shows how visual analytics can be used to summarise the structure of complex social networks. It presents a case study of Liverpool business networks at the end of the eighteenth century which demonstrates how a large binary data set of inter-personal relationships can be used to analyse the structure and dynamics of historical social networks. Chapter 7 develops and tests a model of the equilibrium distribution of population across towns and villages, using decadal data on local population units from the UK Census of Population 1801–1891 for the counties of Northamptonshire and Rutland. Chapter 8 develops and tests a theory of the role of women in land ownership, with special reference to nineteenth-century England, by modelling competition between men, women and institutions in the property market. For this purpose a database of 24,000 individual plots of land is created. Chapter 9 presents the first comprehensive database on the production and use of steam ploughing engines in English agriculture, 1859–1930. It examines the rise and decline of steam ploughing by analysing both spatial and temporal patterns in diffusion. Chapter 10 investigates changes in consumption patterns in the eighteenth- and nineteenth-century London by examining Old Bailey criminal records. Using a remarkable database linking individual burglars to the commodities they stole and the date of the theft, it shows how patterns of theft can inform recent historical debates over fashions and trends in consumer tastes.

The remainder of this chapter reviews the methodology that is common to all the following chapters. It presents the key concepts used in building economic models and in developing hypotheses regarding the long-run development of the economy. It outlines the statistical techniques that can be used to estimate the parameters of such models and to test hypotheses associated with them.

1.3 Fundamental concepts of data analysis in economic history

Observational nature of the data

A fundamental difficulty in establishing a relationship between economic variables is the observational nature of the economic data. Suppose that we want to know whether changes in price of wheat have an effect on the price of barley. One would expect there to be such an effect because these two commodities have similar uses. We cannot conduct an experiment, as we would do in natural sciences, by deliberately changing the wheat price and recording the response in the barley price. In practice, we have a set of observations on both and, furthermore, the recorded prices are likely to have been affected by numerous other factors. It is often impossible to isolate the effect of these factors, even when we happen to have the relevant data. Therefore, statistical methods must be adapted in order to extract reliable information from the historical data in the most efficient way.

Use and limitations of descriptive statistics

A common way of summarising the properties of historical data on economic variables is the use of descriptive statistics, such as the sample mean, as the measure of central tendency, and the sample variance (or the standard deviation), as the measure of dispersion, or the spread of observation about the sample mean. Other frequently used statistics are the range and the minimal and the maximal value in the sample. While these characteristics of the data are often helpful, in many situations they do not reflect certain patterns in the data that can be the most important for the research question. A typical situation is a trend in prices or seasonal fluctuations in trade volumes. Furthermore, the value of descriptive statistics can be driven entirely by one outlier and, therefore, give a poor picture of the bulk of the sample.

Joint properties of a set of several variables are often summarised using pairwise Pearson correlation coefficients. The Pearson correlation coefficient is a handy and intuitive indicator of the co-movement in variables, and a sufficiently large (positive or negative) value of the correlation between two variables, say, the price of wheat and the price of barley, may suggest an economic relationship that can be further investigated. It is, however, a measure of linear relationship between variables: it can be very small for a non-linear relationship (such as a U-shape, for example). Furthermore, its use with time series data is often problematic. For example, it is now well known that if each of the two variables exhibits a stochastic trend (accumulation of random shocks) the correlation coefficient in the sample is likely to be large and significant even when the two variables are generated by independent shocks and are, therefore, unrelated. To uncover important patterns in the data and explore links across several variables one has to go beyond descriptive statistics to more comprehensive methods of analysis.

Use of probability theory to formalise the process of inference

In this and subsequent chapters notations and terminology commonly accepted in the analysis of economic data are used. Here we give a brief summary and refer the reader to the standard texts, such as Greene (2012) or Wooldridge (2006) for more details.

It is typically assumed that the observed data were generated by an underlying process, the so-called data-generating process (DGP) that contains a deterministic component and a stochastic, or random, component. The random component is a random variable, that is, a variable whose value is not known prior to the realisation of some random event. A random variable is characterised by its probability distribution function: $F_Y(y)$ is the probability that random variable Y takes a value less than or equal to y, or $F_Y(y) = \Pr[Y \leq y]$, where Pr denotes probability. $F_Y(y)$ is sometimes referred to as the cumulative distribution function. For a discrete random variable that can take a countable number of discrete values the probability distribution function has discontinuities ('jumps'). The probability of Y taking a particular discrete value y_i is given by the probability mass function, $f_Y(y) = \Pr[Y = y]$. If the distribution function is continuous Y is said to be a continuous random variable, and it can take any value in a continuous range. The probability of a continuous random variable Y taking any particular value is zero, but we can define a probability that Y takes value in a certain range: $\Pr[y_1 < Y \leq y_2] = F_Y(y_2) - F_Y(y_1)$. If the distribution function is continuous and differentiable we can define a corresponding probability density function that is a derivative of the distribution function: $f_Y(y) = dF_Y(y)/dy$.

The joint distribution of two random variables, say, X and Y, is characterised by the joint probability distribution function, $F_{X,Y}(x,y) = \Pr[Y \leq y \text{ and } X \leq x]$. This generalizes to more than two variables. We are often interested in the distribution of a particular variable, say, Y, given the realisation of other variables in the data, say, $X_1 = x_1, \ldots, X_K = x_k$. This is characterised by the conditional distribution: $\mathrm{F}_{\mathrm{Y|X1,\ldots,XK}} = \Pr[Y \leq y | X_1 = x_1, \ldots, X_K = x_k]$.

It is often convenient, for compactness, to use vector notations. An ordered set of numbers, $\{y_1, \ldots, y_n\}$, is denoted by $\mathbf{y}$, and $y_1, \ldots, y_n$ are called the elements of vector $\mathbf{y}$. It is a column vector if the elements are arranged vertically into a column, and a row vector if the elements are arranged horizontally into a row. A transpose of a column vector is a row vector (the same elements arranged now into a row), and vice versa; the usual notation is $\mathbf{y}'$. For a row vector $\mathbf{u}$ and a column vector $\mathbf{v}$ of the same length the inner product (or a dot product) is defined as $\mathbf{uv} = u_1 v_1 + \ldots + u_n v_n$. The inner product of a column vector and its transpose is the sum of squared elements of this vector: $\mathbf{u'u} = u_1^2 + \ldots + u_n^2$.

Vectors of the same length can be arranged into a matrix. In general, an n-by-m matrix contains n rows and m columns. The elements of a matrix are labelled by their positions in the rows and columns: a_{ij} is an element of matrix **A** in the intersection of i-th row and j-th column.

Analysis of variance

Often the objects of interest are the sources of variation in the data. Analysis of variance, or ANOVA, is a statistical tool for decomposing the sample variance in a given variable according to different factors contributing to the variation. Suppose, for example, that we have data on the value per acre of N plots of land in each of k counties, and we wish to assess whether there are significant variations between counties in the value of agricultural land. We can use the values from the plots in the same county to calculate a county mean. To identify a general 'county effect' we need to find out whether the differences between k group means are statistically significant. One possibility is to perform pairwise t-tests for each of the $k(k-1)/2$ pairs of group means for the null hypothesis of the all the group means being the same against the alternative of the means being different. This, however, leads to a high probability of a Type I error (falsely rejecting a true null hypothesis): if α is the level of significance set for each pairwise t-test, with $k(k-1)/2$ independent tests the probability of falsely rejecting a true null hypothesis is $1-(1-\alpha)^{k(k-1)/2}$. This rises quickly with the number of groups: for example, for $\alpha=0.05$ and $k=4$ this probability is 0.265, whereas for $k=6$ it is 0.537. Alternatively, we can use ANOVA to test whether the difference in the means across several groups of data simultaneously is statistically significant. It has an advantage over the pairwise t-tests in that the comparison is simultaneous and so the probability of Type I error does not accumulate.

The null hypothesis in the ANOVA test is that all group means are equal, and equal to the grand mean, and the alternative hypothesis is that at least one group mean is different from the others. To perform the test an F-statistic is constructed using the variance within each group (around the group mean), σ_W^2, and the variance between groups (around the grand mean), σ_B^2. These are usually arranged in a table, as illustrated in Table 1.1.

Table 1.1 ANOVA for k independent groups

Source	*Sum of squares*	***DF***	*Mean squares*
Between	$S_B = \sum_{i=1}^{k} (x_i - x)^2$	$k-1$	$\sigma_B^2 = S_B/(k-1)$
Within	$S_B = \sum_{i=1}^{k} (n_i - 1)\sigma_i^2$	$N-k$	$\sigma_W^2 = S_W/(N-k)$
Total	$SS_B + SS_W$	$N-1$	

The null and the alternative hypotheses can be equivalently stated as the following:

$$H_0 : \sigma_B^2 \le \sigma_W^2 ; H_1 : \sigma_B^2 > \sigma_W^2$$

The F-statistic is the ratio of two variances, and under the null hypothesis it has an F-distribution with ($k-1$, $N-k$) degrees of freedom:

$$F = \frac{\sigma_B^2}{\sigma_W^2} \sim F(k-1, N-1)$$

The null is rejected in favour of the alternative if the value of the statistic computed from the sample exceeds the critical value of the distribution for the given level of significance. The critical values are tabulated in the standard statistics textbooks; these tables can also be easily found online. This approach can be extended to multiple factors, taking into account the interaction between factors (Lind *et al*., 2012, Chapter 12).

1.4 Importance of economic theory

A good practice in the analysis of economic data is to start with the relevant economic theory, in order to understand what patterns and relationships in the data we expect to find under different scenarios, before embarking on the empirical work. In a formal approach we can use an analytical model, formulated as an equation or a set of equations for the variables of interest, or we can use an intuitive approach by making an educated guess about the kind of relationship between the variables of interest, or the former can be complemented by the latter.

For example, in the study of commodity prices (see Chapter 2) we want to identify whether particular commodities (e.g. barley, oats, peas and wheat) were complements (in common use) or substitutes (in competing use), or whether there was a hierarchy of markets, where a particular commodity or a group of commodities was dominant in driving the prices of other commodities. An appropriate place to start is the basic economic model of perfectly competitive interrelated markets where prices equilibrate supply and demand for all commodities simultaneously. In equilibrium the prices of each commodity depend on the prices of the others. Because markets clear instantaneously, the model is static. It is plausible, however, to assume that the production decision (planting, hiring workers) must be made before the price is known, and that the price expectations are formed based on the previous prices. Incorporating this intuitive assumption into the model adds a dynamic structure.

One can also consider an alternative mechanism for the price formation: prices administered by producers. Namely, in each market the dominant producer sets a target price that depends on the previous prices, as well as on the past non-price factors, without knowing the current prices for other commodities, and in each period price adjusts partially towards its target value. This is different from the price-taking behaviour in the perfectly competitive markets, and the behaviour of the producers is described by a dynamic equation that is not derived from fundamentals but is based on an intuitively plausible assumption of administered prices. More structure can be added by assuming that supply and demand, and, therefore, prices, are subject to random shocks that can persist.

1.5 From economic model to econometric model

After the research questions and the hypotheses of interest have been formulated and the appropriate economic model set up, we can move on to data collection and analysis. The economic model, together with assumptions about the behaviour of economic agents, along with the assumed structure of random disturbances, leads to an econometric model, comprising an equation or a system of equations, to which an appropriate method of regression analysis can be applied for estimation and inference. Suppose, for example, that we wish to test a simple version of the Quantity Theory of Money (see Chapter 3). According to the theory, a one per cent change in the money stock leads to a one per cent change in the price level (other things being equal). We can regress price on the money stock and other relevant variables and test the significance of the estimated coefficients. To test the Quantity Theory we test whether the coefficient linking the logarithm of the price to the logarithm of the money stock is equal to one. To test whether the money stock has influence on price at all we test whether the same coefficient is zero.

Typically, to test a hypothesis one has to construct a test statistic whose distribution (exact or approximate) is known if the null hypothesis is true; the regression results are used to compute this test statistic, and the null hypothesis is rejected in favour of the alternative if the value of the test statistic obtained from the data sample falls in the so-called rejection range. This range is associated with a small probability that the test statistic will lie within it when a random draw is made from the distribution under the null; the relevant probability is the 'significance level' of the test. The critical values defining the range in which the test statistic must fall in order to reject the null depend upon the nature of the alternative(s), and in particular whether the alternative(s) are 'one-sided' or 'two-sided' (i.e. if deviations from the null in both directions are allowed).

When interpreting regression results one must remember two important points.

- Estimation is made under the *ceteris paribus* assumption. That is, any factors that have not been included in the estimated equation are assumed to be constant and, therefore, any prediction based on the estimated model is only valid if these excluded factors remain the same. Furthermore, each of the estimated model parameters, or the coefficients in the regression equation, gives the effect of the associated explanatory variable upon the dependent variable holding all other variables constant.

- Sample data by themselves only indicate association (or correlation) among variables and contain no information about causality. The existence and the direction of the cause and effect are postulated in the economic (structural) model and cannot be inferred directly from the data; the data analysis can only suggest evidence in support of or contrary to the economic model.

1.6 Some practical issues in single-equation regression analysis

Multicollinearity

In a single-equation regression model the dependent variable, Y, is assumed to be a function of the explanatory variables, $X_1, \ldots, X_K$. The exact functional form can be suggested by economic theory, but more often than not the functional form is unknown, and in many practical applications it is reasonable to approximate it, for simplicity, by a linear function,

$$Y = \beta_0 + \beta_1 X_1 + \ldots + \beta_K X_K + \varepsilon$$

where ε is the random error. The random error can be thought of as the combination of several factors:

- variables that may have an effect on Y but are not included in the model (either because they are not in the data or because they can be ignored in the context of the problem);
- measurement errors; and
- the deviation of the true functional form from the assumed one.

The main assumptions of the classical linear regression model rule out any problem associated with the sources of the random error (specifically, these assumptions include the linear functional form being correct and the random error being uncorrelated with the explanatory variables). In this case the regression coefficients can be estimated by ordinary least squares (OLS), and the resulting estimators are unbiased and efficient (have the smallest sampling variance in the respective class of estimators).

A difficulty with the practical implementation of this procedure arises when two or more explanatory variables are strongly correlated with each other. This property of the data is called multicollinearity; perfect multicollinearity occurs when there is an exact linear relationship between two or more explanatory variables. In the latter case some variables are redundant and must be removed from the equation (in fact, perfect multicollinearity violates one of the assumptions of the classical linear regression model that can be loosely formulated as ‘no redundant explanatory variables’).

With imperfect multicollinearity the regression equation can still be estimated by OLS, but the estimated coefficients have large sampling variability (large standard errors) and may all appear statistically insignificant. Intuitively, if two or more explanatory variables tend to move together in the sample, it is difficult to disentangle their individual effects upon the dependent variable. It is important to remember that multicollinearity is not a problem with the method of estimation, but is rather a property of the data set, and the only way to solve this 'problem' is to obtain better data.

It is often the case that on preliminary inspection, a scatter plot of a dependent and an explanatory variable may suggest that they do not fit a linear pattern. This does not necessarily mean that the linear regression model cannot be used. In fact, 'linear' here refers to the parameters (regression coefficients), whereas the variables and their combinations can be transformed in a non-linear fashion.

Non-linear transformation allows for more flexibility in fitting the data and can improve explanatory power. For example, including both a variable and its square as explanatory variables in a regression makes it possible to capture a non-monotonic response in the dependent variable. Using logarithms of variables allows estimation of the effect of a proportional change, rather than change in levels. Logarithmic transformation, however, cannot be applied if a variable takes the zero or negative values in the sample.

Omitted variables that are correlated with included variables

When an excluded variable is correlated with the included explanatory variables OLS is no longer a valid estimation procedure. Technically, because the excluded variables form part of the random error, the latter becomes correlated with the included explanatory variables. Intuitively, because of the correlation between the included and the excluded variables the estimated effect of an included variable may be confounded with the effect of an excluded variable. This is often referred to as omitted variable bias. It is important to remember that bias occurs in the estimated effects of only those included variables that are correlated with excluded variables; the estimated effects of other variables remain unbiased.

Endogeneity

An endogenous variable, broadly defined, is a variable whose value in the model is determined by other variables. A typical economic model relates a set of exogenous variables, whose values are given or can be controlled, to one or more endogenous variables. Endogeneity is a natural feature of economic systems, e.g. where an endogenous variable, such as a price or quantity, is determined by an equilibrium between demand and supply, and demand and supply are each influenced by a set of exogenous variables. In historical research it is appropriate to regard most variables as endogenous, particular over a long period of time, during which policies and institutions are liable to change in response to previous events. This creates a problem, however, because there are many endogenous variables to be explained and very few measureable exogenous variables with which to explain them. The problem is often 'solved' by improvisation: some of the endogenous variables are treated as exogenous, and are used to explain the remaining endogenous variables. In the limiting case one variable – the variable that is mainly of interest – is treated as endogenous, and all the other variables, whether exogenous or not, are treated as if they were exogenous. Treating an exogenous variable as endogenous creates a problem, however – a problem known in the literature as the 'endogeneity problem'.

The endogeneity problem is another case, like omitted variables, in which there is correlation between the explanatory variables and the random error that impacts on the dependent variable. This arises because the random disturbance that affects the dependent variable influences one or more of the explanatory variables too. Endogeneity, like omitted variables, introduces bias into OLS estimators of the coefficients which does not disappear even in a very large sample.

The problem can be resolved by modelling the source of endogeneity explicitly and transforming the equation(s) to be estimated into a form where the key assumption is no longer violated. When endogeneity arises from equilibrium relationships between dependent and explanatory variables, it is appropriate to model the equilibrium using a system of structural equations suggested by economic theory. The structural equations can be transformed linearly into a set of reduced form equations in which endogenous variables only appear on the left-hand side. Under suitable conditions the reduced form equations can then be estimated separately by OLS. Estimates of structural parameters can be recovered, under certain conditions, from the estimates of the reduced form parameters by reversing (inverting) the linear transformation. This is described in more detail in section 1.9.

Another approach is to replace endogenous explanatory variables by their linear predictors (fitted values from a linear regression) based on the exogenous ones, known as 'instruments'. The instrumental variables are correlated with the endogenous explanatory variables but are not influenced by the dependent variable. This makes it possible for the instrumental variables to be uncorrelated with the error that influences the dependent variable while acting as acceptable proxies for endogenous explanatory variables. In time series analysis lagged values of endogenous explanatory variables are often used as instruments, although other methods are also available (see Greasley and Oxley, 2011).

Autocorrelation

Autocorrelation arises when a variable is correlated with its own past values. Although this is quite common, it can create problems, particularly when the error term in a regression is autocorrelated. In a time series regression autocorrelation in the error component makes OLS estimators of the coefficients inefficient but not necessarily biased. A more serious problem, namely endogeneity, arises when a lagged dependent variable is used as an explanatory variable in the presence of autocorrelation; this is similar to the omitted variable problem.

Autocorrelation can be detected from the pattern in the residuals from the OLS regression. There are a number of tests for autocorrelation that are conveniently included with OLS estimation in the standard software packages. If we interpret excluded variables as unobserved shocks, autocorrelation in the OLS residuals suggests the presence of persistence in the shocks. One could argue, of course, that autocorrelation is essentially the result of misspecification of the model, and that in a correctly specified model there should be no pattern in the random error. This particular type of misspecification can be examined using the so-called common factors test (see Chapter 20 in Greene, 2012 for discussion).

Heteroskedasticity

Heteroskedasticity, or unequal variance of the random error for different observations, can be viewed as a misspecification of the functional form of the regression equation. Sometimes the variance of the random error is monotonically related to a particular variable (for example, expenditure on certain consumer goods may vary more, the higher is the household income), although non-monotonic patterns are also possible. Under heteroskedasticity the OLS estimators of the coefficients in the regression model are unbiased, but the estimators of the standard errors, and therefore the confidence intervals, are incorrect. This can be fixed by applying the generalized least squares (GLS) estimation procedure, whereby the model is, essentially, re-specified (rescaled) to eliminate the suspected pattern in the error variance. See Chapter 9 in Greene (2012) for details.

Another example of a mis-specified functional form leading to heteroskedasticity is when a linear function is fitted to a quadratic relationship. In this case the OLS residuals closer to the points of intersection of the true (quadratic) function and the assumed (linear) function will tend to be smaller, and those farther away from the intersection points will tend to be larger, suggesting a corresponding pattern in the error variance. To resolve this problem it is necessary to explore more flexible, non-linear functional forms for the regression model.

The presence of an observation with a very large OLS residual, or an outlier, should not necessarily suggest that the error variance for this observation is larger than for the rest of the sample. First, an outlier can arise from a mistake in the data (e.g. a misplaced decimal point). Second, even when the error variance is the same for all observations, it is always possible to observe extreme values in a finite sample. Finally, an outlier could be produced by a different data generating process – that is, belong to a different population. The main problem here is that an outlier can sometimes drive the entire regression result. Once again, this can be a mis-specification of the model: if there is a good reason for a particular observation or a set of observations to lie 'outside the model' then this should be incorporated into the model explicitly. For example, the effect of a natural disaster or a war in the time series data can be modelled using a dummy variable, as described in section 1.8 below.

There is a technical procedure for correcting the estimated standard errors for autocorrelation and heteroskedasticity, which is included in many standard statistical software packages. However, certain forms of autocorrelation and heteroskedasticity (such as autoregressive conditional heteroskedasticity, for example) need to be modelled explicitly. A correctly specified model should include an equation describing the relationship between the variables of interest and an equation describing the pattern in the error variance.

Deviations from normality in the distribution of error

The assumption of the normal distribution of the error term in a regression permits the derivation of the exact distribution in a finite sample of a number of useful statistics (for example, statistics for a t-test or an F-test) and the construction of exact confidence intervals for the OLS estimates of the model coefficients. When the normality assumption is violated it is still possible to test various hypotheses using statistics with known asymptotic distributions, because the distribution of a statistic in an infinitely large sample does not depend on the underlying error distribution. However, asymptotic properties can be a poor approximation in a finite sample. One common situation of non-normality is existence of so-called fat tails (high kurtosis) in the distribution of the error, i.e. a higher probability of larger (by absolute value) realisations of errors than that for normal distribution. Examples of distributions with fat tails are the t-distribution and the logistic distribution.

Small sample size

Insufficient sample size can be a practical problem with implementing regression analysis. There is no threshold for an ‘acceptable’ sample size; technically, we need at least as many observations as the model parameters that need to be estimated. Large samples justify the use of an asymptotic (normal or chi-square) distribution for various test statistics. For many useful test statistics their exact distribution in a sample of any finite size is known, and as long as we are willing to accept the assumptions under which these exact distributions are derived (typically, the assumption of the normal distribution of the error term), the size of the sample does not matter – apart from the possibility, of course, that a larger sample may contain more information that could be utilized. Advice to apply certain techniques only when some size of data set is available usually reflects concerns about the robustness of the techniques when certain underlying assumptions do not hold, rather than any breakdown in the techniques when their assumed conditions hold.

Non-random sample selection

An important assumption in the standard regression analysis is that the data sample is random, and therefore representative of a population. If the data were selected by applying to the underlying population some observed or unobserved criterion that deviates from random selection then the sample may not represent the population, and the inference from the estimation cannot be extended to the population in general. For example, the data on the real estate transactions may only include properties with a value above certain level (truncation). Selection in the sample may occur in relation to some explanatory variable or in relation to the dependent variable; in the latter case OLS estimation is invalid. Alternative estimation techniques involve specifying an additional equation that describes the process of selection.

Another problem can be caused by the way in which the data are recorded. For example, in a survey incomes at or above certain level may be reported as a single point, such as '£100,000 or above' (censoring). A censored sample may still be representative of population, but the type of distortion in the data makes OLS inapplicable and calls for the alternative estimation techniques somewhat similar to those for a truncated sample. See Chapter 19 in Greene (2012) for details.

1.7 Quantitative and qualitative variables

Very broadly, variables in the economic history research belong to two types, quantitative and qualitative (also known as categorical) variables. A quantitative variable has a numerical measurement and can be used in the data analysis directly as recorded or after some appropriate transformation.

A qualitative variable places an observation in a particular category that has no numerical measure. Suppose we want to include a region as an explanatory variable and believe that the division into the North-west, North-east, South-west and South-east is appropriate. In that case we need to define a binary (0/1) variable for each category, so that, for example, variable NW_i takes value 1 if observation i belongs to the North-west and 0 otherwise. In a linear regression one category must be omitted whenever the regression equation includes a constant; this omitted category serves as a reference point, or 'control', for the interpretation of the estimated coefficients for other categories. In time series data a qualitative variable can denote a time period (say, before and after); this can be one point in time or a number of consecutive points. To include such a variable in the data analysis we define a binary variable that takes value 1 if the observation falls into the specified time period and 0 otherwise (see Chapter 4).

Once the qualitative variables are converted into binary variables they can be included in an econometric model along with the quantitative variables. A linear regression model can be estimated in the usual way so long as the dependent variable is quantitative, but it becomes problematic when the dependent variable is qualitative. For example, the dependent variable may be represented by count data, yes/no choices, various rankings, or choices among ordered categories, as well as choices among unordered categories (in the latter case the distinction is made between the choices based on the characteristics of an individual and those based on the attributes of the available options). In principle, if there are only two possible categories for a qualitative dependent variable, one can convert it to a binary variable and estimate a linear probability model using OLS. This is, however, not a satisfactory method: the values of the dependent variable predicted by the model have the interpretation of probabilities, and in a linear model this number can be greater than one or negative. Frequently used qualitative response models that avoid this problem are the probit and the logit models. They 'squash' the dependent variable into the range between zero and one and therefore avoid the problem of predicted probability values outside this range. Here we provide a brief outline of the formal modelling of a binary choice; more details on this model and on a more general case of a multinomial choice can be found in the standard econometrics texts.

Suppose a set of factors $\boldsymbol{X}$ are believed to explain a discrete choice, such as $Y=1$ (Yes, or in the category of interest) versus $Y=0$ (No, or not in the category of interest). This is modelled as a probability of Y taking one of these two values, conditional on the realisation of the factors:

$$\Pr[Y=1|\boldsymbol{X}=\boldsymbol{x}]=F(\boldsymbol{x},\boldsymbol{\theta}) \text{ and } \Pr[Y=0|\boldsymbol{X}=\boldsymbol{x}]=1-F(\boldsymbol{x},\boldsymbol{\theta})$$

where $\boldsymbol{\theta}$ is a vector of unknown parameters whose values are to be estimated. Any continuous probability distribution function can be used for $F(\cdot)$. Two commonly used probability models are the probit, where $F(\boldsymbol{x},\boldsymbol{\theta})=\Phi(\boldsymbol{x}'\boldsymbol{\theta})$ is the normal distribution, and the logit, where $F(\boldsymbol{x},\boldsymbol{\theta})=\Lambda(\boldsymbol{x}'\boldsymbol{\theta})=\exp(\boldsymbol{x}'\boldsymbol{\theta})/(1+\exp(\boldsymbol{x}'\boldsymbol{\theta}))$ is the logistic distribution. Both distribution functions are symmetric and bell-shaped, with the logistic distribution having heavier tails (more like the t-distribution). Other symmetric or asymmetric (such as Weibull, loglog) distributions are used in some applications. The linear probability model is, in fact, a particular case where $F(\boldsymbol{x},\boldsymbol{\theta})=\boldsymbol{x}'\boldsymbol{\theta}$.

The regression equation is estimated by the maximum likelihood method, and the procedure is built into many popular econometrics software packages. One has to be careful with the interpretation of the estimation results. More specifically, the conditional mean of the dependent variable is given by the regression equation:

$$E[y|\boldsymbol{x}]=1\times F(\boldsymbol{x}'\boldsymbol{\theta})+0\times[1-F(\boldsymbol{x}'\boldsymbol{\theta})]=F(\boldsymbol{x}'\boldsymbol{\theta})$$

This is a non-linear model; therefore, the marginal effects of the explanatory variables, $\partial E[y|\boldsymbol{x}\vdots]/\partial \mathbf{x}$, are not the same as the model parameters, $\boldsymbol{\theta}$. Furthermore, the marginal effects are not constant, as they would be in a linear regression model, but instead depend on the values of the explanatory variables: $\partial E[y|\boldsymbol{x}\vdots]/\partial \mathbf{x}=f(\boldsymbol{x}'\boldsymbol{\theta})\,\boldsymbol{\theta}$ where $f(u)\equiv dF(u)/du$.

For the purpose of the reporting of the results the marginal effects can be computed either by averaging the marginal effects across observations or by evaluating the effect at the sample averages of the data. For a binary (dummy) explanatory variable the marginal effect (ME) is calculated at the sample average for all other explanatory variables (Greene, 2012, Chapter 17):

$$ME=\Pr[Y=1|\boldsymbol{x}\backslash \mathrm{D};\ \mathrm{D}=1]-\Pr[Y=1|\boldsymbol{x}\backslash \mathrm{D};\ \mathrm{D}=0]$$

1.8 Stochastic and deterministic trends and structural breaks in time series

Analysis of time series data often requires a methodology that is different from what is normally applied to cross-sectional data. Sometimes, as the first step, we need to transform the data in an appropriate way before proceeding to the analysis of the relationship among the variables. Transformation into a stationary form may be required when the data appears to be non-stationary. The definition of stationarity in the strong sense is rather technical (and, furthermore, in practice it is almost impossible to verify strong stationarity), but for many applications stationarity in the weak sense suffices. A time series is said to be weakly stationary if it has a constant mean and a constant variance, and the covariance between different lags depends only on the distance between the lags and not on time.

Often a variable of interest appears to have a time-dependent mean, or, in other words, exhibits a trend. More generally, one can think of a variable as being a superposition of three components:

$$Y_t = \text{Trend} + \text{Stationary component} + \text{Noise}$$

A trend can be defined as a permanent component of a time series. It is important to distinguish between a deterministic trend and a stochastic trend. A time series $\{Y_t\}$ is said to exhibit a deterministic trend if in every period Y_t (or its d-th difference) changes by a fixed amount,

$$\Delta Y_t = Y_t - Y_{t-1} = \delta.$$

The first difference is defined as $\Delta^1 Y_t = \Delta Y_t$, and the d-th difference is $\Delta^d Y_t = \Delta^1(\Delta^{d-1} Y_t)$. Clearly, a variable with deterministic trend is non-stationary because its mean changes over time. On the other hand, $\{Y_t\}$ is said to exhibit a stochastic trend if in every period Y_t (or its d-th difference) is expected to change by a fixed amount,

$$Y_t - Y_{t-1} = \delta + \varepsilon_t$$

where $E[\varepsilon_t] = 0$. With a stochastic trend each shock ε_i, $i = 1, 2, \ldots$ has a permanent effect (shift) on the conditional mean of Y_t, but these shifts are random. The equation above describes the so-called random walk if $\delta = 0$. A random walk variable is non-stationary, even though it has a constant mean of zero, because its variance is not constant (it grows linearly with time). A variable can exhibit a general trend, i.e. both a deterministic and the stochastic trend can be present, such as in the equation above for $\delta \neq 0$; this describes a random walk with drift.

Often the object of interest is the unexplained component of the time series. To isolate this component one needs to eliminate the trend and the stationary components, which can be done by estimating and subtracting each of these components from the series. Once the trend is removed, the stationary component can be estimated, for example, by using a Box–Jenkins approach. The appropriate methods for eliminating the trend are de-trending for a deterministic trend and differencing for a stochastic trend. De-trending involves estimating and subtracting a deterministic function of time (usually a polynomial), whereas differencing is applied directly to the series. De-trending a variable with a stochastic trend and differencing a variable with a deterministic trend leads to errors in the subsequent analysis. It is, therefore, extremely important to model the trend correctly in order to carry out valid inference and hypothesis testing and to calculate a forecast. To establish stationarity we can use a combination of tools, such as the correlograms (autocorrelation functions and partial autocorrelation functions) and the unit root tests. This is, however, difficult in small samples or when structural changes, or breaks, are present.

A structural break can be the object of interest on its own. For example, a law introduced in a particular year, or a pandemic disease, could have affected economic activity temporarily or permanently. Structural breaks of this sort are modelled by a binary, or a dummy variable. A one-time shift in Y_t that occurred in period $\hat{t}$ is represented by a pulse dummy variable:

$$D_t^P = \begin{cases} 0, t = 1, \ldots, \hat{t} - 1, \hat{t} + 1, \ldots, T, \\ 1, t = \hat{t}. \end{cases}$$

This has a permanent effect on the level of Y_t containing stochastic trend. On the other hand, a permanent shift in Y_t that occurred in period $\hat{t}$ is represented by a level dummy variable:

$$D_t^P = \begin{cases} 0, t = 1, \ldots, \hat{t} - 1, \\ 1, t = \hat{t}, \hat{t} + 1, \ldots, T. \end{cases}$$

Empirically, a stationary process with a permanent jump can be easily mistaken for a unit-root (stochastic trend) process with a one-time jump. This distinction is especially difficult to make when the length of the data sample before the jump or after the jump is relatively short. Another important issue is the dates of the (potential) structural changes: typically, we hypothesize the date of the change from the historical events and use the data to test this hypothesis, rather than inferring the date of the change from data analysis. More on this topic can be found in Enders (2010, Chapters 2, 4) and in Chapter 4 below.

1.9 Simultaneous equations models and the identification problem

Simultaneity, or endogeneity, arises when the variables of interest are determined simultaneously in equilibrium (see section 1.6 above). The classical example is the determination of market prices for substitutable or complementary commodities. In this section we briefly outline the problem using a simple model of two equations and two variables and refer the reader to the standard texts for more details and more general cases.

Suppose that economic theory suggests the following structural model:

$$y_t = b_{10} - b_{12}z_t + \gamma_{11}y_{t-1} + \gamma_{12}z_{t-1} + \varepsilon_{yt}$$

$$z_t = b_{20} - b_{21}y_t + \gamma_{21}y_{t-1} + \gamma_{22}z_{t-1} + \varepsilon_{zt}$$

This could be the model of price formation for two commodities, with y_t being the price of wheat and z_t being the price of barley, based on certain assumptions about the interaction between two markets. Here bs and γs are unknown constant coefficients, and εs are unobserved structural shocks. Assume, for simplicity, that $\{\varepsilon_{yt}\}$ and $\{\varepsilon_{zt}\}$ are uncorrelated white noise random variables, that is, $E[\varepsilon_{yt}]=E[\varepsilon_{zt}]=0$, $Var(\varepsilon_{yt})=\sigma_y^2$, $Var(\varepsilon_{zt})=\sigma_z^2$, $Cov(\varepsilon_{yt}, \varepsilon_{ys})=0$, and $Cov(\varepsilon_{yt}, \varepsilon_{zt})=Cov(\varepsilon_{yt}, \varepsilon_{zs})=0$ for all t and for all $s\neq t$. The collection of unknown bs, γs, and $\sigma_{y,z}{}^2$ are the structural parameters of the model.

Variables y_t and z_t in this model are endogenous because they are determined simultaneously in a system of equations. Ordinary least squares (OLS) applied to a single equation with endogenous explanatory variable(s) is invalid as it yields inconsistent estimators of the regression coefficients. To overcome this problem we can solve the system of equations to eliminate the endogenous variables from the right-hand side. Thus, we obtain a reduced-form model:

$$y_t = \alpha_{10} + \alpha_{11}y_{t-1} + \alpha_{12}z_{t-1} + e_{1t}$$

$$z_t = \alpha_{20} + \alpha_{21}y_{t-1} + \alpha_{22}z_{t-1} + e_{2t}$$

The lagged variables in the right-hand side are predetermined (realised before the endogenous variable in the left-hand-side) and therefore do not create the simultaneity problem. We can estimate the coefficients of the reduced-form system by OLS, which will produce consistent estimates. Using the OLS residuals from the two equations we can estimate the variances and covariances of the reduced-form shocks. The reduced-form shocks are linear combinations of the structural shocks, and their variances and covariances are given by

$$\mathrm{Var}(e_{1t}) = \frac{(\sigma_y^2 + b_{12}\sigma_z^2)}{(1-b_{12}b_{21})^2}$$

$$\mathrm{Var}(e_{2t}) = \frac{(\sigma_z^2 + b_{21}\sigma_y^2)}{(1-b_{12}b_{21})^2}$$

$$\mathrm{Cov}(e_{1t}, e_{2t}) = -\frac{(b_{21}\sigma_y^2 + b_{12}\sigma_z^2)}{(1-b_{12}b_{21})^2}$$

Next, we can use the OLS estimates of the reduced-form parameters to estimate consistently the structural parameters. However, in general, the structural equations are under-identified: there are fewer equations relating the 'known' estimated reduced-form parameters and the 'unknown' structural parameters, than there are unknown parameters.

In general, to identify a model described by a system of k equations we need to impose $(k^2-k)/2$ restrictions on the structural parameters. These must be based in economic theory. For example, we can impose a particular causal chain by assuming that the current value of y (price of barley) is affected by the current and past values of z (price of wheat), as well as its own past values, but z is only affected by its own past values and the past values of y (wheat is a 'dominant' commodity). Such as assumption means $b_{21}=0$, and the system becomes recursive. The corresponding decomposition of the residuals is referred to as the triangular, or Cholesky decomposition. Technically, for k equations $k!$ (k factorial) possible orderings are possible, leading to the different estimated parameters, and one has to choose the ordering that can be plausibly suggested by economic theory. An alternative restriction is the assumption of the symmetric response: $b_{21}=b_{12}$. One can also impose a restriction on some coefficient, say, $b_{12}=1$, or on a variance, say, $\sigma_y^2=1$, but theory does not typically offer much guidance on the magnitudes of the parameters. In more sophisticated structural models researchers use various short-run restrictions (typically in the form of exclusion, i.e. $b_{ij}=0$) and long-run restrictions (e.g. accumulated response to a certain shock is zero in the long run), as well as the sign restrictions. It is possible that theory offers 'too many' restrictions. An over-identified model can be estimated by the generalised method of moments (GMM). See Enders (2010, Chapter 5); Greene (2012, Chapters 10 and 13) for details.

1.10 Partial adjustment and adaptive expectations models

In many economic contexts one can plausibly assume that the variable of interest has a desired, or a target level, to which the variable is adjusting gradually, or partially:

$$Y_t^* = f(X_t;\varepsilon_t),$$

$$Y_t - Y_{t-1} = (1-\lambda)(Y_t^* - Y_{t-1})$$

This can be a description of a price-setting behaviour of a dominant producer in an imperfectly competitive market. The last equation can be rewritten as

$$Y_t = (1-\lambda)Y_t^* + \lambda Y_{t-1} = f(X_t;\varepsilon_t) + \lambda Y_{t-1}$$

To estimate the model we can assume that the target level is described by a linear function of the factor(s), and, therefore,

$$Y_t = \alpha + \beta X_t + \lambda Y_{t-1}.$$

The interpretation of the model parameters is straightforward. The short-run multiplier is β; it shows the contemporaneous effect of a change in the explanatory variable. The long-run effect is $\beta/(1-\lambda)$; this is the effect of a one-off change in the explanatory variable on the target value.

A recursive substitution shows that the partial adjustment equation is equivalent to representing Y_t as a linear function of infinitely many lags of X_t, with the lag weights declining geometrically. An alternative motivation of an econometric model with infinite number of lags decaying geometrically is the economic model of adaptive expectations. Suppose a producer needs to decide how many acres of wheat to plant, Y_t, based on his current expectations of the wheat price next year, $E_t[X_{t+1}]$. The process of formation of the expectations is modelled as the adaptation of the previously expected price to the observed price:

$$Y_t = \alpha + \beta E_t[X_{t+1}] + \varepsilon_t,$$

$$E_t[X_{t+1}] = \lambda E_{t-1}[X_t] + (1-\lambda)X_t.$$

In other words, if the current realised price is higher than it was expected in the previous year, the expectation for the next year's price is adjusted upwards, and vice versa. Parameter λ measures the relative weight of the previous expectation. Two extreme cases are $\lambda=1$, when the expectations are never revised and the current observation is ignored, and $\lambda=0$, when past information is ignored.

The process of formation of the expectation can be rewritten as an equation with infinitely many lags,

$$E_t[X_{t+1}] = (1-\lambda)(X_t + \lambda X_{t-1} + \lambda^2 X_{t-2} + \cdots)$$

The corresponding econometric model, described by a system of equations

$$Y_t = \alpha + \beta Z_t(\lambda) + \varepsilon_t,$$

$$Z_t(\lambda) = X_t + \lambda Z_{t-1}(\lambda)$$

is estimated by non-linear least squares, assuming X_t was in the long-run equilibrium (this is plausible for a long data set).

1.11 Missing observations

Missing observations is a difficulty often faced in data analysis especially, but not only, in economic history research. This can be because of the nature of the data collection process: for example, in time series different variables were recorded (observed) with different frequencies, or in survey data respondents fail to answer some questions. Missing observations may be a conceptual problem. It is important to distinguish two cases for why an observation is unavailable. In the first case the reason is unknown and unrelated to the completeness of other observations. This is referred to as the ignorable case, and the observations are said to be missing at random. In the second case the reason is systematically related to the situation being modelled. This is the self-selection case, and the observations are said to be missing by design.

In the first case the issue is purely technical and can be ignored. We can drop the variable with many missing observations or, in a cross-section or a panel data set, drop the cross-sectional units with missing data and use the complete observations for the purpose of estimation. Furthermore, we can try to extract more information from the incomplete observations in order to improve the efficiency of estimation, for example, by filling the gaps. Missing values of the dependent variable can be replaced by the predicted values. This, however, creates a bias in the estimated coefficients which is difficult to quantify. There are several methods for filling the gaps in the explanatory variables. One can replace missing values by the sample average (or by the average of the appropriate sub-sample, depending on the structure of the data). An equivalent method is to replace missing values by zeros and add to the regression a binary (dummy) variable that takes value 1 if the respective observation is missing and 0 otherwise. Alternatively, we can replace missing values by the predicted values of the explanatory variable obtained from estimating a 'reverse' regression of the explanatory variable on the dependent variable using the available observations. However, there is no systematic evidence of this approach being better than filling the gaps with the averages.

In the second case the complete observations are qualitatively different from the incomplete observations. Effectively, one needs to model the reason for observations not being there simultaneously with modelling the patterns in the observed data. There is a range of models that are appropriate in various situations: truncation, censoring, sample selection (treatment effect) and duration models (Greene, 2012, Chapter 19).

1.12 Standard econometric packages

There are a large number of standard computer packages for econometric and statistical analysis that can be easily mastered and used in the economic history research. An excellent overview of the publicly available econometrics software is given in Renfro (2004), who identifies four categories of econometrics software:

- independent econometric software packages, e.g. AUTOBOX, EasyReg, EViews, LIMDEP, MicroFit, MODLER, PcGets, PcGive, RATS, STAMP, STATA, TSP, WinSolve;
- econometric programming libraries, e.g. BACC, MLE++, MODLER;
- econometric and mathematical programming languages, e.g. GAUSS, OX; and
- applications of econometric and mathematical programming languages, e.g. ARFIMA package for OX, DPD for OX, G@RCH, MSVAR, TSMod.

Some software falls in more than one category. The author provides a comprehensive description of the current capabilities of many packages, as well as the links to the online resources and references to manuals and other relevant publications.

Some basic statistical tools are built into Excel and Access, but these quickly become insufficient even for a moderately sophisticated research. Popular software for statistical analysis such as SPSS or STATISTICA is more advanced, but does not cover all aspects of regression analysis. Modern off-the-shelf packages are extremely popular with researchers and research students because they are easy to use and because they do not seem to require a great deal of knowledge of statistical and econometric theory. It is easy to get started; most of the modern packages have friendly drop-down menu, and so there is no need to write an elaborate computer code. The output is conveniently arranged in tables, and other ways of visualisation of the data and the output, such as various plots, are readily available.

One must remember, however, the limitations of such software. As with any complex numerical computational procedure, the software may have various numerical 'bugs'. Often some of the default output is irrelevant and, on the other hand, what we need may not be reported. Most importantly, the estimation procedure from the drop-down menu may have default options that are inappropriate for the research question or for the data used. Sometimes, however, it is sufficient to check carefully and modify the options as appropriate. For example, in EViews a vector autoregressive (VAR) model (a reduced-form system) is estimated by default with Cholesky decomposition using the order in which variables were initially typed into the input window. A mistake, of course, can be avoided by typing the variables in the correct pre-selected order, assuming the recursive structure is appropriate in the context of the model. Outside the recursive structure, however, not all types of restrictions are allowed, and structural decompositions not available for the vector error correction (VEC) model (an extension of VAR allowing for the estimation of the short-run and the long-run effects); in certain situations default signs of coefficients are imposed. Experienced researchers are aware of these limitations and can often overcome the problems by writing a computer code specific for the research question and the data; this can be done within EViews and many other software packages that are also menu-driven. A good understanding of the underlying economic model and of the applicability of an econometric technique is, therefore, essential.

References

Enders, Walter (2010) *Applied Econometric Time Series*, 2nd edn, Chichester: Wiley.

Greasley, D. and L. Oxley (2011) *Economics and History: Surveys in Cliometrics*, Oxford: Wiley-Blackwell.

Greene, William H. (2012) *Econometric Analysis*. 7th edn, London: Pearson.

Lind, Douglas A., William G. Marchal and Samuel A. Wathen (2012) *Basic Statistics for Business and Economics*, 8th edn, Maidenhead: McGraw-Hill.

Renfro, Charles G. (ed.) (2004) A compendium of existing econometric software packages, *Journal of Economic and Social Measurement*, 29, 359–409.

Wooldridge, J.M. (2006) *Introductory Econometrics: A Modern Approach*, Mason, OH: Thomson/South-Western.

2 Long-run price dynamics

The measurement of substitutability between commodities

Mark Casson, Nigar Hashimzade and Catherine Casson

2.1 Introduction

Prices figure prominently in economic history but, despite this, there has been relatively little methodological investigation of what can or cannot be inferred from movements in prices. Price indexes of basic commodities have been widely used to measure changes in real wages and the cost of living (Allen, 2001). Historians of the Industrial Revolution have used changes in relative prices to explain structural changes in the economy (Allen, 2009); global historians have related spatial price convergence in commodity prices to lower transport costs (Findlay and O'Rourke, 2003); financial historians have used asset prices to investigate the efficiency of money and capital markets (Neal and Atack, 2008), while monetary historians have related aggregate price levels to the money supply (Mayhew, 1995). More controversially, prices have been used to analyse long-run cycles (Fischer, 1996).

In most of these studies, however, the emphasis is on tracking and analysing a single price (Labys, 2006). The prices of individual commodities are either analysed separately (e.g. a study of the wheat market), or aggregated using appropriate weights to construct a single price index (Beveridge, 1939; Lloyd, 1973). Price convergence studies analyse the prices of the same commodity at different locations. By contrast, there has been little study of the relative prices of different commodities at the same location. The statistical methods used to analyse price convergence cannot be applied to analyse relative prices because the long-run market equilibria are determined by different processes. The prices of the same commodity follow the law of one price whereas the prices of different commodities do not; e.g. because of differential productivity growth, the prices of novel commodities tend to fall over time relative to the prices of mature commodities.

This chapter focuses on the prices of different commodities at the same location. It considers what can be inferred from a panel of annual price data on different commodities covering a reasonably long period of time. In particular, it examines substitutability and complementarity between commodities. It seeks to identify pairs of commodities that are substitutes for one another and other pairs that are complements and, among the substitutes, to distinguish close substitutes from distant ones. In the process of this investigation, a number of other issues arise, including the importance of long-term trends in prices, and whether prices are administered by producers or are determined competitively by the balance of supply and demand (for historical context see Stone, 2005, and Threlfall-Holmes, 2005, pp. 75–135).

The chapter considers whether all prices influence other prices equally, or whether there is a hierarchy of commodities in which some commodity prices influence other prices but are not influenced by them. It argues that such hierarchies do indeed exist, and shows how their structures can be analysed systematically. It shows that commodities at the top of the hierarchy are generally weak substitutes for many commodities, while commodities lower down the hierarchy are strong substitutes for fewer commodities. The analysis is applied to long-run annual commodity price data for England 1210–1914.

2.2 Methodology

The significance of prices in historical narratives

Price data is obviously important in analysing the behaviour of a market economy, but there are special reasons why historians have placed particular emphasis on prices.

- Price data is widely available; prices were recorded from a very early date, including in contracts, account books, diaries, exchequer records, and many such records have survived. As a result, historians often possess long runs of annual price data on a range of commodities for which no corresponding annual quantity data is available.
- Price data for one location can be used to infer prices at other locations; by contrast, quantity at one location cannot normally be inferred from quantity at another location because local circumstances will differ. Thus estimates of national prices can be generated from localised evidence more easily than estimates of national output.
- Prices play an important role in the writing of economic history as a causal explanation of narrative evidence. Thus historians wishing to explain changes in the amount of land under cultivation invoke changes in prices of crops as a causal factor, while those wishing to explain the adoption of steam power invoke changes in the price of coal as a causal factor. Such explanations are inevitably partial, however, because they do not explain why prices changed in the first place.

One reason why prices change is that other prices change; a market economy is a complex interdependent system. Coal, coke, wood and peat, for example, are all fuels, and changes in the price of one of them cannot be fully understood without some awareness of changes in the prices of the others. According to the classical theory of markets, enunciated by Adam Smith (1776) and refined by

Alfred Marshall (1890), prices are ultimately influenced by demand and supply. But because the economic system is interdependent, the price of one commodity can be influenced by changes in the demands and supplies for other commodities. To identify the impact of a change in the demand for one commodity, say bread, it is necessary to net out the impact on the price of changes in the prices of other commodities, such as cake and ale.

The fundamental changes affecting specific markets are often difficult to observe directly. Changes in consumer tastes may be driven by short-term fads and fashions, or by long-term changes in family life which are difficult to quantify and to date precisely. Changes in relative prices may suggest the kinds of unobserved changes that could have occurred, but it is important to know how far such speculation is valid.

Even where fundamental changes are observable, they may not have been recorded on a systematic basis until recently, or the records may not have survived. For example, although weather conditions are widely recognised as an important determinant of annual crop production, historians have often felt compared to infer weather conditions from prices rather than explain prices by weather conditions. This has led some historians into circular arguments, whereby bad weather is inferred from high prices and high prices are then explained by bad weather (Hoskins, 1964, 1968).

If quantity data is not available, and the causes of fundamental changes in demand and supply cannot be directly observed, then the question arises as to what can be inferred from price data alone. The answer is a surprising amount.

If two commodities are substitutes in demand (e.g. bread and cake) then an increase in the price of one will cause consumers to switch their demand to the other, and thereby increase its price. Mutual substitution between two commodities will therefore generate positive correlation between their prices. Conversely, negative correlation can emerge when the two commodities are complements (e.g. bread and butter). An increase in the price of one will reduce the demand for the other, and so the price of one will fall as the price of the other rises.

It is possible to assess which pairs of commodities react most to one another: the stronger the reaction, the closer the substitution (if the correlation is positive) or the complementarity (if it is negative). It is also possible to assess how far these reactions are constant (i.e. time-invariant). If some of these reactions are time-invariant over very long periods then they may be regarded as constants of the economic system. Identifying constant substitutability between specific pairs of commodities could contribute greatly to the interpretation of long-run historical change in a market economy.

Price data can also shed light on the persistence of unobservable shocks to the economy (Nerlove, 1958). A shock persists when its impact on price lasts for longer than a year. If grain can be stored, for example, then an unexpected increase in demand in one year may lead to stocks being drawn down, creating a need to replenish them in future years; thus the increase in price generated by a shock may persist for several years. Furthermore, using the substitutabilities and complementarities between commodities, it is possible to impute these shocks to specific commodity markets. The price of a commodity may vary both because of shocks impinging directly on its own market and because of shocks to other markets transmitted to it through the prices of substitutes. With a knowledge of substitutabilities it is possible to reverse out from a set of price data on a set of commodities the markets that are the principal sources of shocks to the system.

The limitations on what can be inferred from a long-run data set on commodity prices are determined by the number of variances and covariances that can be computed from the data. In general, it is possible to measure as many parameters of the market system as there are variances and covariances calculated from the data set. Since it is possible to correlate the price of a commodity with the price of another commodity (or, indeed, its own price) at an earlier or later date, it is possible to calculate many different covariances corresponding to different lengths of lag, and therefore to estimate a considerable number of parameters.

There is a problem, however, which is sufficiently serious that it has discouraged many researchers from pursuing this approach. Any mutual interaction between two commodity prices involves two parameters, each of which measures the impact of one price on the other, but there is only one covariance between them. With only one statistic but two parameters, it is possible only to estimate the value of one reaction conditional on some assumed value of the other. This is known as the identification problem (Fisher, 1966). In conventional regressions estimated by ordinary least squares it is assumed that interaction is one-sided, and that causation runs unambiguously from one price to another. An alternative, however, is to assume that the two reactions are equal, and to estimate them both on this basis. Other criteria can also be used if necessary, but in this particular chapter symmetric responses only are considered. For further details see Appendix A.0.

Three techniques

There are three main techniques of analysing price statistics, and all of them are discussed in this chapter. The most sophisticated involves structural models of multi-market interaction of the type sketched out above. These models involve equations derived from basic economic principles. The theory determines the variables that appear in the equations and the way that the variables are related to each other. The models also contain free parameters which can be estimated from the data. The estimates are derived using general principles of inference such as maximum likelihood or least squares. The validity of the model as a whole is tested using overall measures of goodness of fit.

It is also possible to use multi-purpose models, such as vector auto-regression (VAR), which can be useful for a variety of purposes. VAR can be applied to the analysis of prices provided that some rather strict assumptions are made. These additional assumptions represent the cost that the researcher pays for using off-the-shelf models. In the present context the key assumption is that prices interact with each other only with a lag and do not influence each other in the same period.

Finally, it is possible to use descriptive statistics, such as means, variances and correlation coefficients to analyse patterns of price variation. This may include the estimation of linear trends and the correlation of prices, not only across commodities, but also over time. The difference between descriptive statistics and the previous technique is that they do not involve estimating the parameters of an economic model.

In theory these techniques are substitutes for each other, but in practice they can be regarded as complements. This is because much historical data analysis is exploratory in nature. Descriptive statistics provide a useful way for the researcher to familiarise themselves with the data and identify any potential problems, such as common trends in the prices of different commodities or possible instabilities in the market system. A VAR analysis is a useful preliminary to the estimation of a structural model because the initial stages of estimating a full structural model may involve the estimation of a VAR. In practice, therefore, it is often useful to apply three approaches in order of increasing sophistication, and this is what is done below.

2.3 Sources

The source is Clark's (2004, 2013) data on English prices and wages 1209–1914, which synthesises previous series constructed by Thorold Rogers, Beveridge, Farmer, Phelps-Brown and Hopkins, and others. It is based on the records of large estates, monasteries, dioceses, university colleges, market towns and other reputable institutions. The recorded prices do not usually reflect the actual prices paid by consumers but are wholesale prices received by sellers; the margin between wholesale and retail price may vary across space and time, e.g. higher margins in booms and lower margins in recessions. Prices relating to the medieval period usually cover the period Michaelmas to Michaelmas (autumn to autumn) and undated prices are attributed to the January of the following year. Some aggregation is involved, which can create statistical problems, but there is no easy way to resolve this problem. When prices vary across locations, weights are used, where appropriate, to combine information from different localities. This means that measurements may be less accurate in years where few local observations are available. Clark's data is particularly useful for assessing volatility because he did not interpolate missing observations; linear interpolation would tend to understate volatility.

Clark's data is in nominal terms and is available in both sterling and an international silver standard. The data can be converted to real terms using a consumer price index as deflator, but this creates complications because some of the commodities included in this study carry a significant weight in such indexes.

Clark provides annual time series for a wide range of commodities, but not all the series are suitable for long-run volatility analysis. Eight series were identified according to the following criteria. Series should comprise long runs of data, commencing as early as possible (e.g. the thirteenth century), and have relatively few missing observations. Series should not be compiled by linking different

Table 2.1 Commodities selected for study, ranked in order of number of observations

Commodity	*No. obs.*	*Start year*	*Finish year*
Wheat (bushel)	677	1209	1914
Oats (bushel)	658	1209	1914
Wool (lb.)	637	1209	1914
Barley (bushel)	633	1209	1914
Peas (bushel)	626	1209	1902
Suet (lb.)	581	1209	1869
Hay (ton)	568	1258	1914
Cheese (lb.)	562	1209	1869

series from very different types of source. The commodities should not have been highly taxed or regulated for prolonged periods. Commodities should be of economic importance, being produced and consumed on a large scale; this makes the results of intrinsic interest, and reduces the risk that for thinly traded products there are errors in the recording of price. The commodities should belong to a well-defined group, which in the present case comprises agricultural products, representing alternative uses of a given plot of land. Table 2.1 shows that the number of observations in each price series ranges from 677 (barley) to 562 (cheese); the earliest starting date is 1209 and the latest is 1258 (hay), while the latest end date is 1914 and the earliest 1869 (cheese).

In line with modern econometric studies of price, the analysis is carried out in terms of the logarithm of price. This is because the magnitude of price shocks tends to be proportional to the level of price and price inflation tends to generate an exponential trend. Taking the logarithm converts proportional changes to absolute changes, exponential trends into linear trends and relative prices into price differentials. In the remainder of the paper 'price' always refers to the logarithm of the price unless the context indicates otherwise.

2.4 Descriptive statistics

Many historians have traditionally adopted a grounded approach based on descriptive statistics. There is a wide range of descriptive statistics, which is not always fully exploited in historical research. This section investigates how far descriptive statistics can provide satisfactory measures of substitutability and complementarity. It shows that partial correlations (as opposed to conventional zero-order correlations) can provide significant insights into price interactions, even though they do not correspond to meaningful estimates of parameter values.

The procedure set out below is designed to explore the data in a systematic way, and may be summarised as follows:

1 Correlate pairs of prices and, assuming normality, examine significance for each pair. To detect substitutability identify significant positive correlations

between pairs of prices, and identify complementarity from negative correlations. The simplest measure of correlation is the Pearson zero-order coefficient, namely the square root of the ratio of the covariance of the prices to the geometric mean of the price variances.

2 Adjust for a deterministic trend. Correlate the deviations from trend.

- Is the trend the same for different commodities? If so, this suggests general inflation or deflation;
- Is the trend different for different commodities? If so, this suggests incremental long-run structural change;
- Does it reduce the correlations? If so, it suggests that trend factors have similar influences on all commodities.

3 Examine autocorrelations in the deviations from trend. Apply an Augmented Dickey–Fuller (ADF) test with and without a deterministic trend. Assess whether there are unit roots, i.e. whether the expected value of the current price is equal to the previous year's price. If so, the multi-market system may not be stable, and this could have serious implications for the estimation procedures described below.

4 Correlate first differences; these are particularly meaningful if the possibility of a unit root cannot be rejected.

5 Regress prices on lagged own-price and a trend, and correlate the residuals. This a refinement of correlating first differences, that is appropriate where there are persistent shocks but no unit roots. If the results differ from those for first differences then it suggests that the decay of persistent shocks needs to be taken into account.

6 Introduce partial correlation coefficients. If there are eight commodities then the highest-order partial correlations that can be calculated are of the sixth order. Partial correlations are applied to either levels, changes in levels or residuals from lagged-own price regressions, depending on which seems most appropriate in the light of the previous results. Partial correlation removes the effects of changes in the prices of all commodities other than the two being considered. In a multi-commodity context it removes the complication of indirect price effects channelled through variation in prices other than the two being considered. However, the assumptions underlying partial correlations between different pairs of prices are different, so that a set of partial correlations cannot provide an internally consistent analysis of price behaviour. To achieve this it is necessary to use a formal multi-market model. Nevertheless partial autocorrelations provide what is perhaps the most useful summary of substitutability that does not involve recourse to a formal model.

Correlating prices without any adjustment generates many correlations of over 0.95, with the lowest (between wheat and wool) being 0.827. De-trending all the series using a linear time trend estimated by ordinary least squares (OLS) regression reduces the correlations significantly; the largest correlation coefficient is now 0.931 (between barley and oats), and a few correlations become negative.

All the time trends are positive and significant at 1 per cent. Most of the trends indicate a steady 2 per cent inflation, with the exception of wool, which has less than 1 per cent inflation. These results suggest an interplay between general inflation driven by monetary factors and a steady decline in the relative price of wool driven by structural change.

There is also evidence of stochastic trends. The residuals from the trend regressions exhibit significant autocorrelation. The zero-order correlation between de-trended prices with a one-year lag is between 0.75 and 0.90 for all commodities. Significant positive zero-order autocorrelations persist for over 30 years, suggesting the possibility of unit roots. ADF tests indicate that unit roots cannot be rejected, except in the case of wool (Table 2.2). This applies with or without a time trend included in the test procedure; including a trend reduces significance levels, favouring rejection of a unit root, but not to below threshold levels.

Table 2.3 shows that the highest zero-order correlations on price changes involve barley, wheat, peas and oats. These 'big four' remain dominant as substitutes throughout the analysis. Correlations are 0.680 between oats and peas, 0.670 between barley and wheat, 0.629 between oats and barley, 0.593 between barley and peas and 0.581 between oats and wheat.

Price changes take the previous price as the benchmark against which current price is measured. If this benchmark is replaced by a data-driven weighted average of previous own-prices then somewhat lower correlations are obtained. If price changes are replaced by the residuals from regressions of current price on lagged own-prices and a time trend (using lags of up to three years) then the highest correlations are 0.619 between oats and peas, 0.474 between wheat and oats, 0.408 between wheat and peas, 0.303 between barley and peas and 0.281 between barley and oats. There is only one negative correlation, between cheese and wool, and this is not significant.

Table 2.4 shows that when partial correlations are taken the correlation coefficients fall even further, and barley emerges as the dominant substitute. The highest correlations are 0.418 between barley and wheat, 0.399 between peas and oats, 0.351 between barley and oats and 0.213 between barley and peas. There are now several negative correlations, although only one, between wheat and suet, is significant at 10 per cent.

2.5 VAR technique: a simple model of administered prices

In a VAR regression for prices, price reactions always involve a lag. In the context of bread and ale, for example, a VAR regression allows the price of bread to respond to the previous year's price of ale, and the price of ale to respond to the previous year's price of bread, but does not allow the current price of bread to interact with the current price of ale. Because it eliminates any interaction in the current period, the model may well be mis-specified. This mis-specification will be indicated by non-zero correlations between the residuals for the bread price and ale price generated by the VAR. This section

Table 2.2 Augmented Dickey–Fuller tests for a unit root in the presence of a constant and a linear trend

	Barley	*Cheese*	*Hay*	*Oats*	*Peas*	*Suet*	*Wheat*	*Wool*
Significance	0.169	0.727	0.822	0.445	0.489	0.976	0.344	0.002
Own-price lag 1 year	–0.057	–0.025	–0.037	–0.041	–0.060	–0.010	–0.049	–0.075
	(0.004)	(0.081)	(0.130)	(0.023)	(0.028)	(0.520)	(0.014)	(0.000)
Deterministic time trend	0.0012	0.0006	0.0008	0.0009	0.0012	0.0003	0.0009	0.0007
	(0.008)	(0.085)	(0.260)	(0.031)	(0.043)	(0.361)	(0.044)	(0.001)
Change in own price lag 1 year	–0.145	–0.495	–0.543	–0.336	–0.309	–0.488	–0.095	–0.204
	(0.001)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.023)	(0.002)
Change in own price lag 2 year	–0.407	–0.229	–0.464	–0.367	–0.371	–0.300	–0.343	
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
Change in own price lag 3 year	–0.253	–0.041	–0.447	–0.318	–0.362	–0.255	–0.290	
	(0.000)	(0.320)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
Change in own price lag 4 year	–0.145		–0.332	–0.199	–0.247	–0.164	–0.212	
	(0.000)		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	
Change in own price lag 5 year	–0.102		–0.299		–0.140		0.064	
	(0.012)		(0.000)		(0.001)		(0.106)	
Change in own price lag 6 year			–0.190				–0.134	
			(0.000)				(0.001)	
Change in own price lag 7 year			–0.212					
			(0.000)					
Constant	–0.385	–0.142	0.225	–0.367	–0.312	–0.044	–0.174	0.250
	(0.020)	(0.139)	(0.053)	(0.047)	(0.085)	(0.496)	(0.174)	(0.002)
R^2	0.204	0.217	0.294	0.211	0.224	0.218	0.193	0.087
F statistic	22.939	26.600	21.910	27.810	23.027	24.162	19.061	19.548
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
DW statistic	2.001	1.842	1.955	2.039	1.996	2.006	2.017	1.961

Note
The significance reported in the top line relates to a test of the unit root null hypothesis. The figures in brackets relate to a null hypothesis that the variable concerned has no impact on change in price in the absence of a unit root**.** Lag lengths for changes in price are optimised within an E-views programme.

Table 2.3 Zero-order Pearson correlations on price changes

	Barley	*Cheese*	*Hay*	*Oats*	*Peas*	*Suet*	*Wheat*	*Wool*
Barley	1.000							
Cheese	0.136	1.000						
	(0.001)							
	555							
Hay	0.088	0.118	1.000					
	(0.037)	(0.011)						
	566	466						
Oats	0.629	0.160	0.141	1.000				
	(0.000)	(0.002)	(0.001)					
	656	553	562					
Peas	0.593	0.091	0.090	0.680	1.000			
	(0.000)	(0.034)	(0.039)	(0.000)				
	624	540	532	621				
Suet	0.072	0.170	0.021	0.074	0.110	1.000		
	(0.083)	(0.000)	(0.640)	(0.077)	(0.009)			
	576	528	496	572	558			
Wheat	0.670	0.192	0.119	0.581	0.530	0.033	1.000	
	(0.000)	(0.000)	(0.004)	(0.000)	(0.000)	(0.430)		
	663	560	568	658	626	580		
Wool	0.115	–0.058	0.022	0.087	0.076	0.005	0.093	1.000
	(0.006)	(0.185)	(0.617)	(0.029)	(0.065)	(0.916)	(0.020)	
	629	528	536	625	592	549	634	

Note
Each correlation coefficient is computed using all available data on a given pair of commodities.

Table 2.4 Partial correlations of price changes

	Barley	*Cheese*	*Hay*	*Oats*	*Peas*	*Suet*	*Wheat*	*Wool*
Barley	1.000							
Cheese	–0.002	1.000						
	(0.976)							
Hay	0.069	0.121	1.000					
	(0.165)	(0.014)						
Oats	0.351	0.092	0.011	1.000				
	(0.000)	(0.063)	(0.820)					
Peas	0.213	–0.060	0.016	0.399	1.000			
	(0.000)	(0.229)	(0.750)	(0.000)				
Suet	0.057	0.134	–0.048	–0.057	0.097	1.000		
	(0.253)	(0.007)	(0.338)	(0.250)	(0.050)			
Wheat	0.418	0.126	–0.003	0.078	0.106	–0.089	1.000	
	(0.000)	(0.013)	(0.952)	(0.115)	(0.033)	(0.073)		
Wool	0.035	–0.173	–0.014	0.024	0.035	0.039	0.08	1.000
	(0.478)	(0.141)	(0.777)	(0.633)	(0.482)	(0.436)	(0.876)	

Note
Number of observations: 408.

investigates the results of estimating price interactions using a VAR, and shows how analysis of the residuals from the VAR can provide useful insights. It also highlights the practical limitations of the approach where price data is concerned.

By assuming that prices interact only with a lag, a VAR regression avoids the identification problem noted above. A simple economic explanation of such lags is that prices are administered by producers, so that at the start of each year producers set a price which is not changed until the following year. This means that producers cannot react to the prices of other commodities except with a lag. Each producer adopts a pricing rule by which they relate their administered price to the history of prices up to that date. The rule they adopt reflects whether there is a production lag, and whether they can hold over inventory from one year to the next. The various possibilities are considered in the Appendix A.1, where the implications for prices of different modes of producer behaviour are spelled out. This analysis only considers models without inventory, however; inventory-holding raises complex issues relating to price speculation that lie outside the scope of this chapter.

Suppose that all the producers of a given commodity follow the same rule; then the behaviour of prices simply mirrors that rule. If current price is always set with reference to previous prices, then whatever rule is chosen will be directly reflected in the way that prices behave. Although markets will not be in equilibrium, the disequilibrium will be reflected in quantities rather than prices, i.e. in shortages, or surpluses, or in inventory adjustment. If the rule is linear in previous prices then whatever rule is used is a special case of a general rule in which the current price of each commodity is a linear function of the lagged prices of all commodities.

The application of the rule each period involves a random shock. Under suitable conditions the parameters of the general rule can be estimated by linear regression. The residuals of the regression will reflect the properties of the random shocks. If the shocks affecting different commodities are uncorrelated, then the residuals should be uncorrelated. If different commodities are subjected to common shocks, however, then they may well be correlated. If a single shock has a similar impact on all commodities, for example, then the residuals will all tend to be positively correlated with each other.

An appropriate estimation procedure in this case involves two well-known techniques: a structural VAR regression, followed by a principal components analysis of the residuals. The principal components analysis is particularly useful for identifying whether there may be more than one important common shock in the system. The steps are:

1 Regress prices on lagged own-prices and lagged other-prices and a trend. Analyse the cross-price impacts using Table A.1 in the Appendix.
2 Examine the correlations between the residuals.
3 Calculate the principal components derived from the covariance matrix of residuals.

The results are reported in Table 2.5. The columns refer to the eight commodities whose current prices form the dependent variables in the separate regressions. The independent variables are listed in the rows of the table. These are in four blocks: the first comprises the prices of all eight commodities lagged one year, the second comprises all eight prices lagged two years and the third comprises all eight prices lagged three years. The fourth block comprises exogenous variables that potentially impact on all markets; in the present context exogenous factors are simply proxied by a linear time trend.

A maximum lag of three years was selected because investigation revealed that it is only in the fourth year that the coefficients tend to become almost entirely insignificant. Estimating longer lags would reduce the degrees of freedom because there are missing observations scattered throughout the data. Missing data were not interpolated because the interpolation method employed could influence the estimated price dynamics. Because the current price of each variable always appears as a dependent variable, and never as an independent variable, each regression can be estimated independently using OLS. To reduce the degrees of freedom lost through missing observations, each regression was estimated independently rather than as part of a VAR package.

The results are presented using a standard format where the significance level is reported as a probability value under each estimated coefficient. Measures of overall goodness of fit are reported at the bottom of the table. These need to be interpreted with caution, as they relate to regressions involving levels of price rather than changes in levels, and so include a large amount of variation that is explained simply by similarities in long-term trends.

The results show that each price is significantly influenced by its own previous price, and that previous prices as a whole have an important impact on the current price. For each commodity the sum of the own-price coefficients is less than one, which indicates stability in the system, notwithstanding the earlier ADF results; the difference arises because a standard ADF test does not include the wide range of relevant variables included in the present regressions.

The results suggest that substitution is widespread and that complementarity is relatively rare. Given that all the commodities can be classified as food or agricultural, this is not surprising. The results also suggest significant asymmetries in the price responses. This is particularly noticeable for the one year lags. Thus the price of wheat has a significant positive influence on the prices of barley, cheese, hay, oats and peas, but is itself only influenced by the price of peas. Peas and oats, on the other hand, tend to be influenced by other prices (cheese, hay, oats and wheat) but not to have much influence on other prices. The impacts of prices lagged two periods are more symmetric, and are generally negative, while for three years lags the impacts are largely insignificant.

In terms of Table A.1, these results suggest that, if prices are indeed administered as the model supposes, then model 6 affords the best account of the results. The combination of significant positive coefficients on one year lags, significant negative coefficients on two years lags and insignificant coefficients on three years lags suggest that there are persistent shocks but no production lags. It also

Table 2.5 Estimation of administered pricing model (VAR)

	Barley	*Cheese*	*Hay*	*Oats*	*Peas*	*Suet*	*Wheat*	*Wool*
Constant	–0.002	–0.113	3.66	–1.812	–0.186	1.024	0.456	1.964
	(0.998)	(0.812)	(0.001)	(0.010)	(0.844)	(0.040)	(0.626)	(0.001)
Barley Lag 1	0.344	–0.063	–0.305	0.048	0.141	–0.004	–0.121	0.063
	(0.001)	(0.214)	(0.001)	(0.521)	(0.123)	(0.937)	(0.131)	(0.349)
Cheese Lag 1	0.119	0.432	0.082	0.235	0.282	0.048	0.034	0.142
	(0.224)	(0.000)	(0.518)	(0.005)	(0.008)	(0.477)	(0.749)	(0.123)
Hay Lag 1	0.099	0.054	0.355	0.122	0.179	0.006	0.073	0.021
	(0.114)	(0.079)	(0.000)	(0.016)	(0.004)	(0.846)	(0.294)	(0.513)
Oats Lag 1	0.272	–0.034	0.243	0.407	0.248	0.074	0.075	0.024
	(0.009)	(0.541)	(0.031)	(0.000)	(0.014)	(0.230)	(0.489)	(0.776)
Peas Lag 1	0.036	0.073	–0.006	–0.021	0.270	0.050	0.167	0.040
	(0.620)	(0.108)	(0.948)	(0.736)	(0.004)	(0.229)	(0.036)	(0.493)
Suet Lag 1	–0.145	0.086	0.021	–0.056	–0.160	0.421	–0.193	0.010
	(0.065)	(0.083)	(0.842)	(0.461)	(0.072)	(0.000)	(0.030)	(0.872)
Wheat Lag 1	0.362	0.101	0.170	0.282	0.310	0.114	0.864	–0.130
	(0.000)	(0.009)	(0.025)	(0.000)	(0.000)	(0.801)	(0.000)	(0.008)
Wool Lag 1	0.044	0.133	0.047	–0.021	0.028	0.009	–0.061	0.651
	(0.515)	(0.000)	(0.495)	(0.711)	(0.685)	(0.866)	(0.364)	(0.000)
Barley Lag 2	–0.060	–0.012	0.131	–0.102	–0.152	–0.008	–0.091	–0.013
	(0.571)	(0.839)	(0.244)	(0.223)	(0.134)	(0.889)	(0.309)	(0.873)
Cheese Lag 2	–0.062	0.156	–0.066	–0.044	–0.098	0.073	–0.008	–0.014
	(0.563)	(0.028)	(0.601)	(0.627)	(0.366)	(0.312)	(0.945)	(0.887)
Hay Lag 2	–0.130	–0.042	0.073	–0.091	–0.110	–0.067	–0.001	–0.091
	(0.039)	(0.167)	(0.300)	(0.067)	(0.057)	(0.051)	(0.983)	(0.017)
Oats Lag 2	0.042	0.091	0.005	0.205	–0.047	0.054	0.169	0.074
	(0.717)	(0.111)	(0.962)	(0.040)	(0.638)	(0.385)	(0.071)	(0.410)
Peas Lag 2	–0.187	–0.123	–0.119	–0.196	–0.068	–0.004	–0.223	–0.170
	(0.006)	(0.015)	(0.191)	(0.001)	(0.408)	(0.925)	(0.007)	(0.004)

Suet Lag 2	–0.064	–0.040	0.047	–0.076	–0.167	0.106	0.051	–0.194
	(0.387)	(0.464)	(0.663)	(0.320)	(0.084)	(0.078)	(0.543)	(0.006)
Wheat Lag 2	–0.238	0.023	0.029	–0.270	–0.210	0.057	–0.367	0.096
	(0.004)	(0.645)	(0.733)	(0.000)	(0.012)	(0.217)	(0.000)	(0.134)
Wool Lag 2	0.021	–0.058	–0.020	0.035	0.040	0.004	–0.016	0.113
	(0.767)	(0.186)	(0.808)	(0.552)	(0.628)	(0.949)	(0.826)	(0.256)
Barley Lag 3	0.177	0.004	–0.030	0.073	0.131	0.004	–0.008	–0.005
	(0.061)	(0.947)	(0.761)	(0.337)	(0.197)	(0.930)	(0.922)	(0.937)
Cheese Lag 3	–0.021	0.268	0.117	–0.015	0.014	–0.050	0.158	0.031
	(0.828)	(0.000)	(0.226)	(0.848)	(0.879)	(0.388)	(0.183)	(0.681)
Hay Lag 3	0.026	–0.019	0.034	0.032	–0.4	–0.008	–0.036	0.025
	(0.667)	(0.543)	(.640)	(0.514)	(0.952)	(0.832)	(0.467)	(0.34)
Oats Lag 3	0.026	0.081	0.018	0.157	–0.051	–0.046	0.033	–0.071
	(0.798)	(0.221)	(0.874)	(0.035)	(0.600)	(0.442)	(0.733)	(0.398)
Peas Lag 3	0.068	0.020	–0.062	0.042	0.076	–0.029	0.143	0.062
	(0.350)	(0.679)	(0.503)	(0.514)	(0.368)	(0.522)	(0.054)	(0.270)
Suet Lag 3	0.197	–0.0142	0.006	0.148	0.229	0.196	0.016	0.144
	(0.008)	(0.436)	(0.952)	(0.039)	(0.008)	(0.001)	(0.846)	(0.052)
Wheat Lag 3	0.060	–0.067	0.080	0.093	0.122	–0.013	0.157	0.085
	(0.423)	(0.117)	(0.291)	(0.109)	(0.101)	(0.779)	(0.014)	(0.126)
Wool Lag 3	–0.048	–0.012	–0.052	–0.040	–0.077	0.033	0.051	0.058
	(0.419)	(0.734)	(0.450)	(0.402)	(0.267)	(0.459)	(0.381)	(0.280)
Time	0.002	0,000	0.007	0.001	0.000	0.001	0.002	–0.001
	(0.045)	(0.421)	(0.000)	(0.324)	(0.982)	(0.038)	(0.050)	(0.220)
R^2	0.971	0.990	0.973	0.981	0.967	0.979	0.973	0.940
Adj R^2	0.969	0.990	0.971	0.80	0.965	0.978	0.971	0.936
F	451.67	1351.1	470.79	696.11	392.84	625.01	484.46	208.02
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
DW	2.003	1.902	1.870	2.022	1.976	2.080	1.941	1.939
No.obs	363	353	354	363	361	357	363	357

suggests that producer price expectations are stationary, i.e. they depend only on the previous year's price and not on prices before that. If there were production lags with stationary expectations there would be a tendency for rises in prices to alternate with falls in prices, and there is no evidence in the data to support this.

The existence of persistent shocks suggests that prices are not purely influenced by transitory shocks, but by changes whose legacy lasts for several periods. These could represent changes in tastes, technology and institutions, but the results provide no specific clues as to what they might actually be. The absence of production lags needed to be treated with caution, however. It is possible that lags appear to have no impact simply because commodities are easily stored, so that short-term price fluctuations are being smoothed out by the accumulation and decumulation of inventory.

Table 2.6 indicates high levels of correlation between the residuals from the commodity regressions. This could simply indicate the consequences of omitting current-period price interaction from the model, and the results presented below suggest that this may indeed be the explanation. If it is not the explanation, then it is likely that the answer lies in common shocks instead. Table 2.7 presents the results of a principal components analysis, which identifies a set of orthogonal shocks which together account for residual covariances. Each shock corresponds to one of the unobservable factors in the table. Factor 1 alone accounts for over 40 per cent of all the residual variation, and all of the commodities carry a positive loading. This suggests that there may be a common shock that affects the entire economy, such as a monetary shock or a taxation shock. On closer examination, however, it can be seen that the loadings are high on barley, oats, peas and wheat – the 'big four commodities' previously identified. This suggests that if there is a common shock then it related to these specific commodities, all of which tend to be grown on the same kinds of soils and in similar parts of the country. It is therefore possible that the common shock may have something to do with the weather. On the other hand, if there is no common shock, then the results can simply be explained by the fact that prices are not administered, but are competitively determined, and these four commodities just happen to be close substitutes for each other.

2.6 Competitive prices: a model of hierarchical markets with commodity-specific shocks

This leads directly to the final stage of the analysis, which is to estimate a competitive multi-market model. This model postulates that each year there is general equilibrium across these eight commodity markets (but not necessarily across the entire economy). The model allows for mutual interactions between the current prices of all commodities, but limits the number of possible scenarios using identification restrictions (for a general discussion see Sargan, 1988). The eight commodities are arranged in a hierarchy in which the prices of higher-ranked commodities influence the prices of lower-ranked commodities but the prices of lower-ranked commodities do not influence the prices of the

Table 2.6 Correlations between the residuals of the administered price regression

	Barley	*Cheese*	*Hay*	*Oats*	*Peas*	*Suet*	*Wheat*	*Wool*
Barley	1.000							
Cheese	0.209	1.000						
	(0.000)							
Hay	0.158	0.118	1.000					
	(0.004)	(0.032)						
Oats	0.704	0.240	0.182	1.000				
	(0.000)	(0.000)	(0.001)					
Peas	0.536	0.196	0.139	0.593	1.000			
	(0.000)	(0.000)	(0.011)	(0.000)				
Suet	0.059	0.211	0.058	0.098	0.119	1.000		
	(0.285)	(0.000)	(0.294)	(0.074)	(0.030)			
Wheat	0.597	0.191	0.118	0.469	0.477	0.035	1.000	
	(0.000)	(0.001)	(0.032)	(0.000)	(0.000)	(0.529)		
Wool	0.126	0.039	–0.056	0.124	0.078	0.087	–0.008	1.000
	(0.022)	(0.478)	(0.310)	(0.024)	(0.157)	(0.113)	(0.892)	

Note
Number of observations: 332.

Table 2.7 Principal components of the variances of the residuals of the administered price regression

Component	*1*	*2*	*3*	*4*	*5*	*6*	*7*
Eigenvalue	1.324	0.601	0.378	0.281	0.230	0.216	0.144
Cumulative proportion of variance	0.403	0.586	0.701	0.787	0.857	0.923	0.967
Loadings of eigenvectors							
Barley	0.509	–0.118	0.024	0.258	–0.208	0.459	–0.175
Cheese	0.110	0.025	0.050	–0.063	0.380	0.220	0.875
Hay	0.229	0.960	0.109	0.067	–0.027	–0.087	–0.026
Oats	0.434	–0.066	0.111	–0.102	–0.224	0.445	0.045
Peas	0.506	–0.123	0.039	–0.696	–0.085	–0.456	0.011
Suet	0.054	0.019	0.179	–0.241	0.794	0.276	–0.447
Wheat	0.474	–0.134	–0.370	0.531	0.352	–0.417	–0.010
Wool	0.061	–0.161	0.895	0.299	0.007	–0.277	0.038

higher-ranked ones. The prices of commodities at the same level interact mutually, but commodities only interact in pairs.

The hierarchical ranking cannot be arrived at on the basis of the data alone. The researcher must also apply general historical knowledge, e.g. it is to be expected that internationally traded products will lie at the top of the hierarchy, and locally traded products towards the bottom. Data analysis has a role, however. It is likely that pairs of commodities whose prices are highly correlated with each other but only modestly correlated with others will form interacting pairs. Lagged interactions are also relevant; e.g. if the lagged value of one price influences the current value of some other price, and the lagged value of that price influences the current value of the first price then it may be conjectured that the current prices of the two commodities interact. The relative volatility of prices is also relevant; if the price of one commodity is much more volatile than the price of another then some of the variation in the more volatile price may be due to the influence of the other price, suggesting that a more volatile price is more likely to depend on a less volatile price than the other way round; this also suggests that interaction is most likely when the volatilities of two prices are similar.

The key steps are:

1 Develop a hierarchy involving pairs of products, based on prior assumptions and reflection on previous findings.
2 Follow the procedures for estimating the hierarchical model as set out in the appendix. The result is a set of price regressions in which pairs of prices interact within an underlying recursive structure. Apart from the interactions, the regressions can be interpreted in the usual way. As a result of the estimation method used, all residuals are uncorrelated so, unlike the administered price regression it is unnecessary to analyse them; the measures of goodness of fit need to be interpreted with care, however, because they are measured net of the fit accounted for by interactions between the selected pairs of variables. The results can be analysed using Table A.1.

The hierarchy proposed is illustrated in Figure 2.1. It places wool at the top. The price of wool (rank 1) influences the prices of barley and oats, which interact with each other (rank 2). These three prices then influence another interacting pair: wheat and peas (rank 3). The five commodities then influence the prices of the other three (rank 4). These three comprise an interacting pair of cheese and hay, and a solitary commodity, suet.

The results are presented in Table 2.8. Once again, they suggest that for these specific food and agricultural products substitution is widespread, and that complementarity is quite exceptional. Comparing the coefficients for different lagged impacts suggests that most of the impacts occur in the current period. Where no impact is assumed in the current period, it is quite common for a one-year lagged price to appear significant, which suggests that the assumption of no current impact may be too strong.

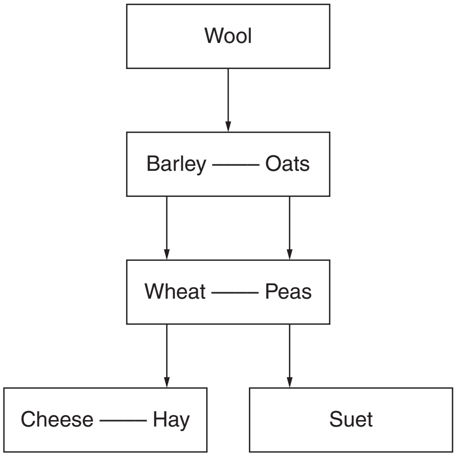

Figure 2.1 Schematic diagram of the hierarchical model.

The positive and significant coefficients on lagged own prices confirm the earlier suggestion of persistence of shocks, but apart from this it is difficult to discriminate between the alternative models reviewed in Table A.1. Many of the cross-commodity lagged effects are positive, rather than negative as predicted by most of the models, and further work is required to investigate this issue. Overall, however, the very strong current-price effects revealed by the hierarchical model suggest that this model is superior to the structural VAR. Future work should therefore be addressed to refining the hierarchical model rather than to refining the VAR.

2.7 Extensions

The lack of long-run data on exogenous variables means that exogenous factors need to be proxied by time trends and structural breaks. Time trends have been included in the preceding regressions to proxy for long-run changes in money supply and GDP. Structural breaks could be introduced to reflect the impacts of known events, such as the imposition of commodity taxes, changes in tariffs, the creation or dissolution of state monopolies, the outbreak of civil war and the enforcement of regulations affecting demand, such as sumptuary laws.

The residuals from the hierarchical model can be used as measures of commodity-specific random shocks. Examining the variances of the residuals makes possible a comparison of the magnitude of shocks affecting individual commodity markets. By analysing the variances of the residuals rather than the variances of the prices themselves it is possible to allow for the fact that some of the impacts of shocks will be transmitted to other commodities and that shocks to other commodities will impact on the commodity price.

Squaring each residual from a commodity price regression generates a time series of instantaneous annual volatility. The derived series of squared deviations can then be regressed on a time trend to test for long-term changes in variance (i.e. heteroskedasticity), and regressed on dummy variables to test for structural breaks in volatility.

Similar exercises can be carried out on the residuals from the administered price regressions. Using these residuals it is possible to examine trends in the covariation of the residuals and to test their constancy over time. The hierarchical model cannot be used for this purpose because its estimation procedures assume that these covariances are constant.

2.8 Conclusions

This chapter has set out a new methodology for analysing price behaviour in a long-run historical context. The methodology embraces conventional methods involving descriptive statistics, structural VARs and principal components, and a new approach based on the principles of structural econometric modelling. The theory of identification is extended to include mutually interacting prices embedded in a recursive structure, which leads to a model involving a hierarchy of markets. A specific hierarchy has been proposed for England, which interprets wool as a generally pervasive influence on prices, although not a strong influence on any particular price. The most important interacting pair of prices comprises barley and oats, whose behaviour then influences another interacting pair, namely wheat and peas.

The next step is to develop a hierarchical model embracing all 24 commodities represented in the database, and to repeat the exercise.

There is good reason to believe that the substitutability of certain commodities can usefully be regarded as a constant of the economic system, although the degree of substitutability may well vary from one country to another. Much conventional historical narrative does not engage with substitutability, because it ignores the fact that substitutability diffuses shocks that originate in one part of the economy to other parts of the economy, so that causes and effects can be difficult to match up. In some historical interpretations one particular substitution effect may be singled out as being of paramount importance, which may also result in distortion, as the chosen effect is exaggerated while other substitution effects continue to be ignored.

Although there has been considerable research on the prices of individual commodities, as noted at the outset, there has been very little historical research on the price system as a whole. As a result, there remain great opportunities to do more work, and it is hoped that the techniques set out in this chapter will encourage others to pursue this research agenda.

Table 2.8 Estimation of hierarchical market model

	Wool	*Barley*	*Oats*	*Wheat*	*Peas*	*Cheese*	*Hay*	*Suet*
Wool		0.091	0.072	–0.098	0.013	0.007	–0.126	0.065
		(0.105)	(0.059)	(0.067)	(0.827)	(0.866)	(0.098)	(0.164)
Barley			0.405	0.515	0.210	0032	0.083	–0.035
			(0.000)	(0.000)	(0.006)	(0.555)	(0.399)	(0.461)
Oats		0.405		0.075	0.492	0.090	0.146	0.024
		(0.000)		(0.360)	(0.000)	(0.148)	(0.202)	(0.731)
Wheat					0.108	0.037	–0.009	–0.016
					(0.000)	(0.357)	(0.909)	(0.733)
Peas				0.108		0.029	0.056	0.058
				(0.000		(0.498)	(0.565)	(0.178)
Cheese							0.039	0.220
							(0.083)	(0.003)
Hay						0.039		0.021
						(0.083)		(0.524)
Suet								
Wool	0.651	–0.034	–0.100	–0.004	0.045	0.126	0.116	–0.076
Lag 1	(0.000)	(0.610)	(0.062)	(0.959)	(0.476)	(0.013)	(0.197)	(0.281)
Barley	0.063	0.314	–0.099	–0.315	0.058	–0.066	–0.331	–0.002
Lag 1	(0.349)	(0.003)	(0.095)	(0.001)	(0.501)	(0.218)	(0.002)	(0.977)
Oats	0.024	0.121	0.307	–0.102	–0.032	–0.103	0.118	0.059
Lag 1	(0.776)	(0.137)	(0.000)	(0.217)	(0.733)	(0.096)	(0.278)	(0.384)
Wheat	–0.130	0.245	0.141	0.605	0.022	0.011	0.082	–0.045
Lag 1	(0.008)	(0.000)	(0.001)	(0.000)	(0.748)	(0.829)	(0.391)	(0.399)
Peas	0.040	0.051	–0.037	0.143	0.229	0.071	–0.018	0.033
Lag 1	(0.493)	(0.379)	(0.445)	(0.023)	(0.002)	(0.110)	(0.859)	(0.424)
Cheese	0.142	–0.002	0.164	–0.099	0.145	0.389	–0.002	–0.060
Lag 1	(0.123)	(0.976)	(0.012)	(0.208)	(0.059)	(0.000)	(0.990)	(0.459)
Hay	0.021	0.049	0.084	0.014	0.085	0.020	0.301	–0.021
Lag 1	(0.513)	(0.304)	(0.008)	(0.771)	(0.150)	(0.568)	(0.000)	(0.491)
Suet	0.010	–0.119	0.005	–0.082	–0.084	0.103	0.092	0.368
Lag 1	(0.872)	(0.051)	(0.933)	(0.241)	(0.214)	(0.037)	(0.377)	(0.000)
Wool	0.113	0.024	0.030	–0.083	0.004	–0.066	–0.023	0.024
Lag 2	(0.256)	(0.700)	(0.542)	(0.206)	(0.950)	(0.210)	(0.812)	(0.727)
Barley	–0.013	–0.026	–0.079	–0.031	–0.080	0.001	0.097	–0.010
Lag 2	(0.873)	(0.768)	(0.227)	(0.683)	(0.356)	(0.980)	(0.408)	(0.868)
Oats	0.074	–0.051	0.182	0.165	–0.162	0.069	0.057	0.006
Lag 2	(0.410)	(0.579)	(0.016)	(0.050)	(0.074)	(0.269)	(0.608)	(0.930)

Wheat	0.096	–0.129	–0.182	–0.204	–0.004	0.077	0.114	0.064
Lag 2	(0.134)	(0.070)	(0.000)	(0.003)	(0.950)	(0.134)	(0.186)	(0.177)
Peas	–0.170	–0.087	–0.102	–0.145	0.097	–0.090	–0.081	0.073
Lag 2	(0.004)	(0.124)	(0.023)	(0.034)	(0.229)	(0.070)	(0.391)	(0.103)
Cheese	–0.014	–0.090	–0.036	0.021	–0.028	0.149	–0.074	0.021
Lag 2	(0.887)	(0.313)	(0.611)	(0.815)	(0.769)	(0.045)	(0.573)	(0.788)
Hay	–0.091	–0.083	–0.032	0.050	–0.034	–0.027	0.109	–0.031
Lag 2	(0.017)	(0.071)	(0.307)	(0.219)	(0.399)	(0.414)	(0.144)	(0.450)
Suet	–0.194	–0.006	–0.029	0.129	–0.129	0.018	0.053	0.147
Lag 2	(0.006)	(0.912)	(0.624)	(0.045)	(0.097)	(0.760)	(0.626)	(0.018)
Wool	0.058	–0.037	–0.024	0.140	–0.054	0.000	–0.004	0.015
Lag 3	(0.280)	(0.475)	(0.543)	(0.007)	(0.375)	(0.994)	(0.955)	(0.754)
Barley	–0.005	0.144	0.003	–0.112	0.055	0.002	–0.094	0.023
Lag 3	(0.937)	(0.065)	(0.962)	(0.073)	(0.483)	(0.970)	(0.363)	(0.670)
Oats	–0.071	–0.043	0.142	–0.030	–0.115	0.064	0.032	–0.050
Lag 3	(0.398)	(0.622)	(0.010)	(0.701)	(0.207)	(0.368)	(0.787)	(0.408)
Wheat	0.085	0.019	0.068	0.111	0.045	–0.096	0.073	–0.016
Lag 3	(0.126)	(0.745)	(0.312)	(0.029)	(0.426)	(0.022)	(0.323)	(0.745)
Peas	0.062	0.044	0.009	0.139	0.022	0.008	–0.070	–0.070
Lag 3	(0.270)	(0.442)	(0.857)	(0.014)	(0.773)	(0.865)	(0.470)	(0.129)
Cheese	0.031	0.029	0.019	0.212	–0.030	0.278	0.125	–0.091
Lag 3	(0.681)	(0.743)	(0.777)	(0.013)	(0.692)	(0.000)	(0.252)	(0.173)
Hay	0.025	–0.036	0.040	–0.013	–0.015	–0.016	0.022	0.004
Lag 3	(0.534)	(0.468)	(0.283)	(0.744)	(0.731)	(0.581)	(0.765)	(0.911)
Suet	0.144	0.135	0.051	–0.147	0.101	–0.073	–0.048	0.208
Lag 3	(0.052)	(0.023)	(0.364)	(0.014)	(0.120)	(0.214)	(0.624)	(0.001)
Constant	1.964	0.483	–1.939	0.866	0.707	0.124	4.409	0.686
	(0.001)	(0.618)	(0.001)	(0.267)	(0.331)	(0.807)	(0.001)	(0.165)
Time	–0.001	0.002	–0.000	0.000	–0.001	–0.000	0.006	0.001
	(0.220)	(0.059)	(0.881)	(0.661)	(0.297)	(0.752)	(0.000)	(0.097)
R^2	0.940	0.947	0.973	0.982	0.975	0.990	0.973	0.981
Adj R^2	0.936	0.943	0.971	0.980	0.973	0.989	0.971	0.978
F	208.02	228.71	463.733	619.82	457.35	1013.7	371.36	470.99
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
DW	1.939	1.952	1.983	1.994	1.959	1.893	1.876	2.110
No.obs	357	357	357	355	355	337	337	332

Note

The regression is estimated in three stages. For each pair of interacting prices, partial variances and covariances are computed, conditional on all relevant variables dictated by the assumed hierarchical structure, and their interaction coefficient is calculated. For each commodity a compensated price is computed by subtracting from its price the price of the interacting commodity, weighted by the interaction coefficient; this compensated price is then regressed on all relevant variables, giving the coefficients reported above. Finally, the interaction coefficients are introduced into the table, showing a significance determined by the partial correlation from which they are calculated. The measures of goodness of fit reported in the table do not take account of the component of the variance of an uncompensated price that is explained by the interacting price.

Appendix: two formal models of multi-market price determination

A.0 General observations on the modelling of price interactions

The concept of substitution

The basic notion of substitution is that the demand for a commodity increases and/or its supply decreases when the price of some other commodity increases. This is because consumers switch to the cheaper commodity, while suppliers divert production to the more expensive commodity which is more profitable to produce. Thus when commodities are substitutes an increase in the price of one commodity tends to increase the price of the other. Substitution is generally measured in terms of proportional prices, i.e. by the percentage increase in the price of one commodity generated by a unit percentage increase in the price of the other. The larger the percentage increase, the closer (or stronger) the substitution.

Substitution is normally reciprocal, in the sense that if commodity 1 is a substitute for commodity 2 then commodity 2 is also a substitute for commodity 1. Substitution is not necessarily symmetrical, however: the impact of the price of commodity 1 on the price of commodity 2 is not necessarily the same as the impact of the price of commodity 2 on the price of commodity 1. If one commodity is versatile, for example, and has many uses, while the other is specialised and has only a single use then the price of the specialised commodity is likely to be more sensitive to the price of the versatile commodity than the other way round. A reduction in the price of a versatile commodity can undermine the entire demand for a specialised commodity, whereas a reduction in the price of a specialised commodity affects only one of several uses for a versatile commodity and therefore has little impact on its price.

Income effects also generate asymmetries. Suppose, for example, that one commodity, such as wheat, accounts for a high proportion of consumer expenditure, while another commodity, such as peas, accounts for only a low proportion. An increase in the price of wheat may substantially reduce real consumer incomes, and thereby lead to a substantial fall in the demand for peas, which would not occur if the reduction in income were fully compensated through lower taxes. If income effects were compensated then the price of peas might rise in line with the price of wheat, but if they are not then it is likely to fall.

Complementarity is the opposite of substitution. Commodity 1 is a complement to commodity 2 when an increase in its price reduces the price of commodity 2. Bread and butter are commonly said to be substitutes, because butter is usually spread on slices of bread. If the price of bread goes up, the demand for butter will tend to go down, and this will in turn drive down the price of butter.

The concepts of substitutability and complementarity apply to pairs of goods which are either both inputs to consumption, both inputs to production or both outputs of production. They do not apply to pairs of inputs and outputs, whose prices generally move together simply because the price of the input is a component of the cost of the output.

The nature of shocks

The variation of prices is driven by exogenous shocks, and different types of shock have different implications.

Quantity or price. Shocks often impact on price through their impacts on quantities (e.g. bad weather may destroy the annual crop of some commodity, pushing up the price), but they can also influence price directly (e.g. monetary shocks that cause inflation).

Supply or demand. Shocks may impact on either demand or supply, e.g. changes in tastes generally impact on demand, while changes in technology or labour supply tend to impact on supply. In practice it is difficult to distinguish demand shocks from supply shocks, e.g. an increase in price could be due either to an increase in demand or a shortage of supply.

General or specific. A shock may influence a single commodity or it may influence several commodities simultaneously. Thus a change in consumer tastes may generate a commodity-specific quantity shock, and an increase in an excise tax may generate a commodity-specific price shock. By contrast, war may generate a general quantity shock, through disruption of production, while depreciation of the currency may generate a general price shock. General shocks that generate equi-proportional effects on all commodities are sometimes called 'common shocks'.

Transitory or persistent. Shocks may influence price only in the period in which they occur, but they can also generate legacy effects, e.g. the impact of the shock may decay slowly over subsequent periods. When commodities are perishable, for example, stocks cannot be held over from one period to the next, and so process in successive periods tend to be independent of each other; when commodities are durable, however, excess supplies can be stored for sale the following period, so that an unexpected fall in demand one period may lead to a fall in prices, not only in the current period, but in subsequent periods too. Persistent shocks can generate stochastic trends, in which the price in any period depends upon the previous price because the previous price embodies shocks that continue to influence price in the current period. Price shocks persist indefinitely if price formation follows a 'unit root' process in which the expected price in any period is equal to the previous price. It is commonly asserted that the daily or weekly prices of financial assets follow a random walk, which is a special type of unit root process, but this does not mean that ordinary commodity prices follow a similar process too.

Observable or unobservable. Price data is much easier to interpret when exogenous shocks are observable. In an historical context, observability means that they are not only measureable and recordable, but that the records survive and can be accessed through archives. Unfortunately, reliable data on exogenous shocks is scarce. Long-period weather data is only available for large areas, such as the Northern Hemisphere, within which localised variation is considerable. Money supply is also difficult to measure because at different times it has been dominated by different forms of money, such as gold, silver, coin, bullion, bank notes and cheque accounts. Long-period annual measurements of national incomes are only just becoming available. As a result, there are many different types of unobservable shock that impinge on the price-formation process, and with data on prices alone it is difficult to distinguish between them. The best that can often be achieved is to create artificial variables such as time trend variables and dummy variables indicating structural breaks.

Administered prices versus negotiated prices

Prices may be administered or negotiated. Administered prices are set at the start of each period and remain in force throughout the period. Prices are normally administered by producers rather than consumers, especially when there are many consumers but few producers, and when producers are better informed about competitive conditions. In addition, prices may sometimes be constrained by regulators (e.g. a statutory maximum price for a standard loaf). With negotiated prices, however, there is no designated price-maker. It is the impersonal forces of the market that make prices, rather than an individual such as a dominant producer or statutory regulator. With negotiated prices, each trader can adjust their price quotations to reflect the offers being made by others. Haggling continues, it is assumed, until all markets are in equilibrium; at this point contracts are exchanged and trade takes place. The price recorded in each market is therefore the equilibrium price.

With negotiated prices, the impact of a shock to any single market is distributed across the entire system. A shock to the market for barley, for example, affects not only on the price of barley, but also the price of wheat and oats. If barley, wheat and oats are substitutes then a rise in the price of barley will increase the demand and/or reduce the supply in the other two markets (provided that supply is elastic), causing prices in those markets to rise as well. These changes induced by the change in the price of barley then feed back onto the price of barley. Price adjustment spreads like an expanding wave-front through the market system and, provided the system is stable, the wave is reflected back upon itself with diminishing intensity until a general equilibrium is attained. In this equilibrium the initial shock to the price of barley has been spread across the system so that all prices have potentially been affected. The impact on the price of barley is less than it would have been in the absence of substitute commodities, but the impact on the prices of the other commodities is greater than it would otherwise have been. With negotiated prices, therefore, the prices of substitutes tend to move in the same direction in the same period, while the prices of complements move in opposite directions instead.

Suppose now that the price in each market is administered by the dominant producer in that market, without consultation with price-makers in other markets. At the time an administered price is set, each price-maker is therefore ignorant of prices other than their own; all they know is the prices that were set for other commodities in previous periods. Thus if two commodities are substitutes, the price of one will always react to the price of the other with a one-period lag; unlike negotiated prices, there will be no reaction between prices within a single period. If there is a specific shock to some particular market then the price in that market will respond while other prices remain unchanged.

Nevertheless, administered prices can move together. If shocks are general rather than specific, then current prices may move in the same direction, not because they react to each other, but because they respond in a similar way to the same general shock. General shocks acting on administered prices can generate price movements that resemble those of negotiated prices responding to specific shocks. If the prices of barley and oats move together, for example, the explanation may be either that shocks are specific, prices are negotiated and the commodities are substitutes, or that both prices respond in a similar way to the same general shock.

It is difficult to distinguish between competing explanations when shocks are unobservable. If an initial shock cannot be directly observed then it is difficult to know whether it is a demand shock or a supply shock, or a general shock rather than a specific shock. If it is a specific shock, it is even difficult to be sure in which market it originated. The only solution is to observe as many shocks as possible. In the absence of direct observations, time trends and dummy variables may be introduced instead, as indicated above.

A.1 A model of administered prices and general random shocks

Consider a set of commodities, indexed $i = 1, \ldots, N$, where at time t ($t = 1, \ldots, T$) the ith commodity has a money price p_{it} (expressed in natural logarithms). These commodities are a small subset of all the commodities traded in the economy.

- The price in each market is set by producers who specialise in supplying that market, all of whom use the same price-setting rule.
- Price is set at the beginning of each period, before demand and supply have been revealed, and remains at this level throughout the period.
- No inventories are held and so, depending on market conditions, some demand may remain unsatisfied or some supply may be wasted.
- Each producer knows the entire price history.
- Shocks may persist. The prices of financial assets are often assumed to follow random walks in which shocks persist indefinitely. For agricultural commodities and manufactures, however, shocks decay fairly rapidly over time. Nevertheless, they may induce some serial correlation in prices.
- Prices of other commodities are forecast using weighted averages of previous prices. When shocks persist indefinitely the most recent price is the best predictor, and when shocks are purely transitory an average of previous prices (i.e. an estimate of the mean) works best. When the impact of shocks decays slowly producers may compromise by using a weighted average of successive prices.

Producers formulate a target price, p^*_{it}, that depends on previous prices and on exogenous changes in non-price factors that occurred the previous period. Let z_{mt} ($m = 1, \ldots, M$) be the value of the mth observable non-price factor from the previous period that affects the administered price in period t; some of these factors may be proxied using a constant term and a time trend. Let v_{nt} be the value of the nth general shock that affects price in period t ($n=1,\ldots, N$); this shock is observed by the producers but is not recorded, and so is treated as an unobservable random shock in the model. Let the parameter g_{in} (i, $n=1,\ldots, N$) measure the impact of the nth unobservable shock on the ith commodity price; since shocks are general, any shock can, in principle, affect all prices.

Given the non-price factors, each producer relates the target price to the expected prices of the other commodities, p^e_{ijt} ($j=1,\ldots, i-1, i+1,\ldots, N$), which they assume to be a weighted average of past prices:

$$p^*_{it}=\sum_j b_{ij}p^e_{ijt}+\sum_m d_{im}z_{mt} \quad (j=1,\ldots, N) \tag{1.1}$$

where

$$p^e_{jt}=k_i p_{jt-1}+(1-k_i)p_{jt-2} \tag{1.2}$$

and $b_{ii}=0$, $0\leq k_i\leq 1$. The parameter k_i determines the weight given by the ith producer to the most recent price; expectations are stationary when $k_i=1$. If $b_{ij}>0$ then j is a substitute for i and if $b_{ij}<0$ it is a complement.

Each period price is adjusted partially towards its target value:

$$p_{it}=q_i(p^*_{it}-p_{it-1})+p_{it-1}+\sum_g g_{in}v_{nt}=q_i p^*_{it}+(1-q_i)p_{it-1}+\sum_g g_{in}v_{nt} \tag{2}$$

where $0<q_i\leq 1$ is the price adjustment parameter; if $q_i=1$ then adjustment is complete and price is set at the target level. Note that while the observable exogenous factors affect the target price the unobservable factors affect only the current price.

Substituting (1) into (2) expresses current price as a linear function of all prices, with a one-year lag on own-price and lags of up to two years on other prices:

$$p_{it}=(1-q_i)p_{it-1}+q_i k_i\sum_j b_{ij}p_{jt-1}+q_i(1-k_i)\sum_j b_{ij}p_{jt-2}+q_i\sum_m d_{im}z_{mt}+\sum_n g_{in}v_{nt} \tag{3}$$

If the exogenous variables are proxied by a constant term and a linear time trend then $\sum_m d_{im}z_{mt}=a_i+d_i t$.

Suppose that all general shocks are persistent, and their impact on administered price decays geometrically over time at the same rate, $h(0\leq h < 1)$; the limiting case $h=0$ corresponds to transitory shocks. The shocks then satisfy the first-order linear difference equation

$$v_{nt}=hv_{nt-1}+u_{nt} \quad (t=1,\ldots, T) \tag{4}$$

Taking weighted sums on both sides of equation (4) using the weights g_{in}, eliminating terms in v using equation (3), and grouping the remaining terms gives

$$\begin{aligned} p_{it}=&(1-q_i+h)p_{it-1}+q_i k_i\sum_j b_{ij}p_{jt-1}-h(1-q_i)\, p_{it-2}+q_i(1-(1+h)k_i)\sum_j b_{ij}p_{jt-2}+\\ &hq_i(1-k_i)\sum_j b_{ij}p_{jt-3}+\sum_m q_i d_{im}(z_{mt}-hz_{mt-1})+\sum_n g_{in}u_{nt} \end{aligned} \tag{5}$$

If the exogenous variables are proxied by a constant term and a linear time trend then $\sum_m q_i d_{im}(z_{mt} - hz_{mt-1}) = (1-h)(a_i + d_i t) - hd_i$.

Equation (5) is a special case of the general form

$$p_{it} = \sum_w \sum_j c_{wij} p_{jt-w} + \sum_m (c_{4im} z_{mt} + c_{5im} z_{mt-1}) + \sum_n g_{in} u_{nt} \tag{6}$$

which expresses the current price of each commodity as a linear function of all commodity prices lagged for up to three periods, a set of non-price exogenous variables lagged up to one period and a set of unobservable random shocks, each of which potentially affects all current prices.

To estimate the parameters of equation (6) it is convenient to assume that the shocks u_{nt} have zero means and constant variances, and are uncorrelated with each other, with previous prices and with the exogenous variables. Let $\mathbf{P} = [p_{it}]$ be a $(T-3) \times N$ matrix of observations on current prices; $\mathbf{X} = ([p_{it-1}], [p_{it-2}], [p_{it-3}], [z_{mt}], [z_{mt-1}])$ a $(T-3) \times (3N+2M)$ matrix of regressors; $\mathbf{C} = ([c_{1ij}], [c_{2ij}], [c_{3ij}], [c_{4im}], [c_{5im}],)'$ a $(3N+2M) \times N$ matrix of coefficients; $\mathbf{G} = [g_{in}]$ a $N \times N$ matrix of impact weights and $\mathbf{U} = [u_{nt}]$ a $(T-3) \times N$ matrix of random shocks. The assumptions imply that $E(\mathbf{U}) = \mathbf{0}$, $E(\mathbf{U'X}) = \mathbf{0}$ and $E(\mathbf{U'U})/N = \Omega$, where $\mathbf{0}$ represents a null vector or matrix and Ω is an $N \times N$ diagonal matrix of variances $\sigma_n{}^2$.

Equation (6) can be rewritten in the standard form for a simultaneous equations system:

$$\mathbf{P} = \mathbf{XC} + \mathbf{UG} \tag{7}$$

Pre-multiplying by $\mathbf{X}$ and taking expected values gives

$$\mathbf{C} = (\mathbf{X'X})^{-1}\, \mathbf{X'P} \tag{8}$$

Equation (8) provides estimators of the impacts on current prices of lagged prices and non-price exogenous variables.

Let $\mathbf{I}$ be the identity matrix; using equations (8) and (9) it can be shown that the sample covariance matrix $\mathbf{V} = \mathbf{P'}(\mathbf{I} - (\mathbf{X'X})^{-1}\ \mathbf{X'})'(\mathbf{I} - (\mathbf{X'X})^{-1}\ \mathbf{X'})\mathbf{P}$ has expected value

$$E(\mathbf{V}) = (\mathbf{UG})'(\mathbf{UG})/N = \mathbf{G'}\Omega\mathbf{G} \tag{9}$$

To identify the elements of $\mathbf{G}$, assume that the general shocks are orthogonal to each other, i.e.

$$\mathbf{G'G} = \mathbf{I} \tag{10}$$

whence $\mathbf{G'} = \mathbf{G}^{-1}$. Equation (10) gives $N(N-1)/2$ restrictions which are just sufficient to identify the elements of $\mathbf{G}$ and the diagonal elements of $\mathbf{\Omega}$. The solution involves rotation in N-dimensional space to identify the principal components $\mathbf{G}$ of $\mathbf{V}$, and the variances of these components, given by the diagonal elements of $\mathbf{\Omega}$.

Estimation procedure

1 Estimate a VAR using lags of up to three years for each commodity price.
2 Interpret the estimated coefficients using columns 1–8 of Table A.1.
3 Save the residuals from each commodity regression and run a principal components analysis on the residual covariance matrix.
4 Identify the components associated with the largest eigenvalues, and interpret them as common shocks affecting subsets of commodities.

A.2 A model of hierarchical markets with commodity-specific random shocks

Derivation of price reaction functions

Suppose now that prices are determined in perfectly competitive markets. It is assumed that

- Demand and supply for each commodity are linear in prices; they depend not only on the own-price of that commodity but the prices of all other commodities too;
- demand responds immediately to price: it is not driven by habit;
- shocks impinge on quantities rather than on prices;
- shocks are commodity-specific: shocks to different commodities are uncorrelated;
- in contrast to the previous model, prices are not administered but, like the previous model, no inventories are held;
- in contrast to the previous model, any producer can produce any commodity;
- no commodity is used to produce any of the other commodities;
- trading takes place once prices have matched demand to supply in every market; thus only equilibrium prices are observed.

Case 2.1: No production lag

Suppose to begin with that there are no production lags, so that supply in each market adjusts immediately to demand.

Demand and supply in market i at time t are respectively:

$$q_{dit} = \sum_m \alpha_{dim} z_{mt} + \sum_j \beta_{dij} p_{jt} + e_{dit} \quad (i, j = 1, \ldots, N;\ t = 1, \ldots, T) \tag{11.1}$$

$$q_{sit} = \sum_m \alpha_{sim} z_{mt} + \sum_j \beta_{sij} p_{jt} + e_{sit} \quad (i, j = 1, \ldots, N;\ t = 1, \ldots, T) \tag{11.2}$$

where α_{di}, α_{si} measure the responsiveness of demand and supply to exogenous non-price factors z_{mt}; β_{dij}, β_{sij} measure responsiveness to prices p_{jt}, and e_{dit}, e_{sit} are random demand and supply shocks respectively; they have zero mean and variance θ_i^2. It is assumed that $\beta_{dii} < 0$, $\beta_{sii} > 0$. Commodity j is a substitute in demand for commodity i if $\beta_{dij} > 0$, and is a substitute in supply if $\beta_{sij} < 0$.

On the left-hand side of equations (11) any monotone-increasing function of q, such as log q, can be substituted for q; this changes the interpretations of the parameters, but does not affect the behavioural implications of the model because quantity is not observed.

Equilibrium in market i implies that $q_{dit}=q_{sit}$, whence

$$\sum_j \beta_{ij} p_{jt} = \sum_m \alpha_{im} z_{mt} + e_{it} \quad (i, j = 1, \ldots, N) \tag{12}$$

where $\alpha_{im}=\alpha_{dim}-\alpha_{dim}$ measure the responsiveness of excess demand to exogenous non-price factors, $\beta_{ij}=\beta_{ijd}-\beta_{ijs}$ are the corresponding price responsivenesses, and $e_{it}=e_{idt}-e_{ist}$ are random shocks to excess demand. Note that β_{ii} is negative because excess demand is a decreasing function of own-price.

Dividing the ith equation through by $-\beta_{ii}$ and rearranging terms gives

$$p_{it} = \sum_m a_{im} z_{mt} + \sum_j b_{ij} p_{jt} + w_{it} \quad (i, j = 1, \ldots, N) \tag{13}$$

where $a_{im}=\alpha_{im}/(-\beta_{ii})$; $b_{ij}=0$ if $i=j$ and $\beta_{ij}/(-\beta_{ii})$ otherwise; and $w_{it}=e_{it}/(-\beta_{ii})$. The parameters a_{im} measure the responsiveness of the price to non-price factors, and b_{ij} the price responsiveness of one price to another; w_{it} are random commodity-specific price shocks with zero mean and variance $\sigma_j^2=\theta_i^2/\beta_{ii}^2$. The price shock w_{it} is the price analogue of the quantity shock e_{it}; normalisation by $-\beta_{ii}$ (a positive term) ensures that a unit increase in excess demand leads to a unit increase in equilibrium price. The parameters b_{ij} are positive (negative) whenever commodity j is a substitute (complement) for commodity i.

When shocks persist, w_{it}, satisfies the first-order equation

$$w_{it} = h_i w_{it-1} + u_{it} \tag{14}$$

In contrast to the administered pricing model, the persistence factor h_i may vary between markets. Applying the method of weighted differencing used before gives a first-order difference equation in prices which depends only on u_{it}:

$$p_{it} = \sum_j b_{ij} p_{jt} - h_i \sum_j b_{ij} p_{jt-1} + h_i p_{it-1} + \sum_m a_{im} z_{mt} - h_i \sum_m a_{im} z_{mt-1} + u_{it} \tag{15}$$

The coefficients on lagged prices are simply the negatives of the coefficients on the corresponding current prices, scaled down by the persistence factor, h_i. With substitution, therefore, coefficients on the current prices of other commodities are positive and coefficients on the lagged prices of other commodities are smaller but negative. The coefficient on the lagged own-price is positive, and equal to the persistence factor, whether there is substitution or not; the persistence model therefore predicts the opposite sign for this coefficient to that predicted by the cobweb model presented below.

A summary of the predictions derived from case 2.1 is presented in columns 1–8 of Table A.1.

Table A.1 Comparisons of the price-responsiveness coefficients in different models

	Model 1	*Model 2*	*Model 3*	*Model 4*	*Model 5*
Price formation	Admin	Admin	Admin	Admin	Admin
Adjustment	Complete $q_i = 1$	Partial $q_i < 1$	Complete $q_i = 1$	Partial $q_i < 1$	Complete $q_i = 1$
One-period production lag	NA	NA	NA	NA	NA
Forecast by stationarity or moving average	Stationary $k_i = 1$	Stationary $k_i = 1$	Moving average $k_i < 1$	Moving average $k_i < 1$	Stationary $k_i = 1$
Persistence	No $h = 0$	No $h = 0$	No $h = 0$	No $h = 0$	Yes $h > 0$
Equation	(3)	(3)	(3)	(3)	(5)
Own price					
One period lag	0	$1 - q_i$ *Positive*	0	$1 - q_i$ *Positive*	h *Positive*
Two period lag	0	0	0	0	0
Three period lag	0	0	0	0	0
Other price					
Current price	0	0	0	0	0
One period lag	b_{ij} *Positive*	$q_i b_{ij}$ *Positive*	$k_i b_{ij}$ *Positive*	$q_i k_i b_{ij}$ *Positive*	b_{ij} *Positive*
Two period lag	0	0	$(1 - k_i) b_{ij}$ *Positive*	$q_i (1 - k_i) b_{ij}$ *Positive*	$-h b_{ij}$ *Negative*
Three period lag	0	0	0	0	0

	Model 6	*Model 7*	*Model 8*	*Model 9*	*Model 10*
Price formation	Admin	Admin	Admin	Negotiated	Negotiated
Adjustment	Partial $q_i < 1$	Complete $q_i = 1$	Partial $q_i < 1$	Complete	Complete
One-period production lag	NA	NA	NA	No	No
Forecast by stationarity or moving average	Stationary $k_i = 1$	Moving average $k_i < 1$	Moving average $k_i < 1$	NA	NA
Persistence	Yes $h > 0$	Yes $h > 0$	Yes $h > 0$	No $h = 0$	Yes $h > 0$
Equation	(5)	(5)	(5)	(15)	(15)

Own price					
One period lag	$1-q_i+h$ *Positive*	h *Positive*	$1-q_i+h$ *Positive*	0	h *Positive*
Two period lag	$-h(1-q_i)$ *Negative*	0	$-h(1-q_i)$ *Negative*	0	0
Three period lag	0	0	0	0	0
Other price					
Current price	0	0	0	b_{ij} *Positive*	b_{ij} *Positive*
One period lag	$q_i b_{ij}$ *Positive*	$k_i b_{ij}$ *Positive*	$q_i k_i b_{ij}$ *Positive*	0	$-hb_{ij}$ *Negative*
Two period lag	$-q_i hb_{ij}$*Negaive*	$1-(1+h)k_i$	$q_i(1-(1+h)k_i)$	0	0
Three period lag	0	$h(1-k_i)$ *Positive*	$q_i h(1-k_i)$ *Positive*	0	0

	Model 11	*Model 12*	*Model 13*	*Model 14*
Price formation	Negotiated	Negotiated	Negotiated	Negotiated
Adjustment	Complete	Complete	Complete	Complete
One-period production lag	Yes	Yes	Yes	Yes
Forecast by stationarity or moving average	Stationary $k_i=1$	Moving average $k_i<1$	Stationary $k_i=1$	Moving average $k_i<1$
Persistence	No $h=0$	No $h=0$	Yes $h>0$	Yes $h>0$
Equation	(15)	(15)	(18)	(18)
Own price				
One period lag	$-b_i$ *Negative*	$-k_i b_i$ *Negative*	h_i-b_i	$h_i-k_i b_i$
Two period lag	0	$(k_i-1)b_i$ *Negative*	$h_i b_i$ *Positive*	$-(1-(1+h_i)k_i)b_i$
Three period lag	0	0	0	$h_i(1-k_i)b_i$ *Positive*
Other price				
Current price	b_{dij}	b_{dij}	b_{dij}	b_{dij}
One period lag	$-b_{sij}$ *Positive*	$-k_j b_{sij}$ *Positive*	$-(b_{sij}+h_i b_{dij})$	$-(k_j b_{sij}+h_i b_{dij})$
Two period lag	0	$(k_j-1)b_{sij}$ *Positive*	$-h_i b_{sij}$ *Positive*	$-(1-(1+h_i)k_j)b_{sij}$
Three period lag	0	0	0	$h_i(1-k_j)b_{sij}$ *Negative*

Note

The signs indicated in the table correspond to cases where $h>0$, $0<k_j<1$, $b_i>0$, $b_{dij}>0$, $b_{sij}<0$; these conditions imply, among other things, that products are substitutes in both demand and supply.

Case 2.2: One-period production lag with two-period weighted average expectations

Suppose now that there is a one-period production lag, e.g. crops need to be planted the season before, or additional workers hired before output can increase. Because production decisions are made before prices are known, price expectations must be formed on the basis of previous prices. It is assumed that all producers forecast prices in the same way, although they may use different weights when forecasting different commodities.

With a one-period production lag, supply in period t is governed by production decision made in period $t-1$, and these must be based on expectations of price, $p^e{}_{jt}$:

$$q_{sit} = \sum_m \alpha_{sim} z_{mt} + \sum_j \beta_{sij} p^e{}_{jt} + e_{sit} \quad (i, j = 1, \ldots, N) \tag{16.1}$$

The exogenous factors z_{mt} now reflect either factors operating in period t that cause actual supply to deviate from planned supply, or factors observed by producers in period $t-1$ that influence plans that are realised in period t. When producers forecast prices the most recent price information available relates to period $t-1$ (as in the administrative pricing model); it is assumed that, given the persistence of shocks, producers adopt a weighted average rule similar to (1.2):

$$p^e{}_{jt} = k_j p_{jt-1} + (1-k_j) p_{jt-2} \tag{16.2}$$

where k_j $(0 \leq k_j \leq 1)$ is the weight applied by all producers to the most recent recorded price in forecasting the price of the jth commodity.

The price reaction equation becomes

$$p_{it} = \sum_m a_{im} z_{mt} + \sum_j b_{dij} p_{jt} - k_i \sum_j k_j b_{sij} p_{jt-1} - k_i b_i p_{it-1} - \sum_j (1-k_j) b_{sij} p_{jt-2} - (1-k_i) b_i p_{it-2} + w_{it} \tag{17}$$

where $a_{im} = \alpha_{im} / (-\beta_{dii})$; $b_{dij} = 0$ if $i=j$ and $\beta_{dij} / (-\beta_{dii})$ otherwise; $b_{sij} = 0$ if $i=j$ and $\beta_{sij} / (-\beta_{dii})$ otherwise; $b_i = \beta_{sii} / (-\beta_{dii})$; and $w_{it} = e_{it} / (-\beta_{dii})$. Using 'weighted differencing' (as before) to remove the effects of persistence gives

$$p_{it} = \sum_j b_{dij} p_{jt} - \sum_j (k_j b_{sij} + h_i b_{dij}) p_{jt-1} + (h_i - k_i b_i) p_{it-1} - \sum_j (1-(1+h_i)k_j) b_{sij} p_{jt-2} - (1-(1+h_i)k_i) b_i p_{it-2} + \sum_j h_i (1-k_j) b_{sij} p_{jt-3} + h_i (1-k_i) b_i p_{it-3} + \sum_m a_{im} z_{mt} - h_i \sum_m a_{im} z_{mt-1} + u_{it} \tag{18}$$

Because of the price lag, equation (18) identifies substitution in demand separately from substitution in supply. If commodity j is a substitute in demand for commodity i then the slope parameter b_{dij} is positive, while if j is a substitute in supply the slope parameter b_{sij} is negative. Thus when commodities are substitutes in demand the coefficients on current prices are positive, and when commodities are substitutes in supply the coefficients on prices lagged three periods are negative. The equation also contains a lagged value of the own-price, which carries a coefficient b_i that depends on the ratio of the own-price responsiveness of supply to the own price responsiveness of demand. Since the assumptions imply that this coefficient is positive, equation (18) implies that the previous own-price, p_{it-1}, will have a negative impact on current price, $h_i/k_i < b_i$. In this case a 'cobweb' two-period price oscillation will occur, in which a high price stimulates supply and lowers the price the following period, while a low price discourages supply and raises the price the following period. For the cobweb process to be stable the roots of the auxiliary equation associated with the own-price components of equation (18) must lie outside the unit circle. In the absence of persistence the coefficient on the two-period lag has the same sign as the coefficient on the one-period lag, and the ratio of the two coefficients corresponds to the relative weights in the forecast. The interactions between the lags can lead to low-frequency sinusoidal oscillations in price.

General equilibrium in the system: derivation and estimation of reduced form equations

For each of the scenarios discussed above there is a general equilibrium in which the price equations for all markets are simultaneously satisfied. The sets of N price equations corresponding to the different cases described above are all special cases of the system:

$$\sum_j B_{ij} p_{jt} = \sum_w \sum_j c_{wij} p_{jt-w} + \sum_m (c_{4im} z_{mt} + c_{5im} z_{mt-1}) + u_{it}$$
$$(i, j = 1, \ldots, N;\ w = 1, 2, 3;\ t = 4, \ldots, T) \tag{19}$$

Note that equation (19) is valid even if some commodities have production lags and some do not.

Equation (19) can be rewritten in the standard form for a simultaneous equations system:

$$\mathbf{PB} = \mathbf{XC} + \mathbf{U} \tag{20}$$

where the data matrices $\mathbf{P}$, $\mathbf{X}$ are the same as before and $\mathbf{C}$ is a different matrix of coefficients with the same dimensions as before. $\mathbf{B}$ is an $N \times N$ matrix of full rank, with diagonal terms equal to unity and off-diagonal terms either $-b_{ij}$ or $-b_{dij}$ as appropriate. $\mathbf{U}$ is a matrix of random shocks with the same dimensions as before, and similar properties: $E(\mathbf{U}) = \mathbf{0}$, $E(\mathbf{U}'\mathbf{X}) = \mathbf{0}$. Shocks are specific rather than general, however, and so the covariance matrix is diagonal: $E(\mathbf{U}'\mathbf{U})/N = \Omega$, where Ω is an $N \times N$ diagonal matrix of variances σ_i^2.

Pre-multiplying by $\mathbf{X}'$, post-multiplying by the inverse matrix $\mathbf{B}^{-1}$ and taking expected values gives

$$\mathbf{CB}^{-1} = (\mathbf{X}'\mathbf{X})^{-1}\, \mathbf{X}' E(\mathbf{P}|\mathbf{X}) \tag{21}$$

Since the left-hand side consists of unknown coefficients and the right-hand side comprises observations, equation (21) can in principle be used to estimate the model by replacing $E(\mathbf{P}|\mathbf{X})$ with $\mathbf{P}$. But because the left-hand side is the product of two matrices, one of which is an inverse, individual parameters cannot normally be identified. If $\mathbf{B}$ were known, however, then $\mathbf{C}$ could be calculated. This can be achieved, in principle, by eliminating $\mathbf{C}$ – a process equivalent to regressing current price on all lagged prices by ordinary least squares (OLS) and calculating the residuals.

Post-multiplying (20) by $\mathbf{B}^{-1}$ and applying the estimator of $\mathbf{CB}^{-1}$ gives

$$\mathbf{UB}^{-1} = (\mathbf{I} - (\mathbf{X'X})^{-1}\mathbf{X'})\mathbf{P} = \mathbf{MP} \tag{22}$$

where $\mathbf{M} = (\mathbf{I} - (\mathbf{X'X})^{-1}\mathbf{X'})$ is a symmetric idempotent matrix, such that $\mathbf{M'M} = \mathbf{M}$. The covariance matrix of the commodity-specific price shocks, $E((\mathbf{B}^{-1})'\mathbf{U'UB}^{-1})/N$, can be obtained by transposing (22), multiplying, and taking expected values:

$$(\mathbf{B}^{-1})'\boldsymbol{\Omega}\,\mathbf{B}^{-1} = \mathbf{V} = E((\mathbf{P}'(\mathbf{I} - (\mathbf{X'X})^{-1}\mathbf{X'})'(\mathbf{I} - (\mathbf{X'X})^{-1}\mathbf{X'})\mathbf{P})|\mathbf{X}) = \mathbf{E}(\mathbf{P'MP}|\mathbf{X}) \tag{23}$$

The OLS estimator is obtained by replacing $E(\mathbf{P}|\mathbf{X})$ with the observed values $\mathbf{P}$. Equation (23) equates two symmetric $N \times N$ matrices: the left-hand side a matrix of unknown parameters in $\mathbf{B}$ and σ_i^2 $(i = 1, \ldots, N)$ alone, and the right-hand a matrix of sample residual covariances, $\mathbf{V}$. With $N(N+1)/2$ independent elements in the covariance matrix, $N(N+1)/2$ unknown parameters can in principle be determined, N of which are variances. The diagonal elements of $\mathbf{B}$ are normalised to unity by construction, and so there are only $N(N-1)$ elements of $\mathbf{B}$ to be determined. There are still too many, however.

There are several approaches to this problem (such as specifying the values of the variances), but only two are really plausible. The best known assumes that price formation is recursive. $\mathbf{B}$ is lower diagonal, so that $\mathbf{B}^{-1}$ is lower diagonal too. The second assumes that price reactions are symmetric, i.e. that $\mathbf{B}$ is symmetric.

- If $\mathbf{B}$ is lower diagonal then there is a strict hierarchy of markets. The price in market 1 is set independently of the current prices of all of other commodities; the price in market 2 is set with reference to market 2, and so on. In general price of each commodity is influenced only by higher ranked (i.e. lower numbered) commodities. Although lagged prices may influence the current price of any commodity, the assumption is still quite strict.
- Symmetry allows all prices to interact with each other, but no price can exert a greater influence on others than the other exert on it. The limitations of this assumption were discussed earlier.

When the two approaches are selectively combined, they can be very flexible. Approximately symmetric responses are analysed by assuming perfect symmetry and asymmetric responses are analysed by assuming recursion.

Symmetric interactions between a pair of commodities

Since the symmetry approach is not well known, it will be discussed first. Consider a two-commodity system. **B** is a 2×2 symmetric matrix with off-diagonal elements b; the diagonal elements are normalised to unity. The two commodities are substitutes if $b>0$ and complements if $b<0$. **P** is a $(T-3)\times2$ vector of current prices and **X** is a $(T-3)\times7$ matrix comprising a unit vector and the prices of both commodities lagged up to three years. The sample variances associated with commodities 1 and 2 are v_{11}, v_{22} and their covariance is v_{12}. Applying equation (23) gives three equations in the three unknowns b, $\sigma_1{}^2$, $\sigma_2{}^2$:

$$v_{11}=(\sigma_1{}^2+b^2\sigma_2{}^2)/(1-b^2)^2 \tag{24.1}$$

$$v_{22}=(b^2\sigma_1{}^2+\sigma_2{}^2)/(1-b^2)^2 \tag{24.2}$$

$$v_{12}=b(\sigma_1{}^2+\sigma_2{}^2)/(1-b^2)^2 \tag{24.3}$$

The solution is

$$b/(1+b^2)=v_{12}/(v_{11}+v_{22}) \tag{25.1}$$

$$\sigma_1{}^2=((1-b^2)/(1+b^2))(v_{11}-b^2v_{22}) \tag{25.2}$$

$$\sigma_2{}^2=((1-b^2)/(1+b^2))\ (v_{22}-b^2v_{11}) \tag{25.3}$$

Equations (25) show that the price response parameter for the pair of substitutes is a non-linear function of a correlation coefficient – not the Pearson coefficient, which divides the covariance by the geometric mean of the variances, but a coefficient that divides the covariance by the sum of the variances. This is equivalent to half the value of a coefficient that divides the covariance by the arithmetic mean of the variances – a property reflected in the fact that the coefficient varies between –0.5 and +0.5, rather than between –1.0 and +1.0 as usual. Solving equations (25) for b gives:

$$b=0 \text{ if } R=0 \tag{26}$$

$$=(1-(1-4R^2)^{1/2})/2R \text{ otherwise}$$

where $R=V_{12}/V_{11}+V_{22})$, $(-0.5<R<0.5)$. Substituting (26) into (25.2) and (25.3) determines the variances $\sigma_1{}^2$, $\sigma_2{}^2$.

Substituting the estimated values of b into **B** and post-multiplying equation (21) by **B** gives the OLS estimator

$$\mathbf{C}=(\mathbf{X'X})^{-1}\ \mathbf{X'PB} \tag{27}$$

Thus for each commodity the lag coefficients predicted in Table A.1 are estimated by a linear transformation, determined by b, of the slope coefficients from the OLS regression of the current price on lagged prices.

Recursive relation between a pair of commodities

The following well-known result for the two-commodity case is presented for purposes of comparison with equation (26) above. If the price of commodity 1 influences the price of commodity 2, but the price of commodity 2 does not affect the price of commodity 1, then **B** is a lower diagonal matrix in which the non-negative off-diagonal element is equal to b. Equations (23) imply that

$$b = v_{12}/v_{22} \tag{29.1}$$

$$\sigma_1{}^2 = v_{11} - (v_{12}{}^2/v_{22}) \tag{29.2}$$

$$\sigma_2{}^2 = v_{22} \tag{29.3}$$

Equations (29) show that b is the OLS slope estimator for a regression of the residuals from the price regression for commodity 1 on the residuals for the price regression for commodity 2. They imply that b can be estimated simply by regressing the current price of commodity 2 on lagged prices and the current price of commodity 1.

Unless like the symmetry case, this result generalises naturally to the N commodity case. When combining the two approaches, therefore, the symmetry case must be embedded within the recursive case, rather than the other way round.

Procedure for estimating the hierarchical model

The procedure described below assumes that commodities are paired. Unpaired commodities are analysed simply by OLS regression of their current price on current and previous values of other relevant variables.

1 Specify the ranking of commodity markets.
2 Regress the current prices of the two highest-ranked commodities separately on both sets of lagged prices (including a constant term), using OLS; record the slope coefficients and calculate the residuals.
3 Calculate the covariance matrix of the residuals and estimate the price interaction coefficient and the error variances using equations (25).
4 Substitute back the price interaction coefficient to estimate the impacts of lagged prices on the current price of each commodity, using equation (27); alternatively re-estimate the system using linear combinations of the relevant dependent variables, with weights determined by the relevant price interaction coefficients.
5 Derive predicted current prices for each commodity and save the residuals.
6 Repeat the procedure 3–5 for commodities 3 and 4, adding the current prices of commodities 1 and 2 to the set of regressors.
7 Move down the ranking of pairs one step at a time, repeating step 5 until all pairs have been analysed. The regressors for the final pair include all current prices other than the pair themselves.

8 Interpret the slope coefficients using column 9–14 of Table A.1. Analyse the residuals for each commodity price.

9 Test for normality, heteroskedasticity, serial correlation and cross-commodity correlation.

References

Allen, Robert C. (2001) The great divergence in European wages and prices from the Middle Ages to the First World War, *Explorations in Economic History*, 38, 411–447.

Allen, Robert C. (2009) *The British Industrial Revolution in Global Perspective*, Cambridge: Cambridge University Press.

Beveridge, William H. (1939) *Prices and Wages in England from the Twelfth to the Nineteenth Century*, London: Longmans and Green.

Clark, Gregory (2004) The price history of English agriculture, 1209–1914, *Research in Economic History*, 22, 41–123.

Clark, Gregory (2013) http://gpih.ucdavis.edu/Datafilelist.htm (accessed 1 March 2013).

Findlay, R. and K.H. O'Rourke (2003) Commodity market integration, 1500–2000, in M.D. Bordo, A.M. Taylor and J.G. Williamson (eds), *Globalization in Historical Perspective*, Chicago: University of Chicago Press, pp. 13–62.

Fischer, David H. (1996) *The Great Wave: Price Revolutions and the Rhythm of History*, Oxford: Oxford University Press.

Fisher, Franklin M. (1966) *The Identification Problem in Econometrics*, New York: McGraw-Hill.

Hoskins, H.G. (1964) Harvest fluctuations and English economic history, 1480–1619, *Agricultural History Review*, 12 (1), 28–46.

Hoskins, H.G. (1968) Harvest fluctuations and English economic history, 1620–1759, *Agricultural History Review*, 16 (1), 15–31.

Labys, Walter (2006) *Modelling and Forecasting Primary Commodity Prices*, Aldershot: Ashgate.

Lloyd, T.H. (1973) *The Movement of Wool Prices in Medieval England*, Cambridge: Cambridge University Press.

Marshall, Alfred (1890) *Principles of Economics*, London: Macmillan.

Mayhew, Nicholas (1995) Modelling medieval monetisation, in B.M.S. Campbell and R.H. Britnell (eds), *A Commercialising Economy: England 1086–1300*, Manchester: Manchester University Press, pp. 55–77.

Neal, Larry and Jeremy Atack (eds) (2008) *The Evolution of Financial Institutions and Markets from the Seventeenth to Twenty-First Century*, Cambridge: Cambridge University Press.

Nerlove, Marc (1958) *The Dynamics of Supply: Estimation of Farmers' Responses to Price*, Baltimore: MD: Johns Hopkins University Press.

Sargan, Denis (1988) *Lectures on Advanced Econometric Theory* (ed. M. Desai), Oxford: Blackwell.

Smith, Adam (1776) *An Inquiry in the Nature and Causes of the Wealth of Nations* (Glasgow edition), Oxford: Oxford University Press, 1976.

Stone, David (2005) *Decision-Making in Medieval Agriculture*, Oxford: Oxford University Press.

Threlfall-Holmes, Miranda (2005) *Monks and Markets: Durham Cathedral Priory 1460–1520*, Oxford: Oxford University Press.

3 The Quantity Theory of Money in historical perspective

Nick Mayhew

3.1 Introduction

This paper uses the Quantity Theory of Money (QT) to analyse price variations in England, 1270–1750. Prices play an important role in economic history, as noted in Chapter 2. The price level is a key influence on the cost of living. For a given level of money wages, the price level determines what baskets of ordinary products a wage-earner can afford to purchase. The price level is measured by taking a weighted average of commodity prices, using weights that correspond to the amount of each commodity contained in a representative basket of goods. Each price is measured in terms of money. A price index constructed in this way measures the rate of exchange between real goods on the one hand and a unit of currency on the other.

The money price level measures the scarcity of goods relative to the scarcity of money. It is therefore influenced by both the amount of goods that sellers bring to market, and the amount of money that buyers bring with them with which to purchase the goods. Prices can rise either because goods are in short supply or because money is in abundant supply. This is one of the basic insights of the QT.

The QT was developed in response to practical problems of stabilising prices through management of the currency. One of the pioneers of the QT was the seventeenth-century English philosopher, John Locke. The sixteenth century witnessed a major debasement of the English currency by Henry VIII, and the seventeenth century witnessed further volatility, due in part to the English Civil War. A systematic formulation of the QT was developed in the early twentieth century, when it was formulated independently by Irving Fisher in the US and by Cambridge economists (Alfred Marshall and his followers) in the UK. The theory was challenged by John Maynard Keynes in the 1930s because he believed that a very rigid application on QT principles contributed to the persistence of unemployment during the Great Depression. Keynes proposed a more flexible version of the theory, but in the 1970s the Chicago economist Milton Friedman was influential in asserting its validity in the long run. Recently, however, some of Friedman's claims have been challenged on the grounds that the relationship between money and prices is not so stable as he asserted.

3.2 Assumptions of the QT

The basic version of the QT focuses on the role of money as a medium of exchange. The role of money as a unit of account is also recognised by the QT, but its role as a store of value is not so important as in Keynesian theory.

The QT involves a sharp distinction between stocks and flows. Money is a stock that circulates in order to support a flow of transactions; indeed, considered as a medium of exchange, the entire rationale for money resides in its circulation. Sellers accept money is exchange for goods purely because they plan to buy other goods with the money they receive. However, they do not need to know, at the time they sell their goods, which particular goods they will buy; money therefore economises on information and forward planning, and avoids the difficulties posed by barter. The ratio between the value of the flow of transactions and the stock of money is the velocity of circulation.

Commodities are implicitly assumed by the QT to be perishable. Commodities are continuously produced and consumed. In contrast to money, which is a stock, commodities are a flow. This flow is generated by the utilisation of stocks of real assets such as labour, land and machinery. The sizes of these real stocks determine the maximum amount of output that the economy can sustain at full employment. The stock flow relationships between money and commodities generate the classic QT equation

$$MV = PY$$

where M is the stock of money, V is the velocity of circulation, P is the price level and Y is output (usually measured by gross domestic product (GDP)). The left-hand side of the equation measures the supply of money available for transactions in a given period, and the right-hand side the value of the transactions that money is required to support. Their equality represents a state of equilibrium in the economy.

Policy implications are derived from the QT by introducing additional assumptions.

- The money supply M is exogenous to the economy and is determined principally by the state;
- Real output Y is fixed in the short run by productive capacity, which reflects a legacy of previous investment decisions; output is equal to capacity because full employment is maintained through flexible labour markets; and
- The velocity of circulation V is fixed by the institutional factors, which change only in the very long run.

These assumptions imply that in the short run price P is determined by money supply M and output Y. Specifically, a one per cent increase in money supply, or a one per cent reduction in output will, other things being equal, induce a one per cent rise in prices. Thus in a fully employed economy price stability requires monetary stability; with a constant velocity of circulation, the rate of growth of the money supply must match the rate of growth of output.

These assumptions also imply that if the money supply is constant and output increases then prices will fall, as there will be a shortage of money, pushing up the price of money relative to goods and equivalently pushing down the price of goods in terms of money. To avoid the deflationary effect of output increases, QT policy is to increase the money supply in line with output in order to maintain price stability.

Modern research on the QT involves estimating the 'demand for money'. The basic idea is that there is a money market in which the money rate of interest on bonds adjusts to equilibrate the demand and supply of money. This reflects the modern view that the velocity of circulation depends on the rate of interest. With interest rate data it is possible to estimate a demand for money function in which the demand for money depends not only on output (as a proxy for the real value of transactions) and the interest rate, but also expected inflation, rates of return on other assets, and so on. This research has identified various adjustment lags which suggest that the money market may be characterised by short-run disequilibria. Interactions between lags complicate the analysis of monetary processes (Sriram, 1999).

Reliable annual interest rate data is not available for the period of this study. The model estimated in this chapter captures the role of interest rates indirectly through their link with other variables – in particular output. It addresses the problem of lags by distinguishing two variants of the QT – a strong QT, in which the money market is in continuous equilibrium, and a weak QT, in which the money market is out of equilibrium. Further details are given in the appendix.

When the QT is applied in a modern context, the measurement of the money stock is complicated by the existence of a sophisticated banking system. A distinction is often drawn between high-powered money such as holdings of bullion and low-powered money such as bank deposits. Under fractional reserve banking, notes and coins may be held as reserves rather than as a circulating medium; payments are made by cheques or credit cards instead. Before 1750, however, banking was under-developed compared to modern standards. A large amount of gold and silver was coined. The remainder was largely used as ornament, such as jewellery and plate. Compared to the twentieth century, therefore, there was a close connection between coinage and the money supply. It was still possible to transact without coin, however, as explained below.

3.3 Statistical sources

Recent work by the Broadberry and Campbell project has generated annual estimates of English population and GDP extending from the thirteenth century to the present (Broadberry *et al*., 2009, 2011). Price data for England has long been available, and although recent work by Robert Allen (2013) and by Gregory Clark (2009) have suggested modifications to the price index constructed by Phelps Brown and Hopkins (1962), all this work, especially as far as the middle ages is concerned, is founded on the original study by J.E. Thorold Rogers (1866–1902). The historian is thus equipped with annual estimates for *P* and *Y* over some seven centuries.

Estimates of money stock have been offered in the last twenty or thirty years, but these have only been for isolated dates in English history. Such estimates exist for various moments through the middle ages (Allen, 2012, p. 344) and through the early modern period (Mayhew, 2013, Table 3). However, the existence of an extended run of mint accounts from the 1220s to the present has made it possible to construct a model of English mint production on an annual basis, which can serve as the foundation of an annual series of estimates of the medieval and early modern English money stock (Challis, 1992, Appendix 1). This paper presents such a series of annual money estimates which, when set alongside the annual price index and annual estimates of GDP, makes it possible to measure the *M*, *P* and *Y* of the QT on a yearly basis.

It is important to recognise the approximate and provisional nature of all these estimates, even though the methods by which they have been calculated often generate apparently precise figures. And yet however approximate they undoubtedly are, there is a broad measure of agreement among those who have attempted such quantitative studies for certain fixed points in the story, and there are real limits to the degree of possible variation which could have occurred between these fixed points. Population and GDP can only plausibly grow (or contract) annually within certain limits. The documented output of the principal English mints provides a reliable guide to the likely size of the coinage, which in turn gives an indication of the probable size of the money stock. In short modelling the size of the economy and of the money stock allows us to set the resulting annual estimates alongside the well-established annual price index to provide a long-run data set of the principal macroeconomic factors of pre-industrial England.

This chapter takes the estimates of population and GDP generated by the Broadberry–Campbell team as a given. No doubt they will be subject to much debate, and revision – not least by the team itself – but for our current purposes, I have used the figures for population and real GDP per head generously supplied to me by Bas van Leeuwen in June 2011. However, since estimates of the coin and money stock are much less widely discussed, it is appropriate to examine them here much more closely.

The construction of the annual monetary estimates begins with the surviving documented mint output figures. This series of accounts is almost unbroken for London, is largely complete for Canterbury during the period of its most significant activity, and also provides some isolated information about occasional bursts of activity at other provincial mints. Since for long periods London was the only mint in operation, the documented mint output often represents 100 per cent of mint production, but that is not always the case. Particularly in the early fourteenth-century London and Canterbury together only produced about 75 per cent of actual output – Durham and Bury St Edmunds being busy at this time.[1] Consequently documented mint output accounts for only part – albeit the dominant part – of total coin production over the period as a whole. It would be possible to make estimates of the undocumented mint production, based on the representation of the respective mints in coin hoards, but for this chapter calculations have been based solely on documented output.[2]

From 1660 the mint accounts run from January to December, but it has been necessary to convert a mixture of accounting periods before this date to provide annual figures. This process necessarily involves assuming that output was spread equally over the whole account period, which was unlikely to have been the case, so creating twelve-month periods has introduced a small inaccuracy.

However, by far the most significant adjustment necessary to the mint account data arises from the fact that although they give excellent information on the quantities of coin struck, they tell us nothing about the amount of coin carried abroad in war, diplomatic payments or trade. Because England's monarchs managed to insist on the exclusive circulation of English coin within the kingdom to a remarkable degree, forcing foreign merchants to convert their own coin into sterling in order to conduct business in England, the mint accounts provide a very good guide to the positive side of the English balance of payments, but say nothing about coin carried out of the kingdom.[3] In addition to coin carried abroad, precious metal coins were subject to wear and clipping. This process meant that occasional recoinages were necessary to standardise the weight of the coinage once more.

This process of recoining the entire circulating medium provides a useful indication for historians of the size of the money stock at the time. Indeed such recoinages are an important element in attempts to make spot-estimates of the size of the currency at key dates – dates which are typically some fifty or so years apart. These exercises show clearly that the total quantity of recoined coin was significantly below the sum of the recorded mint output since the previous recoinage. Much coin which had been struck in the preceding years was no longer available for recoinage, because it had been carried abroad in war or trade, or been melted down for other purposes, or simply worn away in ordinary use.

For the purposes of this chapter all these forms of loss are termed wastage. In order to estimate the likely size of the currency annually it is necessary to add together the known output year by year, but also to deduct estimated wastage of all kinds. The annual estimates of the amount of coin in circulation in England are thus based on the sum of estimated annual output less an allowance for wastage, which is set at either 2 or 4 per cent per annum. The choice of wastage rate has been determined by an awareness of likely periods of heavy expenditure abroad, or of increasing use of bills of exchange – combined with an awareness of the 'spot-estimates' which have already been suggested for particular dates over our period by a range of historians. These 'spot-estimates' thus provide a sort of target at which the annual estimates should ideally arrive.

It will be clear that the annual figures suggested here for the silver and gold coinage from 1270 to 1750, though founded on good documentary evidence, are very much a constructed model which can do no more than indicate the approximate size of the coinage.

However, the size of the coinage itself remains at one remove from the size of the money stock, which needs to take account of various forms of paper, and of credit. There can be absolutely no doubt about the importance of credit in medieval and early modern England, as numerous recent studies testify (Kowaleski, 1995, ch. 5; Britnell, 1986, pp. 98–108, 206–208; Muldrew, 1998; Schofield and Mayhew, 2002; Gray, 2007; Briggs, 2009).[4] However, the supply of credit was itself limited by the supply of coin.[5] Moreover, different forms of credit need to be understood. While cash advances can be seen as contributions to velocity, when credit instruments begin to circulate in their own right they become additions to the money supply.[6] In the absence of reliable data on medieval and early modern interest rates,[7] there seems little prospect of quantifying the role of credit on anything like an annual basis which would permit it to be added to the series on population, GDP, prices and coin stock. For the moment we can only work with coin stock as a proxy for money stock. Thus in the early middle ages one needs to retain an awareness of how labour rents and payments in kind helped an extremely constrained coin supply to service the economy of the time, and how, from the later middle ages, that role of specialised means of payment was increasingly played by bills, bonds and banking. With full awareness of the limitations of the data, the annual estimates of the stock of silver and of gold coins in England derived by the method set out above, together with the annualised mint production figures, are presented in Table 3.1.

In Table 3.1 the first column from the left (1) gives the year. The next two columns (2 and 3) give the silver and gold annualised mint output respectively derived from surviving documents. The next three columns (4, 5, 6) provide the estimated silver coin stock for the years shown. From 1250 to 1278 (Column 6) assumes a wastage rate of 4 per cent per annum. The starting figure is based on an estimate of the size of the silver currency in 1250 directly after the recoinage begun in 1247.

In 1282, after the 1279–1281 recoinage, the silver coin stock was estimated at £603,255. This provides the basis for the new starting figure in 1283, which represents £603,255 plus the recorded mint output for that year, less 4 per cent estimated annual wastage. This series continues in Column 6 from 1283 to 1422.

It should be noted that the more or less comprehensive recoinages of 1247 and 1279 were replaced from the 1340s onwards by partial recoinages in which only old coin of good weight was likely to be drawn into the mints.

From 1423 to 1543 (Column 5) wastage of silver coin is set at 2 per cent per annum. This is an essentially arbitrary estimate of wastage, but the reduction in assumed wastage could reflect the increasing use of bills of exchange and of gold coin in international payments. No estimates have been attempted for the debasement period 1543–1560.The silver coin number for 1560 and 1561 (Column 4) is based on Challis' estimate, and includes much base currency from the debasement period still in circulation, together with good silver coin struck between 1552 and 1559. As a result of Elizabeth's recoinage, the silver stock was reduced in size, but improved in quality. Column 4 thereafter applies a

Table 3.1 New estimates of the money stock: England, 1220–1750

1 Year	*2 Silver: annualised mint output*	*3 Gold: annualised mint output*	*4 Silver from 1560*	*5 Silver stock from 1423 with 2 per cent waste*	*6 Silver stock from 1250 with 4 per cent waste*	*7 Gold stock from 1559 assuming 2 per cent waste*	*8 Gold stock from 1471 with 2 per cent waste*	*9 Gold stock from 1344 and 1527 with 4 per cent waste*
1220	7,749.643							
1221	18,599.14							
1222	17,049.21							
1223								
1224								
1225	10,940.5							
1226	13,246.38							
1227	8,816.25							
1228	8,816.25							
1229	2,204.063							
1230								
1231								
1232								
1233								
1234	15,997.92							
1235	42,091.25							
1236	50,227.25							
1237	46,361.15							
1238	29,836.6							
1239	33,448.33							
1240	48,156.33							
1241	31,244.33							
1242	38,096							
1243	67,580.5							
1244	52,609.33							

1245	46,094.93	
1246	43,579.85	
1247	48,335.83	
1248	114,020.3	
1249	118,865.1	
1250	71,879.56	560,000
1251	58,411.64	593,675.2
1252	65,098.02	635,026.2
1253	69,874	679,499.1
1254	69,956.02	722,275.2
1255	70,858.2	764,242.4
1256	69,197.13	802,869.8
1257	65,940.96	836,696
1258	54,150.86	857,379
1259	55,310.81	878,394.7
1260	43,785.75	887,044.6
1261	71,816.44	923,379.3
1262	49,172.17	935,616.3
1263	18,106.08	916,297.7
1264	14,176.5	893,822.3
1265	39,032.67	897,102.1
1266	36,535.18	897,753.2
1267	21,958.64	883,801.7
1268	21,958.64	870,408.3
1269	21,958.64	857,550.6
1270	21,958.64	845,207.2
1271	10,121.22	821,520.1
1272	9,917.199	798,576.5
1273	8,507.155	775,140.6
1274	12,998.93	757,133.9
1275	13,325.2	740,173.8

continued

1 Year	*2 Silver: annualised mint output*	*3 Gold: annualised mint output*	*4 Silver from 1560*	*5 Silver stock from 1423 with 2 per cent waste*	*6 Silver stock from 1250 with 4 per cent waste*	*7 Gold stock from 1559 assuming 2 per cent waste*	*8 Gold stock from 1471 with 2 per cent waste*	*9 Gold stock from 1344 and 1527 with 4 per cent waste*
1276	18,335.4				728,902.2			
1277	18,335.4				718,081.5			
1278	16,807.45				706,165.7			
1279	94,491.29							
1280	223,015							
1281	111,466.4							
1282	57,523.68							
1283	54,446.61				631,393.5			
1284	39,061.26				645,199.1			
1285	64,636.19				684,027.3			
1286	84,378.93				741,045.1			
1287	96,375.6				807,778.9			
1288	49,647.3				825,115.1			
1289	16,056.6				808,167.1			
1290	10,763.02				786,603.4			
1291	3,523.667				758,662.9			
1292	4,673.25				732,989.7			
1293	4,108.25				707,778.3			
1294	6,279.875				685,747.1			
1295	5,709.5				664,026.7			
1296	5,978.125				643,443.7			
1297	5,557.615				623,263.6			
1298	4,011.93				602,345			
1299	54,311.2				632,562.4			
1300	146,359.3				753,619.2			
1301	52,670.25				776,144.6			

1302	12,334.75	757,433.6
1303	15,416	742,552.3
1304	51,366.75	764,216.9
1305	104,875.3	838,523.5
1306	109,610.3	914,592.8
1307	138,298.5	1,016,308
1308	124,141	1,099,796
1309	119,291.5	1,175,096
1310	38,631	1,166,723
1311	8,657.75	1,128,712
1312	17,983.75	1,101,547
1313	27,638.25	1,085,124
1314	58,682.5	1,100,401
1315	2,5814	1,082,199
1316	8,570	1,047,481
1317	26,077.25	1,031,659
1318	33,202.25	1,023,595
1319	26,404.25	1,009,055
1320	22,403.25	991,096.5
1321	12,674.75	964,127.4
1322	4,387.5	929,949.8
1323	2,034.5	894,786.3
1324	1,347	860,341.8
1325	123.25	826,051.4
1326	166.25	793,175.6
1327	217.25	761,665.8
1328	333	731,532.2
1329	781.5	703,052.4
1330	373.5	675,303.8
1331	869.3333	649,161
1332	478	623,672.6

continued

1 Year	*2 Silver: annualised mint output*	*3 Gold: annualised mint output*	*4 Silver from 1560*	*5 Silver stock from 1423 with 2 per cent waste*	*6 Silver stock from 1250 with 4 per cent waste*	*7 Gold stock from 1559 assuming 2 per cent waste*	*8 Gold stock from 1471 with 2 per cent waste*	*9 Gold stock from 1344 and 1527 with 4 per cent waste*
1333	594.75				599,320.4			
1334	467				575,814.6			
1335	1,352.5				554,134.5			
1336	2,797				534,766.1			
1337	1,363.5				514,739			
1338	1,587.5				495,736.9			
1339	1,892				477,799.4			
1340	1,735.5				460,423			
1341	2,189.5				444,195.5			
1342	7,712				434,139.7			
1343	14,571.75				431,345.9			
1344	43,784.75	41,819.5			457,876.8			41,819.5
1345	21,946	9,558.75			461,507.7			49,323.12
1346	7,124.5	15,392			450,171.9			62,126.52
1347	5,922.75	38,269.5			438,087.8			96,380.17
1348	7,772.25	35,026.5			428,336.5			126,150.4
1349	6,546.45	27,660.25			417,749.5			147,658.2
1350	9,389.467	19,607.67			410,429			160,575.3
1351	36,565.33	107,498.3			429,114.6			257,350.7
1352	85,184.75	61,720			493,727.3			306,307.8
1353	120,770.3	51,802.33			589,917.8			343,785.8
1354	56,682	142,781.5			620,735.8			467,104.6
1355	4,0713	66,477.5			634,990.8			512,238.8
1356	31,350.8	9,714.4			639,688			501,075.1
1357	18,195.95	103,851.4			631,568.6			580,729.3
1358	11,842	108,622.3			617,674.1			661,777.5

1359	9,353	89,218.25	601,946	720,956
1360	6,112.25	99,149	583,736	787,300.8
1361	8,056	188,687	568,120.3	936,948.2
1362	11,410	107,792.3	556,349.1	1,002,951
1363	3,908	46,663.75	537,846.8	1,007,630
1364	5,134	61,538.75	521,261.6	1,026,402
1365	1,479	47,509.25	501,830.9	1,030,955
1366	274.25	104,166.1	482,021	1,089,716
1367	1,097	79,689.5	463,793.3	1,122,629
1368	1,257.75	78,122.5	446,449	1,152,722
1369	1,740	73,421.5	430,261.4	1,177,098
1370	1,505.25	62,793.38	414,496	1,190,295
1371	679.125	36,495.38	398,568.1	1,177,719
1372	313.5	53,254.5	382,926.4	1,181,735
1373	351.625	45,096.88	367,946.9	1,177,758
1374	1,391.5	18,852	354,564.8	1,148,746
1375	3,459.028	12,201.11	343,702.9	1,114,509
1376	1,332.111	8,196.444	331,233.6	1,077,797
1377	1,332.111	8,196.444	319,263.1	1,042,554
1378	1,332.111	8,196.444	307,771.4	1,008,720
1379	1,332.111	8,196.444	296,739.4	976,240.2
1380	1,332.111	8,196.444	286,148.6	945,059.2
1381	1,332.111	8,196.444	275,981.5	915,125.4
1382	1,332.111	8,196.444	266,221.1	886,389
1383	1,332.111	8,196.444	256,851	858,802
1384	1,275.053	11,466.73	247,801.1	835,458
1385	1,103.878	21,277.59	238,948.7	822,466.1
1386	1,103.878	21,277.59	230,450.5	809,994
1387	1,103.878	21,277.59	222,292.2	798,020.7
1388	1,103.878	21,277.59	214,460.2	786,526.4
1389	1,103.878	21,277.59	206,941.6	775,491.8

continued

1 Year	*2 Silver: annualised mint output*	*3 Gold: annualised mint output*	*4 Silver from 1560*	*5 Silver stock from 1423 with 2 per cent waste*	*6 Silver stock from 1250 with 4 per cent waste*	*7 Gold stock from 1559 assuming 2 per cent waste*	*8 Gold stock from 1471 with 2 per cent waste*	*9 Gold stock from 1344 and 1527 with 4 per cent waste*
1390	1,103.878	21,277.59			199,723.6			764,898.6
1391	1,103.878	21,277.59			192,794.4			754,729.2
1392	1,029.241	21,916.49			186,070.7			745,579.8
1393	208.2353	28,944.35			178,827.8			743,543.2
1394	208.2353	28,944.35			171,874.6			741,588.1
1395	350.739	26,081.51			165,336.3			736,962.8
1396	778.25	17,493			159,470			724,277.6
1397	778.25	17,493			153,838.3			712,099.7
1398	778.25	17,493			148,431.9			700,409
1399	658.535	15,501.29			143,126.8			687,273.9
1400	299.3898	9,526.169			137,689.1			668,928.1
1401	299.3898	9,526.169			132,469			651,316.1
1402	299.3898	9,526.169			127,457.6			634,408.6
1403	299.3898	9,526.169			122,646.7			618,177.3
1404	337.2924	8,323.627			118,064.7			601,440.9
1405	363.5	4,890.25			113,691			582,077.9
1406	95.75	4,807			109,235.3			563,409.5
1407	62	2,785.75			104,925.4			543,547.5
1408	6	1,632			100,734.2			523,372.3
1409					96,704.81			502,437.4
1410					92,836.62			482,339.9
1411	242.6667	12,489.25			89,356.11			475,036
1412	3,124.667	148,950.6			88,781.55			599,027.1
1413	5,852.313	123,713.8			90,848.51			693,831.2
1414	7,017.25	78,377			93,951.13			741,319.9
1415	7,017.25	78,377			96,929.64			786,909

1416	7,017.25	78,377		99,789.02	830,674.6
1417	5,789.271	66,742.58		101,355.2	861,520.5
1418	2,105.333	31,839.33		99,322.07	857,625.4
1419	1,596.208	30,909.83		96,881.55	852,993.8
1420	1,824.333	39,336.33		94,757.64	856, 637
1421	2,217.333	65,649.33		93,095.98	885,394.8
1422	7,050.233	146,023.1		96,140.36	990,161.3
1423	8,661.2	172,814.4	102,705.5		1,116,457
1424	13,642.65	136,772	114,021.2		1,203,099
1425	28,587	28,644.78	139,756.1		1,182,474
1426	28,587	28,644.78	164,976.2		1,162,674
1427	28,587	28,644.78	189,691.9		1,143,667
1428	28,587	28,644.78	213,913.4		1,125,419
1429	28,587	28,644.78	237,650.3		1,107,901
1430	28,587	28,644.78	260,912.6		1,091,084
1431	28,587	28,644.78	283,709.6		1,074,940
1432	28,587	28,644.78	306,050.7		1,059,441
1433	21,654	24,129.08	321,150.6		1,040,227
1434	903.25	9,603.167	315,612.8		1,007,837
1435	1,804.4	6,704.133	311,068.8		973,959.7
1436	2,125.6	6,471.95	306,930.5		941,214.4
1437	1,169	5,560.25	301,937.5		908,903.7
1438	2,948.6	5,757.8	298,788.4		878,075
1439	5,146.4	7,275.2	297,856.1		849,936.2
1440	3,534	6,317.25	295,362.3		822,003.3
1441	1,397.034	5,284.233	290,824.2		794,196.1
1442	328.5517	4,767.724	285,329.7		767,005.2
1443	304.9138	4,591.293	279,921.9		740,732.7
1444	253.25	3,721.5	274,571.6		714,676
1445	296.25	2,653	269,370.5		688,635.8
1446	3,854.5	4,676.5	267,760.5		665,579.8

continued

1 Year	*2 Silver: annualised mint output*	*3 Gold: annualised mint output*	*4 Silver from 1560*	*5 Silver stock from 1423 with 2 per cent waste*	*6 Silver stock from 1250 with 4 per cent waste*	*7 Gold stock from 1559 assuming 2 per cent waste*	*8 Gold stock from 1471 with 2 per cent waste*	*9 Gold stock from 1344 and 1527 with 4 per cent waste*
1447	300.2778	1,491.167		262,699.6				640,388.2
1448	467.5556	1,517.333		257,903.8				616,229.3
1449	2,088.917	2,625.75		254,792.9				594,100.8
1450	7,912.083	5,616.25		257,450.8				575,728.4
1451	10,789.33	4,612		262,875.4				557,126.8
1452	7,298.583	4,428.25		264,770.5				539,092.8
1453	5,589.75	2,639.75		264,953				520,063.3
1454	4,656.58	1,554.783		264,217.4				500,753.3
1455	4,280.87	1,300.174		263,128.3				481,971.4
1456	6,066.717	1,393.543		263,811.1				464,030.3
1457	6,369.25	1,422		264,776.8				446,834.2
1458	5,282	1,141.5		264,657.6				430,056.7
1459	6,132.25	714.75		265,374				413,540.6
1460	7,923	1,415.25		267,831.1				398,357.6
1461				262,474.5				382,423.3
1462	2,655.234	728.4468		259,827.1				367,825.7
1463	9,103.66	2,497.532		263,552.2				355,510.3
1464	22,669.59	46,268.86		280,497.3				385,708
1465	49,801.44	133,811.5		323,692.8				498,738.7
1466	37,351.08	100,358.6		353,823				575,133.5
1467				346,746.5				552,128.1
1468	2,777.636	8,363.818		342,533.7				538,072.3
1469	17,383.36	51,810.43		352,718.7				566,287.4
1470	14,652	42,519.75		360,023.3				584,454.9
1471	16,920.5	35,785		369,404.9			618,591	
1472	18,440.25	45,777.5		380,088.3			651,081.1	

1473	12,902.5	31,144.75	385,130.9	668,581.3
1474	11,802.25	33,353	388,994.5	687,895.6
1475	11,650.93	26,550.11	392,632.5	700,156.8
1476	5,776.807	24,891.97	390,441.2	710,547.8
1477	5,161.765	27,774.18	387,690.9	723,555.6
1478	4,216	24,809.75	384,068.7	733,398
1479	4,733.25	25,480.75	381,025.9	743,701.2
1480	3,200.25	27,882.75	376,541.7	756,152.2
1481	2,174.5	17,718.75	371,141.9	758,393.6
1482	3,907.75	15,584.5	367,548.6	758,498.5
1483	7,837	12,008.25	367,877.9	755,096.6
1484	11,155.75	14,589.5	371,453	754,292.4
1485	5,404.75	9,211.75	369,320.6	748,234.1
1486	6,660.5	9,908.75	368,461.4	742,980
1487	3,915.25	8,092.75	364,929.2	736,051.3
1488	5,006	8,151.75	362,536.5	729,318.9
1489	3,831	4,020.75	359,040.1	718,672.9
1490			351,859.3	704,299.4
1491			344,822.1	690,231.5
1492			337,925.7	676,409.2
1493			331,167.2	662,881
1494	2,306.75	5,941	326,804.4	655,455.6
1495	8,002.75	21,121	328,111	663,035.2
1496	5,362	14,196.75	326,803.6	663,687.3
1497	9,955.75	17,791.75	330,024.1	667,489.5
1498	17,101	20,174.5	340,182.6	674,263.5
1499	23,756.75	21,100.25	356,660.6	681,456.5
1500	20,058	19,825.75	369,184.2	687,256.6
1501	20,162.5	26,115.5	381,559.8	699,104.7
1502	17,637.25	29,333	391,213.1	713,868.9
1503	17,405.5	30,537	400,446.2	729,517.8

continued

1 Year	*2 Silver: annualised mint output*	*3 Gold: annualised mint output*	*4 Silver from 1560*	*5 Silver stock from 1423 with 2 per cent waste*	*6 Silver stock from 1250 with 4 per cent waste*	*7 Gold stock from 1559 assuming 2 per cent waste*	*8 Gold stock from 1471 with 2 per cent waste*	*9 Gold stock from 1344 and 1527 with 4 per cent waste*
1504	30,709.75	39,092.25		422,532.9			753,237.8	
1505	43,489.5	59,530.5		456,701.9			796,513	
1506	36,461	93,066.75		483,299.7			871,788.1	
1507	28,723.25	94,879		501,782.5			947,333.8	
1508	20,300.75	121,831.8		511,641.5			1,047,782	
1509	7,729	106,726.5		508,983.1			1,131,419	
1510	2,679.75	64,486.75		501,429.6			1,171,987	
1511	3,457.5	44,597.5		494,789.4			1,192,253	
1512	11,178	38,481.25		495,848			1,206,120	
1513	11,527.75	62,845.25		497,228.3			1,243,586	
1514	4,322.5	34,452		491,519.7			1,252,477	
1515	813.25	44,873.25		482,486.3			1,271,403	
1516	135.75	40,146.75		472,969.6			1,285,319	
1517	250.5	11,473.5		463,755.7			1,270,356	
1518	865	48,153		455,328.3			1,292,629	
1519	356.75	50,267.25		446,571.4			1,316,039	
1520	420.25	33,978.75		438,051.8			1,323,017	
1521	4,674	23,980.25		433,871.3			1,320,057	
1522	15,068.75	13,300.5		439,961.2			1,306,691	
1523	13,170.75	6,855.75		444,069.3			1,227,275	
1524				435,188			1,261,530	
1525				426,484.2			1,236,299	
1526	5,138.667	30,509		422,990.4			1,241,472	
1527	51,964.57	99,259.5		465,455.9				826,210.5
1528	73,097.14	26,546.5		527,782				818,646.7
1529	73,097.14	12,524.75		588,861.5				797,924.6

1530	49,741.29	9,444.75	625,830.8		775,074.6
1531	33,058.55	6,034.5	645,711.5		749,864.7
1532	33,058.55		665,194.7		719,870.2
1533	28,536.26	13,655.5	679,856.3		704,184.6
1534	23,026.5	13,655.5	688,825.1		689,126.5
1535			675,048.6		661,561.5
1536	10,629.75	6,563	671,964.8		641,399.5
1537	37,465.82	27,258.48	695,242		641,911.7
1538	33,374.6	30,277.94	714,044.3		645,302
1539	38,908.75	30,277.94	737,894		648,556.7
1540	25,939.17	15,527.81	748,556.5		637,521.2
1541	29,181.56	8,736	762,183.3		620,406.9
1542	22,133.3	3,990.571	768,630.3		599,421.6
1543	36,654.86	15,962.29	789,179.4		590,768.5
1544	98,735.91	103,549.2			
1545	374,945.9	328,913.6			
1546	450,265.3	290,418.5			
1547	297,494.5	325,110			
1548	340,624.3	148,200.8			
1549	531,668.8	60,949			
1550	370,243.2	5,080.65			
1551	222,629.3	3,743.1			
1552	20,820	1,798.5			
1553	0	0			
1554	79,199	22,068			
1555	76,349	36,900			
1556	0	0			
1557	0	0			
1558	0	0			
1559	19,776	16,191.16		113,867.3	

continued

1 Year	*2 Silver: annualised mint output*	*3 Gold: annualised mint output*	*4 Silver from 1560*	*5 Silver stock from 1423 with 2 per cent waste*	*6 Silver stock from 1250 with 4 per cent waste*	*7 Gold stock from 1559 assuming 2 per cent waste*	*8 Gold stock from 1471 with 2 per cent waste*	*9 Gold stock from 1344 and 1527 with 4 per cent waste*
1560	101,196.7	10,378.57	1,580,000			121,761		
1561	707,163.9	21,973.77	1,350,000			140,860.1		
1562	178,236.7	63,182.5	1,497,672			199,961.7		
1563	35,815.88	13,999.88	1,502,818			209,682.4		
1564	35,815.88	13,999.88	1,507,861			219,208.6		
1565	40,713.25	25,650.25	1,517,603			239,961.7		
1566	85,486.67	29,981.89	1,571,028			264,544.7		
1567	103,717.7	32,663.11	1,641,251			291,263.6		
1568	101,672.7	48,210	1,708,065			332,684.2		
1569	116,394.9	25,118.8	1,787,971			350,646.9		
1570	74,499.64	13,410.3	1,825,221			356,776.1		
1571	64,648.36	13,377.71	1,852,072			362,750.7		
1572	74,234.22	9,343.111	1,887,780			364,651.9		
1573	110,888.8	12,097.56	1,958,695			369,214.5		
1574	31,445.14	7,891.429	1,950,338			369,563.8		
1575	53,725.96	3,557.143	1,963,982			365,658.5		
1576	53,725.96	3,557.143	1,977,354			361,831.4		
1577	53,725.96	3,557.143	1,990,459			358,080.7		
1578	57,472.09	5,114.15	2,006,972			355,931		
1579	104,527.8	12,156.6	2,069,270			360,725.8		
1580	96,419.46	17,413.62	2,122,375			370,576.7		
1581	7,564.316	21,168.63	2,087,341			383,910.4		
1582	63,226.4	1,272	2,107,556			377,478.7		
1583	188,589.1	45,613.73	2,250,222			414,630.6		
1584	207,470.5	37,930.42	2,408,539			443,509.8		
1585	168,806.8	33,821.17	2,525,799			467,784.4		

1586	69,608.25	21,536.83	2,543,499	479,534.8
1587	46,096	32,395.17	2,537,803	501,691.3
1588	35,381.5	23,545.67	2,521,721	514,732.3
1589	40,225.25	17,430.33	2,510,707	521,519.3
1590	60,659.42	12,994	2,519,939	523,823.1
1591	98,176	26,835.67	2,565,753	539,645.6
1592	180,693.9	15,894.83	2,691,518	544,429.6
1593	187,900.4	14,783	2,821,830	548,028.3
1594	187,900.4	14,783	2,949,536	551,555.1
1595	187,900.4	14,783	3,074,687	555,011.4
1596	187,900.4	14,783	3,197,336	558,398.5
1597	35,425.02	8,586.448	3,168,106	555,645.2
1598	21,563.63	8,023.125	3,125,876	552,395
1599	20,710.76	11,604.57	3,083,655	552,719.6
1600	18,152.18	22,348.91	3,039,771	563,567.1
1601	60,103.52	18,794.11	3,037,877	570,714
1602	118,835.4	13,817.4	3,093,578	572,840.8
1603	220,900.2	15,875.46	3,248,189	576,941.9
1604	318,698.8	57,023.02	3,495,550	621,285.6
1605	259,199.7	193,375.1	3,679,655	798,367.5
1606	249,049.6	163,163.5	3,850,130	942,300.3
1607	177,073.9	144,416.1	3,946,660	1,064,982
1608	94,110.25	131,125.3	3,959,955	1,172,185
1609	35,661.5	90,329.25	3,915,704	1,237,264
1610	21,372.75	48,345.5	3, 858,335	1,259,897
1611	20,308.75	64,797.25	3,801,071	1,298,201
1612	9,843.5	245,064.3	3,734,696	1,512,400
1613	13,123.5	320,344.8	3,672,863	1,796,090
1614	5,190	208,265.8	3,604,492	1,964,268
1615	9,691.75	231,762.3	3,541,900	2,152,110
1616	4,179.25	260,825.5	3,475,158	2,364,677

continued

1 Year	*2 Silver: annualised mint output*	*3 Gold: annualised mint output*	*4 Silver from 1560*	*5 Silver stock from 1423 with 2 per cent waste*	*6 Silver stock from 1250 with 4 per cent waste*	*7 Gold stock from 1559 assuming 2 per cent waste*	*8 Gold stock from 1471 with 2 per cent waste*	*9 Gold stock from 1344 and 1527 with 4 per cent waste*
1617	528.25	181,091	3,406,173			2,494,852		
1618	406.5	150,370.5	3,338,448			2,592,318		
1619	117.5	139,840	3,271,794			2,677,515		
1620	266.4	156,756.3	3,206,619			2,777,586		
1621	16,336.31	142,433.6	3,158,496			2,861,619		
1622	24,801.09	137,429.1	3,119,631			2,939,067		
1623	49,061.45	250,369.3	3,105,319			3,125,648		
1624	95,371.88	786,136.5	3,136,677			3,833,549		
1625	80,117.5	686,145.5	3,152,459			4,429,300		
1626	57,347	492,842.3	3,145,609			4,823,700		
1627	20,531.75	290,384.5	3,102,818			5,011,803		
1628	6,106.5	162,531.5	3,046,746			5,070,847		
1629	3,931	183,216.3	2,989,664			5,148,982		
1630	4,695	302,497.3	2,934,472			5,342,450		
1631	51,921.25	193,738.8	2,926,665			5,425,465		
1632	132,384	115,442.3	2,997,868			5,430,089		
1633	180,764.5	91,160	3,115,060			5,410,824		
1634	252,229.5	86,816.75	3,299,944			5,387,688		
1635	166,890.5	94,198.25	3,397,497			5,372,249		
1636	334,133	97,305	3,656,998			5,360,162		
1637	494,245.5	79,918	4,068,218			5,331,279		
1638	409,009.3	72,894	4,387,683			5,296,089		
1639	405,739.8	47,551.47	4,697,554			5,236,768		
1640	477,520.5	55,482.25	5,071,573			5,186,405		
1641	863,460.9	63,444.5	5,816,334			5,144,853		
1642	976,030.1	58,935.25	6,656,517			5,099,712		
1643	872,475.2	54,126	7,378,412			5,050762		

1644	872,475.2	54,126	8,085,869	5,002,790
1645	891,339.2	43,146	8,797,664	4,945,017
1646	788,525.3	42,053.25	9,394,466	4,887329
1647	222,833.2	31,006.99	9,424,953	4,819,969
1648	46,740.96	26,836.32	9,282,260	4,749,869
1649	23,727.42	17,849.96	9,119,868	4,672,365
1650	12,220.65	13,356.77	8,949,447	4,592,007
1651	12,220.65	13,356.77	8,782,434	4,513,257
1652	12,220.65	13,356.77	8,618,762	4,436,081
1653	175,194.9	18,660.2	8,618,077	4,365,647
1654	83,732.25	5,777.75	8,527,774	4,283,996
1655	83,732.25	5,777.75	8,439,276	4,203,978
1656	83,732.25	5,777.75	8,352,548	4,125,561
1657	77,344.97	5,332.18	8,261,295	4,048,275
1658	7,084.909	430.9091	8,103,012	3,967,732
1659	13,090.81	710.18	7,953,781	3,889,073
1660	15,318.08	6,975.8	7,809,717	3,818,128
1661	23,201	4,138	7,676,260	3,745,821
1662	496,678	31,186	8,009,479	3,701,467
1663	305,078	1,231	8,148,266	3,628,644
1664	44,333	53,011	8,028,747	3,608,022
1665	61,722	87,452	7,928,659	3,621,564
1666	37,144	65,218	7,806,487	3,613,046
1667	53,107	125,685	7,702,402	3,663,957
1668	124,940	198,022	7,670,796	3,784,739
1669	44,305	116,588	7,560,799	3,823,301
1670	143,043	121,833	7,549,765	3,866,231
1671	119,800	171,502	7,516,173	3,956,978
1672	268,689	83,661	7,629,165	3,959,827
1673	313,300	119,852	7,783,616	3,998,085
1674	31,888	82,101	7,659,194	3,998,582

continued

1 Year	*2 Silver: annualised mint output*	*3 Gold: annualised mint output*	*4 Silver from 1560*	*5 Silver stock from 1423 with 2 per cent waste*	*6 Silver stock from 1250 with 4 per cent waste*	*7 Gold stock from 1559 assuming 2 per cent waste*	*8 Gold stock from 1471 with 2 per cent waste*	*9 Gold stock from 1344 and 1527 with 4 per cent waste*
1675	5,754	51,741	7,511,649			3,969,317		
1676	315,695	233,612	7,670,797			4,118,870		
1677	452,706	228,551	7,961,033			4,260,473		
1678	22,835	136,856	7,824,191			4,309,382		
1679	270,671	534,409	7,932,964			4,746,915		
1680	180,433	572,046	7,951,129			5,212,582		
1681	95,427	299,119	7,885,625			5,401,467		
1682	39,553	183,535	7,766,675			5,473,302		
1683	216,538	369,317	7,823,548			5,725,767		
1684	53,660	284,488	7,719,664			5,890,050		
1685	94,773	537,338	7,658,149			6,298,840		
1686	59,814	617,411	7,563,603			6,777,926		
1687	250,630	401,301	7,657,949			7,035,642		
1688	76,231	561,309	7,579,496			7,445,012		
1689	96,573	128,442	7,522,548			7,421,985		
1690	1,995	48,722	7,374,052			7,321,293		
1691	3,731	54,497	7,230,227			7,228,274		
1692	4,160	114,499	7,089,699			7,195,918		
1693	9,277	55,968	6,956,997			7,106,848		
1694	160	57,246	6,818,014			7,020,812		
1695	62	717,218	6,681,714			7,583,270		
1696	2,717,086	138,618	2,717,086			7,567,450		
1697	3,660,408	120,447	6,249,944			7,534,139		
1698	463,226	471,567	6,578,907			7,845,592		
1699	60,444	141,377	6,506,564			7,827,229		
1700	14,898	120,212	6,391,032			7,788,493		
1701	116,179	1,190,019	6,377,067			8,798,941		

1702	355	162,069	6,249,874	8,781,790
1703	2,226	1,520	6,127,058	8,607,644
1704	12,422		6,016,690	8,435,491
1705	1,332	4,628	5,897,662	8,271,317
1706	2,889	23,897	5,782,540	8129,309
1707	3,639	27,012	5,670,455	7,993,195
1708	11,628	44,945	5,568,442	7,877,377
1709	78,811	109,826	5,534,307	7,827,459
1710	2,533	165,362	5,426,104	7,832,965
1711	76,781	414,918	5,392,827	8,082,925
1712	5,532	127,048	5,290,392	8,045,774
1713	7,232	584,597	5,191,671	8,457,763
1714	4,855	1,313,907	5,092,596	9,576,237
1715	5,093	1,739,505	4,995,735	11,089,427
1716	5,115	1,057,543	4,900,833	11,904,031
1717	2,939	675,777	4,805,697	12,328,211
1718	7,115	140,642	4,716,555	12,219,476
1719	5,444	688,960	4,627,559	12,650,268
1720	24,279	885,859	4,558,802	13,265,404
1721	7,170	272,500	4,474,652	13,267,146
1722	6,147	594,716	4,391,183	13,584,625
1723	149,107	388,098	4,449,484	13,693,268
1724	5,121	273,809	4,365,513	13,687,736
1725	7,735	58,360	4,285,783	13,471,174
1726	2,592	872,963	4,202,608	14,057,254
1727	2,049	292,779	4,120,564	14,063,032
1728	2,644	53,874	4,040,743	13,834,568
1729	6,371		3,966,172	13,557,877
1730	3,478	91,628	3,890,257	13,376,515
1731	2,182	305,768	3,814,590	13,408,637
1732	2,620	373,473	3,740,866	13,506,468

continued

1 Year	*2 Silver: annualised mint output*	*3 Gold: annualised mint output*	*4 Silver from 1560*	*5 Silver stock from 1423 with 2 per cent waste*	*6 Silver stock from 1250 with 4 per cent waste*	*7 Gold stock from 1559 assuming 2 per cent waste*	*8 Gold stock from 1471 with 2 per cent waste*	*9 Gold stock from 1344 and 1527 with 4 per cent waste*
1733	3,581	833,948	3,669,558			14,053,608		
1734	4,929	487,108	3,600,998			14,249,901		
1735	3,460	107,234	3,532,368			14,069,993		
1736	5,310	330,579	3,466,925			14,112,560		
1737	3,720	67,284	3,401,232			13,896,247		
1738		269,837	3,333,207			13,882,763		
1739	10,528	283,854	3,276,861			13,883,284		
1740		196,245	3,211,323			13,797,939		
1741	9,486	25,232	3,156,393			13,546,707		
1742			3,093,265			13,275,773		
1743	7,440		3,038,691			13,010,258		
1744	7,837	9,812	2,985,598			12,759,668		
1745	1,860	292,966	2,927,708			12,791,582		
1746	136,431	474,492	3,002,857			13,000,752		
1747	4,650	37,146	2,947,357			12,777,140		
1748		338,523	2,888,409			12,853,350		
1749		710,687	2,830,641			13,292,756		
1750		558,597	2,774,028			13,574,326		

Note
Column 4 reports 1560 and 1561 estimated silver stock from Challis, and from 1562 the estimated silver stock assuming 2 per cent annual wastage.

wastage rate of 2 per cent per annum to the accumulating annual silver mint production figures through to 1750.

Estimates of the gold coin stock appear for 1344 to 1470 (Column 9), for 1471 (Column 8), for 1527 to 1543 (Column 9 again) and from 1559 (Column 7).

It is worth pausing to observe how far the resulting silver coin stock estimates match existing 'spot-estimates' of the size of the coinage.[8] It needs to be borne in mind that the current estimates are based on documented mint output, and are thus likely to understate the size of the actual silver coin stock, which also included undocumented output from especially the Durham and Bury St Edmunds mints, and the contribution of the Irish and Scottish mints to the total English silver coin stock.[9]

For the Tudor period the 'spot-estimates' are best provided by Christopher Challis (1978, 1992), while for the Civil War period, Edward Besly (1987) provides a guide to the likely size of the currency. For the late seventeenth and eighteenth centuries estimates of the size of the currency Rondo Cameron (1967) remains the best guide. On the whole these estimates provide an approximate correspondence with the annual figures given in Table 3.1. The match is very far from exact, but the general pattern indicated by both approaches is similar. There is a notable mis-match for gold in the first half of the seventeenth century, when a bimetallic flow of silver in and gold out might suggest that a wastage figure of 2 per cent per annum for gold in this period may be too low. Overall, however, there is sufficient general correspondence to suggest that the proposed model can generate annual figures for the stock of silver and gold coin in England which are broadly plausible.

Nevertheless it must be emphasised that the model does not offer any degree of precision. This is a very broad-brush approach. Given these reservations about the data, it follows that the regression analysis must be essentially experimental in character. But if we proceed on the assumption that the model provides a reasonable approximation to what actually occurred, how far did prices respond to the variables of population, GDP and gold and silver coin stocks?

The QT provides an explanation of the mechanism by which the money stock might influence prices. The QT is not the only explanation of price changes offered by historians, however. It has been argued that population changes were important too. The arguments advanced appear rather spurious, however. Population growth, it is said, increases commodity demand and thereby pushes prices up. This ignores the fact that population also increases the labour force, and so potentially increases supply as well. Indeed, according to the QT, rising population could have a deflationary impact if it increases the volume of business which the money supply must service. The work of the Broadberry and Campbell team shows that output grew faster than population over the period as a whole, so it is very unlikely that demand outstripped supply in the way that the population theory suggests (Mayhew, 2013).

While the traditional demographic explanation of the movement of prices is not convincing, our results also suggest that the positive correlation between prices and population is not statistically significant. Overall, it appears that the interplay (both short and longer term) between prices, population, GDP and money stock is more complex and nuanced than previously thought.

3.4 Regression analysis

Regression analysis was used to examine the QT. Two versions of the QT were investigated: a strong version and a weak version, as indicated above. The strong version implies that prices adjust almost instantaneously to changes in the money supply, while the weak version suggests that adjustment is a relatively slow process. Money supply is identified with notes and coin, and excludes bank deposits.[10]

The dependent variable is the price level. This reflects the assumption of the QT that price is endogenous and money supply and output (GDP) are exogenous. It is possible to 'first difference' the data to make the annual rate of inflation the dependent variable, but this is unnecessary in the present case. Prices are expressed in logarithms, as are the key explanatory variables. Using logarithms, differences in levels reflect proportional changes, and exponential trends are converted into linear trends. Logarithms are particularly useful in analysing the QT because the multiplicative relations on the each side of the Quantity equation are transformed into additive ones, which make them easier to investigate statistically.

The results are reported in Table 3.2. Two versions are given: column A shows the full results and column B the results when two insignificant dummy variables are eliminated. The two sets of results are broadly similar, indicating that the findings are robust. In each cell the estimated coefficient is shown on the top line and the probability value (level of significance) in brackets below. Three levels of significance are indicated: 1 per cent (***), 5 per cent (**) and 10 per cent (*).

The explanatory variables are listed on the left-hand side of the table. Lagged values of the key explanatory variables – money supply and GDP per head – are included. Lagged values of prices are also included; these reduce serial correlation in the residuals and thereby focus attention on the impact of current shocks rather than on the legacy of previous shocks.

The strong version of the QT implies that lags will be unimportant, while the weak version implies that they are crucial. Lags up to and including three years are included; trial and error suggests that the impact of longer lags is relatively insignificant. To simplify the interpretation of the results, the impact of lagged values are reported in terms of the impact of the level lagged one year; the change in the level between the current and the previous year; the change in the level between one year ago and two years ago; and the change in the level between two years ago and three years ago. This does not alter the basic results, since a span of four years (the current year and three previous years) is involved

Table 3.2 Estimation of price regression for England, 1273–1750

Variable	*Version A*	*Version B*
Constant	–0.422***	–0.409***
	(0.000)	(0.000)
Price level lagged one year	0.735***	0.741***
	(0.000)	(0.000)
Inflation rate lagged one year	0.059	0.055
	(0.308)	(0.336)
Inflation rate lagged two years	–0.121***	–0.124***
	(0.008)	(0.006)
GDP in England lagged one year	0.013	0.023
	(0.876)	(0.733)
Change in GDP	–0.540***	–0.544***
	(0.000)	(0.000)
Change in GDP lagged one year	–0.187**	–0.181**
	(0.038)	(0.034)
Change in GDP lagged two years	–0.072	–0.069
	(0.374)	(0.398)
Population in England lagged one year	0.111**	0.112***
	(0.040)	(0.004)
Stock of silver coin lagged one year	0.011**	0.011**
	(0.044)	(0.013)
Change in stock of silver coin	0.019	0.014
	(0.272)	(0.398)
Change in stock of silver coin lagged one year	–0.035*	–0.038**
	(0.065)	(0.037)
Change in stock of silver coin lagged two years	–0.014	–0.016
	(0.420)	(0.360)
Stock of gold coin lagged one year	0.017***	0.017***
	(0.000)	(0.000)
Change in stock of gold coin	–0.023***	–0.024***
	(0.001)	(0.001)
Change in stock of gold coin lagged one year	–0.007	–0.007
	(0.512)	(0.535)
Change in stock of gold coin lagged two years	–0.016**	0.017***
	(0.014)	(0.009)
Year, counting from 1270	–0.000*	–0.000***
	(0.076)	(0.001)
Dummy for unrecorded gold prior to 1344	0.180***	0.173***
	(0.000)	(0.000)
Dummy for change in estimates of silver stock in 1422	–0.007	–
	(0.565)	
Dummy for change in gold stock in 1558	0.038**	0.040**
	(0.020)	(0.011)
Dummy for Great Recoinage of 1696	–0.008	–
	(0.417)	
R^2	0.962	0.961
Adjusted R^2	0.960	0.960
F-statistic	501.08	555.153
	(0.000)	(0.000)
Durbin-Watson statistic	1.983	1.984
Number of observations	443	443

whatever combination of levels and changes is used. Long-run dynamics are captured by the coefficient on the lagged level and short-run dynamics by the coefficients on the changes in levels.

The lagged price level is highly significant and carries a coefficient of about 0.75; the fact that it is positive shows that shocks to price persist significantly, while the fact that it is less than one indicates that they decay over time. This suggests that the price level does not follow a pure random walk or other 'unit root' process. The significant negative coefficient on the inflation rate two years previously can also be explained by the persistence of shocks (see the appendix).

If the velocity of circulation were independent of GDP then according to the QT a 1 per cent increase in GDP would lead to a 1 per cent reduction in the price level, whereas the result for lagged GDP suggests that it does not change significantly at all. It seems that in the long run the deflationary effect of an increase in GDP is offset by an inflationary effect that neutralises it. The results for changes in GDP suggest a somewhat different picture, however. After controlling for the level of GDP, an increase in GDP in the current period reduces prices by 1.22 per cent, which is more in line with the QT, and the reduction increases to about 1.76 per cent if the effect of lagged changes is included as well.[11]

Population has an insignificant positive impact on the price level. Medieval population data is problematic, however, and is not sufficiently accurate to measure year-to-year changes;[12] as a result, annual changes in population have not been included in the regression. The absence of an inflationary impact of population growth implies that an increase in velocity of circulation did not occur when the currency had to circulate between a larger number of people. It appears that the impact of changes in population on prices are fully accounted for by the impact of population change on GDP.[13]

The main results relate to the money stock. They show that both the stock of gold and the stock of silver have significant positive impacts on the price level. The impact of gold is somewhat greater than that of silver, and more significant too; a 1 per cent increase in the lagged gold stock leads to a 0.048 per cent increase in the price level, while a similar increase in the silver stock increases prices by only 0.024 per cent. Both these effects are relatively small, however. The short-run dynamics provide a rather different picture, though, with changes in the stock of gold having very significant positive effects. Overall, the results for the money stock support the weak version of the QT rather than the strong one.[14]

There is a significant negative price trend, but it is very small – only 0.02 per cent per annum. There is little evidence of long-run institutional change affecting the velocity of circulation, other than the changes associated with movements in GDP, as noted above.[15]

Four dummy variables were included to control for potential structural breaks, but only two are significant. The dummy variable for unrecorded gold stock prior to 1344 is positive and highly significant, suggesting that foreign gold coin was almost certainly in circulation prior to the launch of the national gold coinage, and that much of it left the country thereafter. There appears to have been a significant increase in the velocity of circulation after the changes to the money stock in the 1550s.

The overall goodness of fit of the regressions is very high – over 96 per cent of price variation is explained. However, much of the explanation comes from the lagged price effects; if the regressions are re-specified using the rate of inflation as the dependent variable then the proportion of variance explained falls to about 33 per cent. The *F*-statistic shows that the regressions as a whole are highly significant.

3.5 Conclusions

Statistical analysis provides a very reassuringly objective basis for our conclusions, but it is important to remember that the estimates of gold and silver coin stocks are only estimates. Accordingly they should be subject to scrutiny, and improvement of these estimates should be an obvious target for future work. I have resisted the temptation to tweak these figures, because I wanted to explore how far a model based on documented output and the assumption of a fixed wastage rate for long periods might take us. However, the reality was undoubtedly different: some mint output has not left surviving documentation, and wastage rates probably varied a good deal from year to year. Nor should one forget the distinction, discussed above, between estimated coin stocks and the money stock.

Another area for future work might attempt a more nuanced understanding of the behaviour of the coin stock. Historians are occasionally aware of very large treasures held by the crown – particularly Henry VII – or by individual nobles. Moreover, unequal distribution of the available money stock will also have influenced hoarding and velocity.

Nevertheless, these initial results provide a sufficiently plausible model to encourage further work. It would be interesting to take this approach beyond 1750, when it might be possible to explore more precisely the relationship between coin and bullion stocks and bank notes under the Gold Standard. This could enable a more accurate assessment of the Money Stock.

For the moment, however, it seems sufficient to note the support which this analysis provides for a weak version of the Quantity Theory in the medieval and early modern period. This version of QT recognises the importance of lags, and the ‘sticky’ influence of earlier prices on current ones, and also allows a role for GDP, which is so closely related in this period to population. Over the centuries the relationship between the growth of the money supply, GDP and population emerges clearly.

Appendix

Derivation of the regression specification

Basic concepts

Let the aggregate money supply at time t, M_{st}, be a geometric average of the stock of minted silver sterling coin, M_{1t}, and the stock of minted gold sterling coin, M_{2t}:

$$M_{st} = M_{1t}{}^{b} M_{2t}{}^{1-b} \quad (t = 1, \ldots, T;\ 0 \leq b \leq 1) \tag{1}$$

If $b=0$ then only gold counts in the money supply, while if $b=1$ then only silver counts; if $b=½$ then they both carry equal weight.

Let the demand for money at time, M_{dt}, be directly proportional to both price, P_t, and output, Y_t,

$$M_{dt} = (1/v_t) P_t Y_t \quad (t = 1, \ldots, T) \tag{2}$$

where v_t is the velocity of circulation.

Let the velocity of circulation increase with respect to output. This is a 'Keynesian' effect: increased output increases the transactions demand for money and, for a given price level, pushes up interest rates, which raises the opportunity cost of holding money and encourages the faster circulation of money:

$$v_t = v_0 Y_t^{a} \quad (t = 1, \ldots, T;\ 0 < a < 1) \tag{3}$$

It is assumed that money supply and output are exogenous and that price is endogenous.

Strong Quantity Theory (SQT)

The supply and demand for money are in continual equilibrium:

$$M_{dt} = M_{st} \tag{4}$$

Substituting (3) into (2), then (1) and (2) into (4) gives

$$P_t = v_0 Y_t^{-(1-a)} M_{1t}{}^{b} M_{2t}{}^{1-b} \tag{5}$$

Taking logarithms gives a linear price equation:

$$\log P_t = \log v_0 - (1-a) \log Y_t + b \log M_{1t} + (1-b) \log M_{2t} \tag{6}$$

Weak Quantity Theory (WQT)

The WQT is based on an error-correction mechanism. The supply and demand for money are normally in disequilibrium, but there is a tendency towards a long-run equilibrium: price increases when there is an excess supply of money and falls when there is an excess demand. The speed of adjustment is proportional to the discrepancy between supply and demand.

The relative excess of money supply is

$$m_t = M_{st}/M_{dt} \quad (t=1,\ldots, T) \tag{7}$$

In long-run equilibrium $m=1$. In any period, t, the proportional change in price increases continually with respect to the relative excess of money supply in the previous period:

$$P_t/P_{t-1} = m_{t-1}{}^{k} \quad (t=1,\ldots, T;\ 0<k\leq 1) \tag{8}$$

When the supply of money exceeds demand excess money balances are spent, driving up the price level, while when demand for money exceeds supply hoarding reduces expenditure and prices fall. Disequilibrium is most likely to persist if the value of the adjustment parameter k is low.

Substituting (1), (3) and (7) into (8) gives

$$P_t = v_0{}^{k} P_{t-1}{}^{1-k} Y_{t-1}{}^{-k(1-a)} M_{1t-1}{}^{bk} M_{2t-1}{}^{(1-b)k} \tag{9}$$

Whence taking logarithms gives

$$\log P_t = k \log v_0 + (1-k) \log P_{t-1} - k(1-a) \log Y_{t-1} + bk \log M_{1t-1} + (1-b)k \log M_{2t-1} \tag{10}$$

In the limiting case $k=1$ equation (10) becomes similar to equation (6) of the SQT, except that lagged values rather than current values of money supply appear on the right-hand side.

Refinements

There are various ways of refining these models.

1 Include a time trend in the price equations; replace equation (3) by

$$v_t = v_0 exp(gt)\ Y_t^a \quad (t=1,\ldots, T;\ 0<g<1) \tag{11}$$

Exponential growth at a rate g stimulates continuous change in the velocity of circulation due to institutional changes. The logarithm of price acquires a linear trend with slope g.

2 Include a stochastic term w_t in the equation (11) to account for unexplained variations in price. Develop an encompassing model that includes both the SQY and WQT as special cases. This has the general form

$$\log P_t = a + b_1 \log P_{t-1} + b_2 \log Y_t + b_3 \log Y_{t-1} + b_4 \log M_{1t} + b_5 \log M_{1t-1} + b_6 \log M_{2t} + b_7 \log M_{2t-1} + b_8 t + w_t \quad 12)$$

where $b_1 = b_3 = b_5 = b_7 = 0$ if the SQT applies and $b_2 = b_4 = b_5 = b_6 = 0$ if the WQT applies.

3 Allow for persistent price shocks. When analysing monetary behaviour it is quite usual to find positive serial correlation in the estimated residuals, suggesting that price shocks may have persistent impacts. Suppose that the impact of shocks persists, but decays at an annual rate $h(0 \leq h < 1)$; this suggests that w_t is generated by a first-order autoregressive process involving serially uncorrelated shocks, u_t:

$$w_t = hw_{t-1} + u_t \quad (t = 1, \ldots, T;\ 0 < h < 1) \quad (13)$$

If $h = 0$ then the w_t are serially uncorrelated too. Substituting equation (12) into both the left-hand side and right-hand side of equation (13) and rearranging terms gives

$$\log P_t = a + h(1-a) + b_1 \log P_{t-1} + b_2 \log Y_t + (b_3 - hb_2)\log Y_{t-1} + b_4 \log M_{1t} + (b_5 - hb_4)\log M_{1t-1} + b_6 \log M_{2t} + (b_7 - hb_6)\log M_{2t-1} - hb_3 \log Y_{t-2} - hb_5 \log M_{1t-2} - hb_7 \log M_{2t-1} + b_8 t + u_t \quad (14)$$

4 Adjustments may be made to equation (14) to allow for specification errors introduced by omitted variables. These include introducing

- a population variable to allow for changes in behaviour caused by changing income per head, the impact of the Black Death, etc.;
- an additional lag to allow for more complex dynamics; and
- dummy variables to test for structural breaks associated with the introduction of gold coins, recoinages and other fundamental changes to the money base.

Together these adjustments generate the estimated equation in the text.

Notes

1 Martin Allen has pointed out that in addition to provincial mints, Scottish and Irish coins also contributed to the English currency; see Allen (2012).

2 Estimating undocumented output remains an option for further work. At this stage it seemed preferable to work only with the undisputed manuscript sources, since opinions vary about the scale of undocumented output. Nevertheless, it should always be borne in mind that the documented output is only a minimum figure.

3 Medieval kings forbade the export of sterling and organised searches of merchants' goods to enforce the ban, but some sterling was nevertheless carried abroad, and is

found in northern France, Germany, the Low Countries and Scandinavia, where its reputation for good weight and fineness ensured its acceptance. The increasing use of bills of exchange will have helped to reduce the need to export specie, and finds of sterling abroad are much rarer after the fourteenth century.

4 Much of all this was foreshadowed in Postan (1973), chs 1 and 2.

5 See especially the work of Nightingale (1990), based on surviving medieval Chancery Certificates of Debt.

6 On this point see, Mayhew (1995, p. 242). Attempts to quantify the use of sixteenth-century bills are not entirely satisfactory, though Kerridge (1988) collects much valuable evidence.

7 For a very bald summary of the movement of legal interests in sixteenth- and seventeenth-century England, see Mayhew (2000, pp. 60–62).

8 Existing 'spot-estimates' of the currency are collected in Mayhew (2013, Table 3).

9 This important consideration was first pointed out by Allen (2000, pp. 38–44). See also Allen (2012, Table 10.12). Allen gives an upper and lower figure, and I generally incline towards his lower estimates.

10 Note, however, that our money supply data is based solely on our estimate of silver and gold coin stocks, since it is not currently possible to estimate the quantity of monetary paper in use.

11 In practice over our period 1273–1750 M generally rises over the long term, allowing both (nominal and real) GDP and Prices to rise.

12 Although it could be argued that similar reservations should be entered about the estimates for annual money stock, prices and GDP as well, estimates for population before the mid sixteenth century are particularly disputed.

13 Estimates of pre-industrial GDP and population are highly likely to move in step. I have argued elsewhere that rising M could increase P, and that profit inflation could stimulate enterprise and economic growth, in turn permitting population growth (Mayhew, 2013).

14 This finding about the relative effects of gold and silver does not support a recent suggestion of mine that a silver based currency might be more inflationary than a gold based currency of the same value. This suggestion was prompted by the observation that fifteenth- and eighteenth-century prices moved extremely sluggishly, and that gold dominated the currency in both these periods. Coin finds indicate that gold coin did circulate less than silver, typically exhibiting less wear. Perhaps these distinctions failed to emerge in the QT, which reflects the annual turnover of the whole face value of the money stock. A pound face value in gold might be expressed in 3 gold nobles each worth 6 s 8 d, or in 240 silver pennies. Perhaps owners of gold nobles were more careful about spending it compared with spending 80 pence on various occasions. These essentially psychological factors elude the QT.

15 In fact, of course, over the very long term money stock increased markedly, permitting velocity to fall, despite rising GDP and population. See Mayhew (2013) citing Cameron (1967, p. 42) and Bordo and Jonung (1987, pp. 4–12).

References

Allen, Martin (2000) The volume and composition of the English silver currency, 1279–1351, *British Numismatic Journal*, 70, 38–44.

Allen, Martin (2012) *Mints and Money in Medieval England*, Cambridge: Cambridge University Press.

Allen, Robert C. (2013) Consumer price index for London is available at www.nuff.ox.ac.uk/People/sites/Allen/Shared%20Documents/Forms/AllItems.aspx. The London CPI appears in the excel database 'labweb' (accessed 15 April 2013).

Besly, E. (1987) *English Civil War Coin Hoards*, London: British Museum Occasional Paper 5.

Bordo, Michael D. and Lars Jonung (1987) *The Long-Run Behaviour of the Velocity of Circulation: The International Evidence*, Cambridge: Cambridge University Press.

Briggs, Christopher (2009) *Credit and Village Society in Fourteenth-Century England*, Oxford: Oxford University Press.

Britnell, R.H. (1986) *The Growth and Decline of Colchester, 1300–1525*, Cambridge: Cambridge University Press.

Broadberry, Stephen, Bruce M.S. Campbell, Alexander Klein, Mark Overton and Bas van Leeuwen (2009, 2011) British Economic Growth, 1270–1870: An Output-Based Approach (18 December 2011), at www2.lse.ac.uk/economicHistory/pdf/Broadberry/BritishGDPLongrun16a.pdf (accessed 15 April 2013).

Cameron, Rondo (1967) England 1750–1844, in Cameron, R. (ed.), *Banking in the Early Stages of Industrialization: A Study in Comparative Economic History*, Oxford: Oxford University Press, pp. 15–67.

Challis, C.E. (1978) *The Tudor Coinage*, Manchester: Manchester University Press.

Challis, C.E. (ed.) (1992) *A New History of the Royal Mint*, Cambridge: Cambridge University Press.

Clark, Gregory (2009) The Macroeconomic Aggregates for England, 1209–2008. Working papers, University of California, Department of Economics, no. 09–19, revised October 2009. www.ideas.repec.org/p/cda/wpaper/09–19.html (accessed 15 April 2013).

Gray, Hazel (2007) Money Lending in Twelfth-Century England. King's College, London PhD thesis.

Kerridge, Eric (1998) *Trade and Banking in Early Modern England*, Manchester: Manchester University Press.

Kowaleski, Maryanne (1995) *Local Markets and Regional Trade in Medieval Exeter*, Cambridge: Cambridge University Press.

Mayhew, N.J. (1995) Population, money supply, and the velocity of circulation in England, 1300–1700, *Economic History Review*, 48, 238–257.

Mayhew, Nicholas J. (2000) *Sterling: The History of a Currency*, Harmondsworth: Penguin.

Mayhew, Nicholas J. (2013) Prices in England, 1170–1750, *Past and Present*, May.

Muldrew, Craig (1998) *The Economy of Obligation: The Culture of Credit and Social Relations in Early Modern England*, Basingstoke: Macmillan.

Nightingale, Pamela (1990) Monetary contraction and mercantile credit in later medieval England, *Economic History Review*, 43, 560–576.

Phelps Brown, E. Henry and Sheila V. Hopkins (1962) Seven centuries of the prices of consumables, compared with builders' wage-rates, in E.M. Carus-Wilson (ed.), *Essays in Economic History*, London: Edward Arnold, pp. 179–196.

Postan, M.M. (1973) *Medieval Trade and Finance*, Cambridge: Cambridge University Press.

Schofield, P.R. and N.J. Mayhew (eds) (2002) *Credit and Debt in Medieval England, c.1180-c.1350*, Oxford: Oxford University Press.

Sriram, Subramanian S. (1999) Survey of literature on demand for money: Theoretical and empirical work with special reference to error-correction models, Washington DC, IMF Working Paper WP/99/64.

Thorold Rogers, James E. (1866–1902) *A History of Agriculture and Prices in England, 1259–1793*, 7 vols., Oxford: Oxford University Press.

4 Medieval foreign exchange

A time series analysis

Adrian R. Bell, Chris Brooks and Tony K. Moore

4.1 Introduction

This chapter will demonstrate some of the potential historical applications of econometric testing, focusing on a time series analysis of medieval foreign exchange rates. The work is part of a wider research project *Medieval Foreign Exchange, 1300–1500*, funded by the Leverhulme Trust under grant RPG193. The project is currently collecting a new database of medieval exchange rates and the following analysis is an exploratory study based on existing data sets intended to illustrate relevant forms of analysis. The first part of the chapter will briefly discuss the surviving evidence for medieval exchange rates. This is relatively extensive, at least by medieval standards, but there are a number of important caveats that need to be considered carefully. Despite this, exchange rate data offer some of the best opportunities to apply econometric analyses to the medieval economy. Recent studies have looked at volatility (Booth and Gurun, 2008), market integration (Volckart and Wolf, 2006; Chilosi and Volckart, 2011; Kugler, 2011) and implicit interest rates (Booth, 2011). With the exception of some early work by Hyman Sardy (de Roover, 1968), however, economic historians have not really used modern time series analysis to study medieval exchange rates. The second part of the chapter will use such techniques to test for seasonality, non-stationarity and structural breaks. Finally, it will compare the results of these tests against work on modern exchange rates as well as contemporary evidence from medieval merchants' manuals.

4.2 Sources in their historical context

Reliable quantitative data on the medieval economy are hard to find, including key economic indicators. For example, there is very little direct evidence for interest rates (Bell *et al*., 2009) or about trade and financial flows. Other information is more readily available – for instance, agricultural prices and wages can be extracted from manorial accounts (see Chapter 2), while final concords (fines) provide data about property values, as discussed in the following chapter. Another exception is foreign exchange rates (Einzig, 1970). This can be attributed to the confluence of three factors. The first was a general increase in long-distance trade following the 'commercial revolution of the thirteenth century' (Spufford, 1991). The second was the concurrent establishment of independent territorial states across Europe. Minting coins was both a key attribute of sovereignty and a source of income and so rulers sought to enforce the use of their local currency within their borders (Munro, 1979). As a result, merchants conducting long-distance trade needed to engage in foreign exchange. The third was the increasing sophistication of record keeping. In particular, the accurate recording and calculation of exchange rates was vital for merchants in order to keep their books balanced (de Roover, 1944a). While the vast majority of such records have been lost, enough survive to shed considerable light on medieval exchange rates.

Unfortunately, the surviving information about foreign exchange is scattered through a wide range of different sources. The most comprehensive data set of medieval exchange rates was compiled by Spufford (1986) in the 1970s and 1980s from correspondence with scholars across Europe. The *Handbook of Medieval Exchange* covers the period from 1106 to 1510, although the bulk of the data comes from the fourteenth and fifteenth centuries. It includes 13,197 exchange rate observations for 696 different currency pairs extracted from 528 different sources.[1] Unlike the more recent *Handbook of World Exchange Rates, 1590–1914* (Denzel, 2010), Spufford did not attempt to calculate annual average figures for each currency, as the extant data were not amenable to such treatment. Instead he presented individual raw exchange rates as found in the sources.

The aim of the *Handbook* was to provide indicative information about the relative values of different currencies that historians could use when drawing comparisons between countries. This meant that Spufford sought to cover as many currency pairs as possible over as long a period as possible. This inclusive approach necessarily generated a great variety in the nature of the evidence used. At the same time, there were also limitations imposed by the format of the *Handbook* as a printed book. Instead of listing all cross-exchanges between all currencies, Spufford concentrated on listing the exchange rates of each currency against the major international currency; first, the Florentine florin and, from the fifteenth century, the Venetian ducat. This is similar to the use of the dollar today. Moreover, even where more frequent data were available, Spufford only selected a maximum of one rate for each month.

As a result, the *Handbook* presents a number of problems for the purpose of quantitative analysis, although these can better be seen as the flip-side of its strengths. In particular, the *Handbook*'s inclusive approach to data collection, necessary given the scattered nature of the surviving source material, means that the data are not always strictly comparable. Spufford himself stresses this fact and one of the great strengths of the *Handbook* is that it provides information about the type of transaction, its location and, vitally, a reference to the original source, so that the exchange rates can be checked.

Most obviously, the Handbook includes exchange rates taken from a variety of different types of source and relating to different types of exchange transaction (Spufford, 1986, pp. l–liii). For example, spot exchange of one coin for another did not involve a time element whereas bills of exchange did, as will be shown below. The exchange rates used in accounts could also become fossilised at anachronistic rates or could be manipulated, for example to hide interest charges, while the official exchange rates promulgated by governments could be aspirational rather than effective. As a result, the exchange rates quoted in different types of transactions could vary considerably, making it difficult to draw valid comparisons.

Another important factor influencing medieval exchange rates, and a major difference to the modern situation, was the location of the transaction. In part, this reflected the slower speed of medieval communications. It also resulted from patterns of trade, which affected the demand for and supply of money differently at different places. More fundamentally, the bill of exchange, the classic foreign exchange instrument, did not just include a foreign exchange transaction but necessarily involved the extension of credit, since the buyer of the bill paid the seller upfront in local currency but only received the value in foreign currency later. Bills of exchange between different financial centres had different maturities (known as usance), mostly increasing with geographical distance. The exchange rates quoted at different financial centres therefore incorporated a spread to account for the time value of money. The operation of this system was demonstrated by de Roover (1944b) and has been traced back to the later thirteenth century by Blomquist (1990, pp. 362–368). It has even been argued that the primary significance of the bill of exchange was that it enabled merchants to circumvent the usury prohibition on charging interest (Koyama, 2010; Rubin, 2010). Other historians have argued that this over-states the credit aspect of the bill of exchange and downplays its role in facilitating international trade (Leone, 1983). The important point for our current purposes is that exchange rates varied depending on the location (or the direction) of the transfer and so rates at different places cannot be compared directly.

Furthermore, where exchange rates have been taken from records of individual transactions, there may be idiosyncratic factors influencing the rates quoted in that particular case. To take one example, in May and June 1305 the London branch of the Gallerani of Siena sold seven bills of exchange to customers wishing to remit money from England to the papal curia (Bigwood and Grunzweig, 1962). The rates quoted by the Gallerani varied from 5¾ florins per mark sterling on 21 May to 4½ florins per mark sterling on 7 June, a drop of 21.7 per cent in a little over two weeks. However, by 14 June the exchange rate had rebounded back up to 5¾ florins per mark sterling, an increase of 27.8 per cent in one week. It is unlikely that market exchange rates really fluctuated this wildly. A more plausible explanation is that the Gallerani offered different exchange rates to different customers. For instance, the three buyers that received the most favourable rates were all connected to the papacy, while the worst rate was received by an English clerk.

Another fundamental challenge concerns the frequency and distribution of the surviving evidence. Even for the best documented currency pairs, the *Handbook* records an average of only 1–2 observations per year. This means that idiosyncratic factors influencing individual exchange rates could potentially distort long-run trends. Also, there were strong seasonal variations in exchange rates, as will be shown below. Since the variation in exchange rates within each year was usually greater than the change from year to year, long-run developments might be obscured if the data for some years quoted rates from a seasonal peak while that for others came from seasonal lows. There are frequent gaps in the series, which raises problems of interpolation. Finally, the frequency of observations varies dramatically over time, which makes it difficult to conduct many types of statistical analysis.

In this chapter, we take a different approach to constructing a data set of medieval exchange rates. Rather than seeking to cover a long time span, we focus on a relatively short period for which we have abundant data. Instead of following an inclusive data collection policy, our approach is more exclusive and limited to rates from one particular type of source, namely mercantile correspondence. The merchant who had better and more up-to-date information about exchange rates in other banking centres enjoyed an advantage over his uninformed peers. For this reason, merchants often listed the current market exchange rates at the end of their commercial letters. These rates were probably collected from the bill brokers that arranged deals in each city (de Roover, 1968, p. 29). They can be seen as forerunners of the exchange rate currents printed from the sixteenth century onwards (McCuster and Gravesteijn, 1991), which ultimately developed into the modern financial press. The use of exchange rates from commercial correspondence has two main advantages. First, the market rates stripped out some of the idiosyncratic factors that may have influenced the exchange rates used in particular transactions. This makes the data more useful for comparative purposes, although historians should bear in mind Mueller's (1997, p. 588) warning against 'fetishiz[ing]' the rates quoted in mercantile correspondence since the rates used in actual transactions may have varied depending on the relative 'contractual leverage' of the two parties. Second, merchants wrote frequently to their correspondents (on a weekly or even daily basis) and often corresponded with several different cities, providing a greater depth and higher frequency of data.

This chapter draws upon the archive of one merchant in particular, namely Francesco di Marco Datini of Prato, near Florence in Italy. Datini was an extremely successful merchant, and between 1380 and 1410 his network of branches and correspondents covered much of western Europe (Origo, 1963; Nigro, 2010). The contents of the archive are described in Melis (1962) and many of the documents have been digitised and can be consulted online.[2] The potential of the Datini archive has long been recognised, and the commercial letters have been mined for data about exchange rates; de Roover (1968) extracted the exchange rates cited in letters from Barcelona and from Bruges while Mueller (1997) did the same for the letters from Venice.[3] The *Medieval Foreign Exchange* project is currently extending this data set to cover Florence and Genoa, the other two key banking centres in medieval Italy. The enlarged data set has not been finalised, however, so the following statistical analysis is based on the data collected by de Roover and Mueller.

4.3 Descriptive statistics

Time series analysis is a cornerstone of econometrics and the evaluation of a series' properties when viewed on its own should be a precursor to any solid multivariate research. We will first describe the data, in particular focusing on the statistical features of the series; we then move on to discuss the exchange rates' seasonal patterns and outline a framework for determining whether the series are best described as stationary or non-stationary processes. Finally, we consider in detail, using two different approaches, how to determine whether there were structural breaks.

As explained above, the following analysis draws on exchange rates quoted in merchants' letters from Barcelona, Bruges and Venice written between *c.*1385 and *c.*1410. The basic sources are shown in Table 4.1. The first two columns list the currency pair and the location. This effectively shows the direction of the exchange (e.g. Barcelona–Bruges is the exchange rate at Barcelona between the lira of Barcelona and the Flemish écu while Bruges–Barcelona is the rate at Bruges for the same pair). The third column describes how the exchange rate was quoted. The following analysis is based on monthly percentage changes in the exchange rates rather than raw levels but the method of quotation is still very significant for understanding the significance of the growth rates of the exchange rates discussed below. For example, Bruges 'gave certain' to Barcelona and London; that is, the exchange rate was quoted as an uncertain number of pence of Barcelona or pence sterling per écu of Bruges. Thus, a rise in the quoted exchange rate reflects an increase in the value of the écu. Bruges, however, 'gave uncertain' to Genoa, Pisa, Paris and Venice, meaning that exchange rates were quoted as a variable number of Flemish groats per unit of foreign currency. Thus an increase in these exchange rates actually means that the écu was declining in value.

The penultimate two columns show the total number of data points available for analysis and the main date range over which we have observations. We take monthly averages over all data points available for that month. It is clear that with the notable exceptions of Barcelona–Venice, Bruges–Pisa and Venice–Rome, where the samples are very small, for most of the series we have at least twenty years of monthly data or typically 200–300 observations. Unfortunately, it is in the nature of medieval data sources that there are inevitably missing values for some months. In order not to lose too many data points, we interpolate in such cases by rolling forward the value that was available for the previous month. The final column of Table 4.1 shows the number of interpolated values for each series. As can be seen, the number of such missing values is modest, and never more than 10 per cent of the total sample. We ensure that we never roll forward a data point for more than three months in a row – if there are more than three months of missing observations, we truncate the sample at that point.

We work mainly with the growth rates of the exchange rates, since we know that this analysis will be econometrically valid even if the raw rates contain unit roots, as discussed below. Thus we start by presenting the main features of the

Table 4.1 Definitions of currency pairs and available sample

Series	*Where quoted*	*Meaning*	*Total observations*	*Main date range*	*No. of interpolated values*
Barcelona–Avignon	Barcelona	Pence of Barcelona per French franc	229	1386:11–1405:11	25
Barcelona–Bruges	Barcelona	Pence of Barcelona per écu of 22 groats	224	1387:01–1405:11	20
Barcelona–Florence	Barcelona	Pence of Barcelona per Florentine florin	102	1394:10–1405:11	6
Barcelona–Genoa	Barcelona	Pence of Barcelona per florin of Genoa	229	1386:11–1405:11	19
Barcelona–Majorca	Barcelona	Pence of Barcelona per Majorcan real	132	1390:02–1405:09	19
Barcelona–Montpellier	Barcelona	Pence of Barcelona per French franc	218	1387:01–1405:11	15
Barcelona–Pisa	Barcelona	Pence of Barcelona per French franc	192	1387:01–1405:08	6
Barcelona–Venice	Barcelona	Pence of Barcelona per Venetian ducat	80	1399:02–1405:11	1
Bruges–Barcelona	Bruges	Pence of Barcelona per écu of 22 groats	194	1392:01–1410:09	5
Bruges–Genoa	Bruges	Flemish groats per florin of Genoa	196	1389:03–1410:09	12
Bruges–London	Bruges	Pence sterling per écu of 24 groats	190	1389:03–1410:09	12
Bruges–Paris	Bruges	Flemish groats per French franc	195	1389:03–1410:08	10
Bruges–Pisa	Bruges	Flemish groats per florin of Pisa	39	1395:03–1397:08	3
Bruges–Venice	Bruges	Flemish groats per Venetian ducat	196	1389:03–1410:09	14
Venice–Barcelona	Venice	Pence of Barcelona per Venetian ducat	139	1399:03–1410:09	1
Venice–Bologna	Venice	Florins of Bologna per Venetian ducat	269	1384:05–1410:09	14
Venice–Bruges	Venice	Flemish groats per Venetian ducat	251	1384:03–1410:09	2
Venice–Florence	Venice	Pence affiorino of Florence per Venetian ducat	320	1384:02–1410:09	10
Venice–Genoa	Venice	Florins of Genoa per Venetian ducat	319	1384:03–1410:09	16
Venice–London	Venice	Pence sterling per Venetian ducat	65	1403:10–1410:09	7
Venice–Lucca	Venice	Florins of Lucca per Venetian ducat	135	1399:04–1410:06	10
Venice–Milan	Venice	Ducats of Milan per Venetian ducat	136	1384:03–1404:03	6
Venice–Paris	Venice	Grossi d'oro of Venice per French franc	177	1388:08–1410:09	7
Venice–Pisa	Venice	Florins of Pisa per Venetian ducat	266	1384:03–1410:05	14
Venice–Rome	Venice	Florins of Rome per Venetian ducat	99	1394:06–1407:08	1

exchange rates in Table 4.2. The table presents summary statistics for the monthly percentage changes of the twenty-five series that we examine spanning the three main venues of Barcelona, Bruges and Venice.

It is evident that the Venetian means are all positive (except for Venice–Paris where Venice 'gave uncertain' to Paris), indicating the strength of the ducat over the period. In particular, the ducat rose by around 0.2 per cent per month (2.4 per cent per annum) against the Roman florin and by 0.14 per cent per month (1.6 per cent per annum) against the pound sterling. This reflects the maintenance of the gold content of the ducat, as well its increasing displacement of the Florentine florin as the major international coin (Spufford, 1991, p. 321).

The variance estimates, presented in the fourth column, are broadly similar to modern figures, which would be of the order of 3–5 per cent per month. The variance, sometimes known in finance as the volatility of a series, measures the extent to which it moves around over time. It thus shows the spread of the observations around their mean value. However, we can note considerable differences in volatilities across the series – from just 0.8 per cent for Barcelona–Majorca to 9.5 per cent for Venice–Bruges and 13 per cent for Barcelona–Bruges. The former probably reflects the fact that Barcelona and Majorca were both part of

Table 4.2 Summary statistics for exchange rate returns (percentage changes)

Series	*Mean*	*Variance*	*Skewness*	*Kurtosis (excess)*	*Min*	*Max*
Barcelona–Avignon	–0.01	1.96	0.20	1.54	–4.45	4.94
Barcelona–Bruges	0.09	13.02	9.35	119.04	–6.45	46.15
Barcelona–Florence	0.00	2.60	0.54	2.82	–3.92	6.80
Barcelona–Genoa	0.01	3.48	–0.53	2.12	–8.42	5.94
Barcelona–Majorca	0.05	0.78	–0.45	3.52	–3.95	2.84
Barcelona–Montpellier	0.00	2.18	0.33	2.41	–4.57	6.11
Barcelona–Pisa	0.02	3.39	–0.10	3.10	–7.03	8.50
Barcelona–Venice	0.02	1.90	–0.12	0.29	–3.24	3.94
Bruges–Barcelona	–0.05	2.28	0.01	1.35	–4.99	4.53
Bruges–Genoa	0.00	2.50	–0.28	1.00	–5.64	4.49
Bruges–London	–0.04	1.43	0.44	1.54	–4.08	3.85
Bruges–Paris	0.06	0.81	–0.23	2.42	–3.4	3.41
Bruges–Pisa	0.12	1.47	–0.44	1.20	–3.47	2.50
Bruges–Venice	0.12	9.49	2.68	62.85	23.08	30.3
Venice–Barcelona	0.02	1.99	–0.11	1.50	–5.00	4.35
Venice–Bologna	0.00	0.57	0.02	3.63	–2.97	3.51
Venice–Bruges	0.04	2.22	–0.23	2.22	–6.77	4.83
Venice–Florence	0.01	1.07	–0.08	2.78	–3.96	5.19
Venice–Genoa	0.07	1.48	–0.08	1.34	–3.96	3.98
Venice–London	0.14	1.76	0.48	1.50	–2.67	4.44
Venice–Lucca	0.01	0.94	0.10	4.98	–4.32	3.93
Venice–Milan	0.10	5.97	–5.27	48.98	21.85	7.42
Venice–Paris	–0.03	1.86	–0.13	0.60	–4.08	4.34
Venice–Pisa	0.02	1.10	0.29	4.09	–3.88	5.40
Venice–Rome	0.20	2.26	1.03	3.81	–3.04	7.03

the Aragonese realm. The latter are almost certainly caused by outliers in the percentage change series for these currency pairs as a result of currency revaluations in Bruges (Munro, 2012).

In terms of their symmetry or otherwise, some of the series are left-skewed and some are right-skewed; all series are leptokurtic, implying that they have more in the mean, fatter tails but less in the shoulders of the distribution compared with a normally distributed series having the same mean and variance. Again, this feature is very similar to that in standard contemporary asset return series. This leptokurtosis is a key feature of financial time series from a risk management perspective since it implies that extreme movements (i.e. very large swings) are more likely than would be the case under a normal distribution. Barcelona–Bruges is the most skewed and has the highest kurtosis, in particular due to an almost 50 per cent rise between October and November 1390. This reflects a curious delayed impact of the enhancement of the Flemish currency in 1389. The Italian merchants at Barcelona continued to quote the exchange rate with Bruges in terms of the old money until November 1390, when they switched to using the new money (de Roover, 1968, pp. 39–40).

4.4 Seasonal patterns in exchange rates

Since bills of exchange were used both to transfer money for trade and also to borrow or lend money, exchange rates were closely linked to the wider condition of the money market. The sixteenth-century merchant Bernardo Davanzati explained this relationship using the analogy of a hand tightening or loosening its grip on money (Mueller, 1997, p. 305). At times of high demand for cash, the hand would tighten and not release any money except at a higher price. As a result, the exchange rate would rise, that is, sellers of bills of exchange (borrowers) would have to promise more foreign currency to receive one unit of local currency. In this case, merchants described money as being ‘dear’ (*carestia*) or ‘tight’ (*strettazza*). In the contrary situation, when the supply of money exceeded the demand, money was ‘loose’ (*larghezza*) or ‘abundant’ (*dovizia*) and exchange rates fell as buyers of bills of exchange (lenders) were prepared to accept less foreign currency per unit of local currency. According to Giovanni di Antonio da Uzzano, the ‘good rule in making exchange’ was to anticipate changes in exchange rates and not merely to react to them (Pagnini, 1766, p. 153). It is therefore important to know whether there were predictable seasonal patterns in medieval exchange rates.

There are various ways to test for seasonality in time series data. It is possible to employ trigonometric functions or to work in the frequency domain. However, in such cases the quantitative sophistication arises at the expense of interpretability. A much simpler approach, which we apply here, is to use a linear regression including monthly dummy variables

$$\begin{aligned} y_t = {} & \alpha_1 D1_t + \alpha_2 D2_t + \alpha_3 D3_t + \alpha_4 D4_t + \alpha_5 D5_t + \alpha_6 D6_t + \alpha_7 D7_t + \\ & \alpha_8 D8_t + \alpha_9 D9_t + \alpha_{10} D10_t \alpha_{11} D11_t + \alpha_{12} D12_t + u_t \end{aligned} \tag{4.1}$$

where y_t is the exchange rate percentage change series under consideration, $D1_t$, $D2_t \ldots D12_t$ are monthly dummy variables for January, February,…, December, and u_t is an error term, assumed to be normally distributed with zero mean and constant variance. These dummy variables take the value one for the month to which they correspond and zero otherwise – so, for example, $D1_t$ takes the value one every January and zero for every other month. This way, the dummies effectively 'pull out' the observations for their corresponding months and set everything else to zero. Thus each parameter attached to the dummies can be interpreted as the average change in the foreign exchange rate for that month (in percentage terms).

Note that, given the way that it has been specified to contain a full set of twelve monthly dummy variables, this regression must not contain an intercept term to avoid the 'dummy variable trap'. This would have arisen if all possible seasonal dummies given the frequency of the data employed (i.e. twelve for monthly, four for quarterly, etc.) were included in the model together with an intercept. The result would be that the regression could not be run due to multicollinearity (see Chapter 1).

The seasonality results are presented in Table 4.3 for the rates quoted in Barcelona, in Table 4.4 for the rates quoted in Bruges and in Table 4.5 for those quoted in Venice. Overall, these tables indicate very clear patterns. This quantitative evidence can be used to test the qualitative descriptions of the state of the money market by contemporary merchants. In this case, we shall use Giovanni di Antonio da Uzzano's *Practica Della Mercatura* (Pagnini, 1766). It should be noted that Uzzano was writing *c.*1442, roughly fifty years after our data from the Datini letters, and this may explain some of the discrepancies between our statistical reconstruction and Uzzano's experience.

The foreign exchange rates quoted in Barcelona rise in December and most significantly in January and this continues at a more modest pace until around April, followed by large reversals in June through to October, with the most significant falls in the autumn months. In all cases, these exchange rates were quoted as an uncertain number of pence of Barcelona for a certain number of foreign coins. Thus a rise in exchange rates at Barcelona meant that the lira was decreasing in value, reflecting an easing of the money market. Conversely, a fall in exchange rates indicates an increase in the value of the *lira* and thus a tightening of the money market. This agrees with the first part of the account given by Uzzano:

> In Barcelona, money is dear from the first of June through all of August because of the investments in wool from Aragon and the surrounding valleys and because of the purchases of 'grain' [the dye] in Valencia; the money market tightens again in October, after St Luke's day, which is on the 18th, because of the investments in saffron, when the dearness is even greater than in the wool season, and it will last until January; and from then on, money eases every day and the exchange rates return to their former level, and the easiness lasts until the wool season, unless something unexpected happens.
>
> (Pagnini, 1766, p. 156; translated in de Roover, 1968, p. 88)

Table 4.3 Seasonal variations – Barcelona

	Avignon	*Bruges*	*Florence*	*Genoa*	*Majorca*	*Montpellier*	*Paris*	*Venice*
January	0.87***	1.68**	1.63***	1.85***	0.84***	0.94***	2.17***	0.85
	(0.30)	(0.81)	(0.53)	(0.39)	(0.24)	(0.33)	(0.42)	(0.53)
February	0.39	1.06	0.75	0.86**	0.61***	0.52	0.50	0.81
	(0.30)	(0.81)	(0.53)	(0.39)	(0.23)	(0.32)	(0.42)	(0.49)
March	–0.02	0.16	–0.35	0.14	0.12	–0.02	0.54	0.39
	(0.30)	(0.83)	(0.53)	(0.39)	(0.25)	(0.33)	(0.44)	(0.49)
April	0.22	0.29	0.59	0.31	0.22	0.22	0.38	0.86*
	(0.30)	(0.83)	(0.53)	(0.39)	(0.24)	(0.33)	(0.42)	(0.49)
May	0.23	0.04	–1.16**	–0.90**	0.45*	0.24	–0.29	–0.91*
	(0.30)	(0.83)	(0.53)	(0.39)	(0.25)	(0.32)	(0.41)	(0.49)
June	0.34	–0.21	–0.55	–0.71*	–0.02	0.26	–0.47	–0.23
	(0.30)	(0.81)	(0.53)	(0.39)	(0.24)	(0.32)	(0.41)	(0.49)
July	–0.43	–0.77	–0.13	–0.68*	–0.25	–0.24	–0.80*	–0.54
	(0.30)	(0.81)	(0.50)	(0.39)	(0.24)	(0.33)	(0.41)	(0.49)
August	–0.53*	–0.94	–0.48	–0.47	–0.14	–0.84**	–0.36	–0.75
	(0.30)	(0.81)	(0.50)	(0.39)	(0.24)	(0.33)	(0.42)	(0.49)
September	–1.17***	–1.8**	–1.20**	–0.72*	–0.16	–1.18***	–1.04*	–1.12**
	(0.30)	(0.81)	(0.50)	(0.39)	(0.23)	(0.33)	(0.44)	(0.53)
October	–0.84***	–1.1	0.20	–0.65*	–0.81***	–0.67**	–0.65	0.1
	(0.30)	(0.81)	(0.47)	(0.39)	(0.22)	(0.33)	(0.42)	(0.53)
November	0.07	2.17***	0.04	–0.15	–0.31	0.08	–0.09	0.42
	(0.29)	(0.81)	(0.50)	(0.38)	(0.25)	(0.33)	(0.42)	(0.49)
December	0.71**	0.54	0.77	1.28***	0.19	0.67**	0.46	0.38
	(0.30)	(0.83)	(0.53)	(0.39)	–0.25	(0.34)	(0.44)	(0.53)
R^2	0.18	0.10	0.24	0.21	0.25	0.17	0.20	0.24

Notes
Standard errors in parentheses; *, ** and *** denote significance at the 10 per cent, 5 per cent and 1 per cent levels; critical values vary according to the number of observations for each series. All the rates are quoted as an uncertain number of pence of Barcelona per unit of foreign currency.

Table 4.4 Seasonal variations – Bruges

	Barcelona	*Genoa*	*London*	*Paris*	*Pisa*	*Venice*
January	–0.23	1.09***	–0.57**	0.39*	1.16*	0.82
	(0.37)	(0.36)	(0.28)	(0.22)	(0.66)	(0.76)
February	–0.65*	0.56	–0.58**	0.11	0.25	0.49
	(0.36)	(0.35)	(0.28)	(0.21)	(0.57)	(0.73)
March	–0.58	0.33	–0.49*	0.42**	0.51	0.60
	(0.36)	(0.34)	(0.28)	(0.21)	(0.51)	(0.71)
April	0.41	–0.92***	–0.53**	–0.25	–0.15	–0.82
	(0.36)	(0.34)	(0.27)	(0.21)	(0.51)	(0.71)
May	0.3	–0.51	–0.55*	0.01	–0.02	–0.25
	(0.36)	(0.35)	(0.28)	(0.21)	(0.57)	(0.73)
June	0.08	–1.38***	0.14	–0.04	–1.23**	–0.40
	(0.36)	(0.35)	(0.28)	(0.21)	(0.57)	(0.73)
July	0.11	–0.48	0.12	–0.58***	–0.78	–0.68
	(0.37)	(0.36)	(0.28)	(0.22)	(0.66)	(0.76)
August	–0.4	0.41	0.03	0.12	0.38	–1.52**
	(0.37)	(0.36)	(0.28)	(0.22)	(0.66)	(0.76)
September	–0.56	0.08	0.96***	0.36	1.15	2.12***
	(0.37)	(0.36)	(0.28)	(0.23)	(0.81)	(0.76)
October	0.21	0.40	0.84***	0.17	1.49*	0.21
	(0.38)	(0.38)	(0.29)	(0.23)	(0.81)	(0.78)
November	0.95**	–0.12	0.37	–0.14	–0.16	–0.01
	(0.38)	(0.38)	(0.29)	(0.23)	(0.81)	(0.78)
December	–0.11	0.80**	–0.15	0.18	–0.15	1.09
	(0.38)	(0.38)	(0.29)	(0.23)	(0.81)	(0.78)
R^2	0.09	0.20	0.19	0.10	0.37	0.09

Notes
The rates for Genoa, Paris, Pisa and Venice are quoted as an undertain number of Flemish groats per unit of foreign currency while Barcelona and London are quoted as an uncertain number of foreign coins per écu (of twenty-two and twenty-four groats respectively).

The Datini data support the tightness in the summer and autumn months (June to October) but the market seems to ease in the winter, when Uzzano suggests that the market should have continued to tighten and to peak. It is possible that the Spanish saffron trade was less important during the Datini period than it later became, although Datini himself traded in saffron (Origo, 1963, p. 97).

For the foreign exchange rates quoted in Bruges, the pattern is less clear-cut, with most of the changes not being statistically significant. In general, we see that the exchange rates for Bruges–Barcelona and Bruges–London fell in the winter months from December to March (May for London) and rose from April (June for London) until November, while the other rates fell in the summer and rose in the winter and spring. Here it is important to recall that Bruges 'gave certain' to both Barcelona and London and a rise in these exchange rates therefore reflects a rising value for the Flemish *pond groot* and thus a tightening of the money market while a fall in the same rates reflects an easing of the market. However, the exchange rates with Genoa, Paris, Pisa and Venice were quoted as

Table 4.5 Seasonal variations – Venice

	Barcelona	*Bologna*	*Bruges*	*Florence*	*Genoa*	*London*	*Lucca*	*Milan*	*Paris*	*Pisa*	*Rome*
January	0.05	–0.15	0.11	–0.19	–0.33	–0.64	0.14	0.74	0.38	–0.19	–0.01
	(0.40)	(0.15)	(0.35)	(0.18)	(0.22)	(0.59)	(0.27)	(0.68)	(0.32)	(0.20)	(0.57)
February	–0.46	–0.09	–1.03***	–0.45**	–0.33	–0.51	–0.40	–0.70	0.80**	–0.46**	–0.05
	(0.40)	(0.15)	(0.36)	(0.18)	(0.22)	(0.59)	(0.27)	(0.68)	(0.33)	(0.20)	(0.57)
March	–0.68*	0.04	0.11	0.18	0.10	–0.24	0.12	–0.66	0.20	0.10	–0.13
	(0.38)	(0.15)	(0.34)	(0.18)	(0.22)	(0.59)	(0.27)	(0.68)	(0.33)	(0.19)	(0.53)
April	–0.29	0.01	0.04	–0.04	0.20	–0.45	0.29	0.40	–0.32	0.01	0.09
	(0.38)	(0.15)	(0.34)	(0.18)	(0.22)	(0.54)	(0.26)	(0.78)	(0.34)	(0.19)	(0.57)
May	0.01	–0.18	0.01	0.13	0.64***	1.14**	0.09	0.16	0.04	0.01	0.16
	(0.38)	(0.15)	(0.35)	(0.18)	(0.22)	(0.54)	(0.26)	(0.78)	(0.34)	(0.19)	(0.50)
June	0.86**	0.02	1.37***	0.56***	0.62***	0.57	0.60**	1.02	–1.01***	0.71***	0.29
	(0.38)	(0.15)	(0.33)	(0.18)	(0.22)	(0.54)	(0.26)	(0.74)	(0.33)	(0.20)	(0.45)
July	0.77**	0.39**	0.60*	0.58***	0.57***	0.55	0.67**	0.46	–0.89***	0.91***	1.11**
	(0.38)	(0.15)	(0.34)	(0.18)	(0.22)	(0.54)	(0.27)	(0.74)	(0.33)	(0.20)	(0.47)
August	0.33	0.20	–0.04	–0.24	0.17	0.56	–0.71***	–0.10	–0.46	–0.16	0.81*
	(0.38)	(0.15)	(0.33)	(0.18)	(0.22)	(0.54)	(0.27)	(0.74)	(0.32)	(0.19)	(0.45)
September	–1.27***	–0.64***	–0.81**	–1.29***	–1.11***	–0.30	–0.83***	–1.02	0.81**	–1.35***	–0.76
	(0.38)	(0.15)	(0.33)	(0.18)	(0.22)	(0.59)	(0.27)	(0.74)	(0.33)	(0.20)	(0.53)
October	0.42	0.01	0.66*	0.51***	0.39*	0.24	0.28	0.38	–0.39	0.39*	0.61
	(0.40)	(0.15)	(0.34)	(0.18)	(0.22)	(0.59)	(0.27)	(0.78)	(0.34)	(0.20)	(0.57)
November	–0.28	0.17	0.07	0.32*	–0.03	0.68	–0.20	–0.04	0.57*	0.16	–0.43
	(0.40)	(0.15)	(0.34)	(0.18)	(0.22)	(0.59)	(0.27)	(0.78)	(0.34)	(0.20)	(0.57)
December	0.78*	0.17	–0.10	0.08	–0.14	–0.22	–0.01	0.69	–0.09	0.02	0.07
	(0.40)	(0.15)	(0.34)	(0.18)	(0.22)	(0.59)	(0.27)	(0.68)	(0.34)	(0.20)	(0.57)
R^2	0.20	0.11	0.14	0.24	0.16	0.18	0.21	0.07	0.19	0.27	0.13

Notes

All the exchange rates are quoted as an uncertain number of foreign coins per Venetian ducat, except for Venice–Paris.

an uncertain number of Flemish groats for a certain number of foreign coins, and so the reverse logic applies. The evidence of the Datini letters indicates that the Bruges money market tightened in the summer and eased in the winter. Interestingly, this directly contradicts Uzzano's depiction of the seasonal trends in the Bruges' money market *c.*1442:

> In Bruges, money is dear in December and January because of the many ships that are being loaded with commodities and dispatched at this time, and, in August and September, money expands because of the fairs that are being held and that attract merchants who come to purchase and bring in ready cash.
>
> (Pagnini, 1766, p. 156; translated in de Roover, 1968, p. 90)

This discrepancy would repay further historical investigation into the seasonal patterns of trade in medieval Bruges. In particular, were there significant changes between the later fourteenth and early fifteenth centuries, when the Datini letters were written, and the mid-fifteenth century, when Uzzano wrote his manual, that might explain the stark differences set out above?

For the rates quoted in Venice, there are mostly significant falls in September and in February but rises in June and July. As Venice 'gave certain' to most other currencies, this means that the market was tightening in the summer and loosening after September. We may note that Venice–Paris appears to show the opposite pattern, because here Venice 'gave uncertain' to Paris, but in fact it shares the same underlying trend. This supports Uzzano's depiction of the Venetian money market:

> In Venice, money is expensive from May to 8th September, because of the outward bound galleys which leave in July, August and September. The reason why it gets more expensive is because everyone starts to make arrangements and they want to remit more there; and this higher cost is due to the amount of cash the galleys carry, because a great deal of merchandise is sold there at the time of the galleys, which must be paid for just when you have many demands on your purse – and a lot of money goes out of the banks in cash, so cash is always dear there by 1 per cent more than usual. And money is highly priced for all places, and is offered there at various maturities. From 8th July money is highly priced, then there are no more maturities until 1st August, and in this month there is an expansion by ½ to 1 per cent. From 1st August money starts to fluctuate, and is expensive continually until 8th September; and after the 8th all maturities have become due, and all the galleys have gone, so there is no more demand – and the banks are quick to supply and money goes through the floor.
>
> (Pagnini, 1766, pp. 156–157; translation by Dr Helen Bradley)

Mueller (1997, pp. 305–307) adds further details. The galleys to Romania left Venice in mid-July, those to Beirut in mid-August, and finally the Alexandria galleys in late August to early September. Since western Europe had a trade deficit with the East, merchants had to export silver bullion to purchase goods for re-import. This led to a great demand for money in the summer, which drove up exchange rates. Subsequently, after the last galleys had departed, the demand for money reduced and thus exchange rates collapsed in September. Uzzano also mentions a later tightening of the money market in December and January, linked to the departure of galleys to Catalonia (Pagnini, 1766, p. 157). This fits neatly with the sharp rise in the Venice-Barcelona exchange rate in December shown in Table 4.5.

A further important issue is whether we analyse the seasonal patterns in the levels of the series or in their percentage changes. Where a series in levels has little underlying trend, it is likely to make very little difference. For instance, our results in Table 4.5 for Venice–Florence are very similar to the patterns in Graph 8.1 of Mueller (1997, p. 307) where the sample period is very close to ours and Table 4.2 indicates that over time the mean change in the exchange rate was only 0.01 per cent per month. On the other hand, the pattern that we observe for Barcelona-Bruges is quite different to that reported by Hyman Sardy (de Roover, 1968) and this may be a result of his use of levels on a series where the mean percentage increase in the exchange rate was 0.09 per cent per month, almost ten times higher than that for Venice–Florence. When a series is trending heavily over time – either up or down – then any statistical analysis using the levels will not have the meaning that the researcher probably intended. In the case of a series trending upwards, an analysis of the seasonal patterns will give undue weight to the observations at the end, which by definition will be much larger in value than those close to the beginning. Thus we would recommend historians to use percentage changes rather than levels in such cases.

The R^2 figures from these regressions measure the degree to which the seasonal patterns can explain the overall variation in each series. In an earlier analysis of the same data from Bruges and Barcelona, Hyman Sardy found that 'about ten per cent of the fluctuations in the series could be attributed to seasonality' (de Roover, 1968, p. 104). Our work suggests that seasonality played a greater role, accounting for around 20 per cent of the variation in most series. This is highest for Bruges–Pisa (37 per cent of the variation explained, although based on the smallest sample) and for Venice–Pisa (27 per cent explained). This suggests that Pisa was particularly affected by seasonal flows of trade or finance. Seasonality is less significant for Bruges–Paris (10 per cent) and Venice–Bologna (11 per cent). Uzzano explains that there was a particularly close connection between exchange rates at these and a number of other places (Pagnini, 1766, p. 154), which Mueller describes as 'paired cities', since their geographical proximity meant that information could be transmitted between the two in 1–2 days, allowing merchants in one location to react quickly to changes at the other (Mueller, 1997, pp. 588–589).

4.5 Testing for unit roots in exchange rates

A key early question in analysing time series is whether each variable in a sample can be considered stationary or whether it contains a unit root, which is so-called because the root of a characteristic equation for such a process is unity. A stationary series is one with a constant mean, constant variance and constant autocovariance structure for a given lag – in other words, the relationship between the current value of the time series and its previous values remains constant. This is probably the most important characteristic of a series as it has the most significant impact on its properties and also on which type of analysis is most appropriate. There is also the issue that if a series contains a unit root, standard econometric approaches cannot be applied to the data in their raw, levels form, otherwise this would result in 'spurious regression' where entirely independent unit root processes appear to standard econometric approaches to be strongly related. Thus, in such cases, the series must be converted into a percentage changes (i.e. growth rates) form.

However, whether exchange rate series in levels are non-stationary is surprisingly still an open question, with much empirical evidence both for and against. In the 'modern finance' literature, much research has already been undertaken in this regard, even when we focus specifically on exchange rates. Many authors use the augmented Dickey–Fuller (ADF) test (Dickey and Fuller, 1979), described below, on data of monthly or quarterly frequency. Meese and Singleton (1982), for example, test for the presence of a unit root in the log of a number of weekly US dollar-denominated exchange rates from the 1970s. They cannot reject the null hypothesis of a unit root in the log of the levels, but the percentage changes are found to be stationary, a result echoed by Corbae and Ouliaris (1988) in their tests on monthly sterling data.

More recently, Whitt (1992) compared the results of a test for unit roots on the real US dollar exchange rate against a number of others using the ADF approach and Sims' Bayesian test. The ADF test cannot reject the non-stationary null, but the Bayesian test does strongly reject it. Some evidence for stationarity in real exchange rates was also found by Taylor (1990). In a comparative study, Schotman and van Dijk (1991) use the ADF and Sargan Bhargava unit root tests, together with a Bayesian posterior odds approach. The ADF and Sargan-Bhargava tests both lead to non-rejection for all the exchange rates they test, while the Bayesian method does not provide conclusive results.

Finally, Goodhart *et al*. (1993) conduct ADF and Phillips–Perron tests on a number of US dollar-based exchange rates at sampling frequencies from tick-by-tick to daily. The Phillips-Perron procedure is very similar to the ADF test described below except that it incorporates an automatic correction for autocorrelated residuals in the test regression. Their main finding is that daily and hourly series have a unit root, while most minute-by-minute and tick-by-tick series have a unit root with trend. They conclude that temporal aggregation preserves the non-stationarity in exchange rates so that the series are non-stationary whether they are observed hourly, every minute or as the transactions occur.

If a series y_t is initially non-stationary, but becomes stationary on differencing d times, then it contains d unit roots and is said to be integrated of order d. We 'first difference' a series by subtracting the immediately previous value of that series from the current one. It is then possible to second difference a series by applying the same process again to the series that has already been differenced once. A unit root is only one possible form of non-stationarity, an explosive root being the other case, although the latter is rarely considered since it is more difficult to justify from an economic theoretical perspective. Many economic series contain an exact unit root when transformed into logarithms (Banerjee *et al.*, 1993, p. 99). Thus if we find that the raw data contain a unit root, but the percentage changes are stationary, this provides an *ex post* justification for using the percentage changes in subsequent analysis.

This section will now give a brief description of the Dickey–Fuller approach to testing for a unit root in time series. They conducted the early and pioneering work on this topic, and despite numerous advances in the testing theory since then, their technique still constitutes the main workhorse of unit root testing – see Dickey and Fuller (1979) and Fuller (1976). The Dickey–Fuller (DF) tests used in this chapter are the based on a t-ratio from a test regression. It is possible to include either a constant, or a deterministic trend, or both or neither in the test regression. Consistent with the expected features of the data, we elect to adopt the model including an intercept but not a trend. The null (H_0) and alternative (H_1) models for the test we implement are thus

$$H_0\colon y_t = y_{t-1} + u_t \tag{4.2}$$

$$H_1\colon y_t = \varphi y_{t-1} + \mu + u_t,\ \varphi < 1 \tag{4.3}$$

This is a test for a random walk against a stationary autoregressive model of order one (AR(1)) with a drift. We can thus write

$$\Delta y_t = u_t \tag{4.4}$$

under the null hypothesis, where $\Delta y_t = y_t - y_{t-1}$, and the alternative may be expressed as

$$\Delta y_t = \psi y_{t-1} + \mu + u_t. \tag{4.5}$$

The test for whether the series contains a unit root is based on the t-ratio of the y_{t-1} term in the estimated regression and thus the test statistic is defined as

$$\text{Unit root test statistic} = \frac{\hat{\psi}}{\widehat{SE}(\hat{\psi})} \tag{4.6}$$

The test statistic does not follow the usual t-distribution under the null hypothesis, since the null is one of non-stationarity, but rather it follows a nonstandard distribution. Critical values are derived from Monte Carlo experiments in, for example, Fuller (1976). The null hypothesis of a unit root is rejected in favour of the stationary alternative in each case if the test statistic is more negative than the critical value.

The test may also be 'augmented' by the addition of p lags of the dependent variable to the estimated equation, known as the augmented Dickey–Fuller test, which allows for possible autocorrelation of residuals in the regression. The unit root test would not perform well if such a structure was present in the residuals of the test regression but unaccounted for. The model under the alternative in this case can now be written as

$$\Delta y_t = \mu + \psi y_{t-1} + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + u_t \tag{4.7}$$

A problem now arises in determining the 'optimal' number of lags of the dependent variable to add to the estimated equation in order to sufficiently allow for autocorrelation, but not to over-fit. One way to determine the number of lags is to use the frequency of the data as a decision rule (e.g., use four lags for quarterly data, twelve for monthly and so on), but this is likely to lead to considerable over-fitting in the case of monthly data as we have here. Alternatively, the number of lags may also be determined, based on the data, using some kind of information criterion, such as Akaike's or Schwarz's Bayesian criteria. However, given the limited number of data points at our disposal and for consistency and comparability across series, we employ an arbitrary zero and three lags for all currency pairs investigated. We find that the conclusions are not qualitatively affected by this choice of lag length.

4.6 Structural breaks in exchange rates

A structural break occurs when the properties of a series go through a substantial change in behaviour so that previous models that described the relationship between variables subsequently break down. Although the term is often used somewhat loosely in the modern applied econometrics literature, it may be helpful to distinguish between a structural break and an outlier. A structural break is as described in the previous line, where the behaviour of a series or its relationship with other series changes on a long-term or permanent basis, whereas an outlier occurs when the properties of a series change for one or perhaps several periods before fairly quickly reverting back to its previous behaviour. In the context of medieval exchange rates, structural breaks could be caused by wars, financial crises, currency shortages, debasements or recoinages, poor harvests, and so on.

In terms of the econometrics, early tests for structural breaks were conducted in the context of regressions based on stationary time series – for example, the Chow (1960) and Quandt (1960) likelihood ratio tests. However, some more recent approaches have been conducted in the context of tests for unit roots since it has been shown that the standard Dickey–Fuller-type unit root tests presented above do not perform well if there are one or more structural breaks in the series under investigation, either in the intercept or the slope of the regression. More specifically, the tests have low power in such circumstances and they fail to reject the unit root null hypothesis when it is incorrect as the slope parameter in the regression of Δy_t on y_{t-1} is biased towards unity by an unparameterised structural break. In general, the larger the break and the smaller the sample, the lower the power of the test. As Leybourne *et al.* (1998) have shown, in addition unit root tests are oversized in the presence of structural breaks, so they also reject the null hypothesis too frequently when it is correct.

Perron's (1989) work was the first to systematically address the issue of testing for unit roots in the presence of structural breaks. This work is considered important since he was able to demonstrate that if we allow for structural breaks in the testing framework, a whole raft of macroeconomic series that Nelson and Plosser (1982) had identified as non-stationary may turn out to be stationary. Perron argues that most economic time series are best characterised by broken trend stationary processes, where the data generating process is a deterministic trend but with a structural break around 1929 that permanently changed the levels (i.e. the intercepts) of the series.

Perron (1989) proposes three test equations differing dependent on the type of break that was thought to be present. The first he terms a 'crash' model that allows a break in the level (i.e. the intercept) of the series; the second is a 'changing growth' model that allows for a break in the growth rate (i.e. the slope) of the series; the final model allows for both types of break to occur at the same time, changing both the intercept and the slope of the trend. If we define the break point in the data as T_b, and D_t is a dummy variable defined as

$$D_t = \begin{cases} 0 \text{ if } t < T_b \\ 1 \text{ if } t \geq T_b \end{cases} \tag{4.8}$$

The general equation for the most general type of test Perron proposed is

$$\Delta y_t = \mu + \psi y_{t-1} + \beta_1 D_t + \beta_2 (t - T_b) D_t + \lambda t + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + u_t \tag{4.9}$$

For the crash only model, set $\beta_2 = 0$, while for the changing growth only model, set $\beta_1 = 0$. In all three cases, there is a unit root with a structural break at T_b under the null hypothesis and a series that is a stationary process with a break under the alternative.

While Perron (1989) initiated a new literature on testing for unit roots in the presence of structural breaks, an important limitation of his approach is that it assumes the break date is known in advance and the test is constructed using this information. It is possible, and perhaps even likely, however, that the date will not be known and must be determined from the data. More seriously, Christiano (1992) has argued that the critical values employed with the test will presume the break date to be chosen exogenously and yet most researchers will select a break point based on an examination of the data and thus the asymptotic theory assumed will no longer hold.

As a result, Banerjee *et al*. (1992) and Zivot and Andrews (1992) introduce an approach to testing for unit roots in the presence of structural change that allows the break date to be selected endogenously. Their methods are based on recursive, rolling and sequential tests. For the recursive and rolling tests, Banerjee *et al*. propose four specifications. First, the standard Dickey–Fuller test on the whole sample, which they term $\hat{t}_{DF}$; second, the ADF test is conducted repeatedly on the sub-samples and the minimal DF statistic, $\hat{t}_{DF}^{\min}$, is obtained; third, the maximal DF statistic is obtained from the sub-samples, $\hat{t}_{DF}^{\max}$; finally, the difference between the maximal and minimal statistics, $\hat{t}_{DF}^{diff} = \hat{t}_{DF}^{\max} - \hat{t}_{DF}^{\min}$, is taken. For the sequential test, the whole sample is used each time with the following regression being run:

$$\Delta y_t = \mu + \psi y_{t-1} + \alpha \tau_t(t_{used}) + \lambda t + \sum_{i=1}^{p} \alpha_i \Delta y_{t-i} + u_t \tag{4.10}$$

where $t_{used} = T_b / T$. The test is run repeatedly for different values of T_b over as much of the data as possible (a ‘trimmed sample’) that excludes the first few and the last few observations (since it is not possible to reliably detect breaks there). Clearly, it is $\tau_t(t_{used})$ that allows for the break, which can either be in the level (where $\tau_t(t_{used}) = 1$ if $t > t_{used}$ and 0 otherwise); or the break can be in the deterministic trend $\tau_t(t_{used}) = t - t_{used}$ if $t > t_{used}$ and 0 otherwise). For each specification, a different set of critical values is required, and these can be found in Banerjee *et al*.

Much recent work on whether exchange rates contain unit roots has been conducted in the panel context, where the additional information from combining series together can lead to considerable efficiency gains and improvements in power. Relevant research includes Jorion and Sweeney (1996), who reject the unit root null hypothesis for ten US dollar-denominated currencies and seven Deutschmark-denominated currencies. Wu (1996) concludes firmly that US dollar real exchange rates are not unit root processes using a panel of eighteen monthly real exchange rate series; a similar result is observed by MacDonald (1996) and Oh (1996) using longer samples of annual data. It is also possible to allow for structural breaks in the context of a panel unit root process, as conducted by Wu *et al*. (2004), who reject the unit root null hypothesis for a set of South East Asian real exchange rates. However, the nature of our data including considerable differences in the lengths of the series and the sample periods covered mean that it is not possible to use a panel approach.

Perron (1997) proposes an extension of his original technique using a sequential procedure that estimates the test statistic allowing for a break at any point during the sample to be determined by the data. This technique is very similar to that of Zivot and Andrews, except that Perron’s is more flexible, and therefore arguably preferable, since it allows for a break under both the null and alternative hypotheses whereas according to Zivot and Andrews’ model, it can only arise under the alternative. Given its apparent superiority over the previous approaches, we employ the Perron (1997) test here, with the results being presented in Table 4.6.

Table 4.6 Unit root and structural break tests

Series	*DF t-statistic (0 lags)*	*ADF t-statistic (3 lags)*	*Perron break test t-statistic*	*Break date*
Barcelona–Avignon	–4.69***	–4.81***	–0.46	1388:08
Barcelona–Bruges	–2.06	–2.40	–0.82	1389:02
Barcelona–Florence	–2.31	–2.46	–0.45	1398:04
Barcelona–Genoa	–3.95***	–5.01***	0.08	1388:08
Barcelona–Majorca	–4.25***	–5.68***	–1.43	1393:01
Barcelona–Montpellier	–4.81***	–4.85***	–1.29	1389:10
Barcelona–Pisa	–2.78*	–3.11**	–1.20	1393:07
Barcelona–Venice	–1.55	–2.21	–0.86	1399:08
Bruges–Barcelona	–3.10***	–2.39	–0.60	1396:07
Bruges–Genoa	–3.12**	–3.09**	–0.34	1396:07
Bruges–London	–2.33	–2.29	–3.95***	1396.09
Bruges–Paris	–2.23	–1.33	0.62	1396:07
Bruges–Pisa	–0.98	–1.27	–0.83	1395:04
Bruges–Venice	–4.50***	–2.71*	–0.45	1396:07
Venice–Barcelona	–1.87	–1.64	–0.60	1400:03
Venice–Bologna	–6.12***	–3.85***	0.11	1392:11
Venice–Bruges	–1.49	–1.30	–0.82	1399:10
Venice–Florence	–6.05***	–3.98***	0.01	1386:09
Venice–Genoa	–1.18	–0.78	1.04	1386:09
Venice–London	–0.77	–0.48	–3.10**	1406:08
Venice–Lucca	–4.17***	–3.02**	–1.06	1400:04
Venice–Milan	–2.4	–2.12	–0.88	1395:04
Venice–Paris	–1.46	–1.44	–1.25	1400:02
Venice–Pisa	–6.30***	–5.34**	–1.16	1386:03
Venice–Rome	–2.81*	–2.60	–1.17	1399:09

Notes

A constant but no trend is included in all DF and ADF test regressions; the additive outlier model is used for the Perron test.

These are tests for whether the series are 'stationary' in their levels (rather than the percentage changes) – i.e. the question is whether they contain a stochastic trend, otherwise known as a unit root. The null hypothesis is that they do and asterisks denote instances where this hypothesis is rejected. The table presents results for both the pure unit root tests (i.e. the DF and ADF tests) and tests that allow for a single endogenously determined structural break (the Perron test).

As the results in the table show, there is a considerable degree of agreement between the test that includes an arbitrary three lags and the pure Dickey–Fuller (DF) test that does not. The evidence for modern foreign exchange series is more strongly that they are all non-stationary, whereas the results here are much more mixed. Almost half the series here appear to be stationary, including most of the Barcelona series and half of those from Venice.

The Perron test that we employ is for a unit root and a structural break, when the break date is assumed unknown. The null in each case is of a unit root with a break against an alternative of stationarity, and the results are very different to the DF tests. Allowing for a structural break, only two of the series are stationary in their levels, which is arguably much more the result that we would have expected. For the Venice-based series, there seems to be a number having breaks around September 1386 and September 1399–April 1400. Several of the Barcelona series have breaks in August 1388, but here the patterns are not so clear. For rates quoted at Bruges, the most common date is between April and July 1396. This is interesting, in large part because it does not coincide with any of the frequent devaluations and *renforcements* of the Flemish currency (Munro, 2012). De Roover suggested that the rise in the Bruges-Venice rate may have been connected to the capture of John de Nevers, heir to the duchy of Burgundy (of which the county of Flanders was part) by the Ottomans at the battle of Nicopolis. The subsequent payment of John's ransom, set at 200,000 ducats, required large transfers from Bruges and Paris to Venice and thence to the east, which would have driven up the price of ducats (de Roover 1968, pp. 52–53).

An important limitation of the Perron (1997) approach is that it can only be employed to estimate (up to) one structural break. This may be problematic not only in the sense that other breaks cannot be detected even if they are present in the data, but more seriously, if there are multiple breaks then a model assuming that there is at most one will be mis-specified, possibly leading to errors in inference as serious as those if there was one break and we allowed for none. Thus a further extension would be to allow for more than one structural break in the series – for example, Lumsdaine and Papell (1997) enhance the Zivot and Andrews (1992) approach to allow for two structural breaks. It is also possible to allow for structural breaks in the cointegrating relationship between series using an extension of the first step in the Engle–Granger approach – see Gregory and Hansen (1996).

More recently, in a series of papers, Bai and Perron (1998, 2003a, 2003b) develop another technique that tests for structural change, but this time allowing for more than one break. This approach, however, is not conducted in the context of unit root tests and is instead applied to a model using stationary data. In our case, in the absence of data on exogenous factors that may affect the exchange rates, the test for a structural break is based on a simple autoregressive model of order one (AR(1)) on the percentage changes in the exchange rates. Essentially, the test works by estimating a set of models allowing for 1, 2,…, m structural breaks and selecting the number of breaks that minimises the Bayesian information criterion (BIC). It is also possible to use an F-test to examine the null hypothesis that there exists a given number of breaks m against the alternative of $m-1$. The model is thus

$$\begin{aligned} y_t &= x_t'\beta + z_t'\delta_1 + u_t, \quad t = 1,\ldots T_1 \\ y_t &= x_t'\beta + z_t'\delta_2 + u_t, \quad t = T_1 + 1,\ldots T_2 \\ &\vdots \\ y_t &= x_t'\beta + z_t'\delta_{m+1} + u_t, \quad t = T_m + 1,\ldots T \end{aligned} \tag{4.11}$$

where y_t is the dependent variable, x_t is a $p \times 1$ vector of variables whose relationships with y are assumed not to vary over time and z_t is a $q \times 1$ vector of variables whose relationships with y vary because of a set of structural breaks which take place at times T_1, T_2,…, T_m. The vectors β, δ_1,…, δ_{m+1} represent sets of unknown parameter values. We conduct the test allowing for a maximum of five breaks, and a minimum distance between the breaks of three months.

The results from applying the test are given in Table 4.7. For around half of the series, no structural breaks are detected. Interestingly, almost all of the Barcelona series do not have any breaks, while all of the Bruges series except Bruges–Barcelona have a single break, and the results for the Venice pairs are more mixed and with occasional multiple breaks. Thus it appears that, overall, the possibility of multiple breaks need not be of significant concern here. Where breaks do take place, it is evident that there is much less consistency across series in their dates of occurrence than was the case for the unit root-structural break results presented above.

One interesting case study is Venice–Milan. The Perron test reported in Table 4.6 identifies a structural break in April 1395 whereas the Bia–Perron break test reported in Table 4.7 found a structural break in March 1400. This coincides with two different periods of Milanese monetary policy (Mueller, 1997, pp. 590–592). The exchange rate between Venice and Milan was quoted between the gold Venetian ducat and a notional gold Milanese ducat consisting of 32 silver *soldi imperiali*. Until 1395 the domestic exchange rate between ducats and *lire imperiali* in Milan remained stable and the exchange rate between Venice and Milan varied around 3–5 per cent in favour of Venice (i.e. 100 Venetian ducats were equivalent to 103–105 Milanese ducats/3296–3360 *soldi imperiali*). After 1395, however, Giangealeazzo Visconti, ruler of Milan, progressively debased the silver coinage to fund his aggressive foreign policy. As the value of the silver coin began to fall, so the Venice–Milan exchange rate rose. At its nadir in early 1400, the Milanese ducat was equivalent to 48–49 *soldi imperiali*, a decrease of roughly half in the value of the silver currency. In parallel, the Venice–Milan exchange rate increased

Table 4.7 Bai–Perron structural break tests

Series	*Number of structural breaks*	*Break date*	*Second break date*	*Third break date*
Barcelona–Avignon	0			
Barcelona–Bruges	0			
Barcelona–Florence	0			
Barcelona–Genoa	0			
Barcelona–Majorca	0			
Barcelona–Montpellier	1	1395:06		
Barcelona–Pisa	2	1399:02	1398:12	
Barcelona–Venice	0			
Bruges–Barcelona	0			
Bruges–Genoa	1	1396:07		
Bruges–London	1	1407:12		
Bruges–Paris	1	1395:11		
Bruges–Pisa	1	1396:07		
Bruges–Venice	1	1410:03		
Venice–Barcelona	0			
Venice–Bruges	0			
Venice–Bologna	1	1403:07	1400:03	
Venice–Florence	0			
Venice–Genoa	0			
Venice–London	1	1406:11		
Venice–Lucca	0	1404:07	1409:12	
Venice–Milan	2	1400:03	1399:09	
Venice–Paris	1	1409:03		
Venice–Pisa	3	1397:10	1397:06	1398:01
Venice–Rome	1	1403:10		

Note
The number of structural breaks is selected by the Bai–Perron test that minimises the BIC; the tests are constructed in the context of an AR(1) model.

to 52 per cent in favour of Venice (so that 100 Venetian ducats were now worth 152 Milanese ducats/4864 *soldi imperiali*). At this point, Visconti sought to revalue the silver currency in Milan and return to the 1395 valuation of 32 *soldi imperiali* to the Milanese ducat. On 21 February 1400, he issued a decree 'crying up' the value of the silver *lira imperiali* by one-third. By March 1400, the domestic exchange rate in Milan had fallen to 35–36 *soldi imperiali* per Milanese ducat and the Venice-Milan exchange rate had dropped correspondingly to 10–12 per cent in favour of Venice (i.e. 100 Venetian ducats were worth 110–112 Milanese ducats/3520–3584 *soldi imperiali*).

4.7 Conclusion

This chapter has introduced some fundamental statistical techniques that historians could use when analysing time series data. It has set out methods to identify seasonal trends, and to test for unit roots and structural breaks. Econometric analysis offers a number of advantages; it is more precise and, in some ways, more rigorous than merely 'eyeballing' the data. It can identify trends that a more superficial study of the raw numbers might miss or mis-interpret. It can determine whether, from a statistical perspective, an observed pattern is important or not. Moreover, a more formal quantitative analysis can also reduce the number of 'false positives' where the researcher effectively sees a face in the clouds that is not really there. However, such statistical methods are not a replacement for history but need to be used in conjunction with traditional historical research. First, it is vital to consider the historical context when explaining any results produced by the statistical analysis. For example, structural break tests may propose potential dates but the ultimate meaning and significance of any such changes identified depends on the historical reconstruction. Second, and perhaps even more importantly, any statistical analysis is only as good as the data on which it is based. Historians must ensure the integrity of the data and that it is suitable for the analysis proposed.

Notes

1 The data from the *Handbook of Medieval Exchange* has been uploaded to the Medieval and Early Modern Data Bank and can be accessed at www2.scc.rutgers.edu/memdb/search_form_spuf.php.

2 The homepage of the Datini archive is http://datini.archiviodistato.prato.it/www/indice.html.

3 The exchange rates for Barcelona and Bruges were entered into a spreadsheet from the appendix in de Roover (1968). Mueller's data for Venice is available from the Medieval and Early Modern Data Bank (www2.scc.rutgers.edu/memdb/search_form_mueller.php). It should be noted that there are a number of errors in the online data set, especially for the Venice–Barcelona exchange rates. These do not seem to be included in Mueller's own figures and presumably occurred during when the data was being re-entered for upload. We identified data points that deviated significantly from contemporaneous rates, checked them against the original letters from the Datini archive and made corrections where necessary.

References

Bai, Jushan and Perron, Pierre (1998) Estimating and testing linear models with multiple structural breaks, *Econometrica*, 66, 47–78.

Bai, Jushan and Perron, Pierre (2003a) Computation and analysis of multiple structural change models, *Journal of Applied Econometrics*, 18, 1–22.

Bai, Jushan and Perron, Pierre (2003b) Critical values for multiple structural change tests, *Econometrics Journal*, 1, 1–7.

Banerjee, Anindya, Lumsdaine, Rubin L. and Stock, James H. (1992) Recursive and sequential tests of the unit-root and trend-break hypotheses: Theory and international evidence, *Journal of Business and Economic Statistics*, 10, 271–287.

Banerjee, Anindya, Dolado, Juan J., Galbraith, John W. and Hendry, David F. (1993) *Cointegration, Error Correction, and the Econometric Analysis of Nonstationary Data*, Oxford: Oxford University Press.

Bell, Adrian R., Brooks, Chris and Moore, Tony K. (2009) Interest in medieval accounts: Examples from England, 1272–1340, *History*, 94, 411–433.

Bigwood, Georges and Grunzweig, Armand (eds) (1962) *Les Livres des Comptes des Gallerani*, Brussels: Académie Royale de Belgique.

Blomquist, Thomas W. (1990) Some observations on early foreign exchange banking based upon new evidence from XIIIth century Lucca, *Journal of European Economic History*, 19, 353–375.

Booth, G. Geoffrey (2011) Foreign exchange profits in two early Renaissance money markets, *Journal of European Economic History*, 38, 123–144.

Booth, G. Geoffrey and Gurun, Umit G. (2008) Volatility clustering and the bid-ask spread: Exchange rate behavior in early Renaissance Florence, *Journal of Empirical Finance*, 15, 133–144.

Chilosi, David and Volckart, Oliver (2011) Money, states and empire: Financial integration and institutional change in Central Europe, 1400–1520, *Journal of Economic History*, 71, 762–791.

Chow, Gregory C. (1960) Tests of equality between sets of coefficients in two linear regressions, *Econometrica*, 28(3), 591–605.

Christiano, Lawrence. J. (1992) Searching for a break in GNP, *Journal of Business and Economic Statistics*, 10, 237–250.

Corbae, Dean and Ouliaris, Sam (1988) Cointegration tests of purchasing power parity, *Review of Economics and Statistics*, 70, 508–511.

Denzel, Markus A. (2010) *Handbook of World Exchange Rates, 1590–1914*, Aldershot: Ashgate.

Dickey, David A. and Fuller, Wayne A. (1979) Distribution of estimators for time series regressions with a unit root, *Journal of the American Statistical Association*, 74, 427–431.

Einzig, Paul (1970) *The History of Foreign Exchange*, 2nd edn, London: Macmillan.

Fuller, W. A. (1976) *Introduction to Statistical Time Series*, New York: Wiley.

Goodhart, Charles A. E., McMahon, Patrick C. and Ngama, Yerima L. (1993) Testing for unit roots with very high frequency spot exchange rate data, *Journal of Macroeconomics*, 15(3), 423–438.

Gregory, Allan W. and Hansen, Bruce E. (1996) Residual-based tests for cointegration in models with regime shifts, *Journal of Econometrics*, 70, 99–126.

Jorion, Philippe and Sweeney, Richard J. (1996) Mean reversion in real exchange rates: Evidence and implications for forecasting, *Journal of International Money and Finance*, 15, 535–550.

Koyama, Mark (2010) Evading the 'Taint of Usury': The usury prohibition as a barrier to entry, *Explorations in Economic History*, 47, 420–442.

Kugler, Peter (2011) Financial market integration in late medieval Europe: Results from a threshold error correction model for the Rhinegulden and Basle Pound, 1365–1429, *Swiss Journal of Economics and Statistics*, 147, 337–357.

Leone, Alfonso (1983) Some preliminary remarks on the study of foreign currency exchange in the medieval period, *Journal of European Economic History*, 12, 619–629.

Leybourne, Stephen, Newbold, Paul and Vougas, Dimitrios (1998) Unit roots and smooth transitions, *Journal of Time Series Analysis*, 19, 83–97.

Lumsdaine, Robin L. and Papell, David H. (1997) Multiple trend breaks and the unit root hypothesis, *Review of Economics and Statistics*, 79(2), 212–218.

MacDonald, Ronald (1996) Panel unit root tests and real exchange rates, *Economics Letters*, 50, 7–11.

McCusker, John J. and Gravesteijn, Cora (1991) *The Beginnings of Commercial and Financial Journalism: The Commodity Price Currents, Exchange Rate Currents, and Money Currents of Early Modern Europe*, Amsterdam: Nederlandsch Economisch-Historisch Archief.

Meese, Richard A. and Singleton, Kenneth J. (1982) On unit roots and the empirical modelling of exchange rates, *Journal of Finance*, 37(4), 1029–1035.

Melis, Federigo (1962) *Aspetti della Vita Economica Medievale: Studi nell'Archivio Datini di Prato*, Siena: Monte dei Paschi di Siena.

Mueller, Reinhold C. (1997) *The Venetian Money Market: Banks, Panics, and the Public Debt, 1200–1500*, Baltimore, MD: Johns Hopkins Press.

Munro, John H. A. (1979) Bullionism and the bill of exchange in England, 1272–1663 in *The Dawn of Modern Banking: Ten Essays Prepared for a Conference Sponsored by the Centre for Medieval and Renaissance Studies, University of California, Los Angeles*, London: Yale University Press, pp. 169–239.

Munro, John H. A. (2012) Coinage debasements in Burgundian Flanders, 1384–1482: Monetary of fiscal remedies? in Nicholas, David, Bachrach, Bernard S. and Murray, James M. (eds) *Comparative Perspectives on History and Historians: Essays in Honour of Bryce Lyon*, Kalamazoo: Medieval Institute Publications, Western Michigan University, pp. 314–360.

Nelson, Charles and Plosser, Charles (1982) Trends and random walks in macroeconomic time series, *Journal of Monetary Economics*, 10, 139–162.

Nigro, Giampiero (ed.) (2010) *Francesco di Marco Datini: The Man, the Merchant*, Florence: Firenze University Press, Fondazione Istituto Internazionale di Storia Economica 'F. Datini'.

Oh, Keun-Yeob (1996) Purchasing power parity and unit root tests using panel data, *Journal of International Money and Finance*, 15, 405–418.

Origo, Iris M. (1963) *The Merchant of Prato: Francesco di Marco Datini, 1335–1410* rev. edn, Harmondsworth: Penguin.

Pagnini, Gianfrancesco (ed.) (1766) *Della Decima e di Varie Altre Gravezze Imposte dal Comune di Firenze della Monete e della Mercatura de Fiorentini Fino al Secolo XVI: Tomo Quarto Continente La Pratica della Mercatura Scritta da Giovanni di Antonio da Uzzano nel 1442*, Lisbon-Lucca.

Perron. Pierre (1989) The Great Crash, the Oil Price Shock and the Unit Root Hypothesis, *Econometrica*, 57, 1361–1401.

Perron, Pierre (1997) Further evidence on breaking trend functions in macroeconomic variables, *Journal of Econometrics*, 80, 355–385.

Quandt, Richard (1960) Tests of the hypothesis that a linear regression system obeys two different regimes, *Journal of the American Statistical Association*, 55, 324–330.

de Roover, Raymond A. (1944a) Early accounting problems of foreign exchange, *The Accounting Review*, 19, 381–407.

de Roover, Raymond A. (1944b) What is dry exchange? A contribution to the study of English Mercantilism, *Journal of Political Economy*, 52, 250–266.

de Roover, Raymond A. (1968) *The Bruges Money Market Around 1400 (with a Statistical Supplement by Hyman Sardy)*, Brussels: Paleis der Academiën.

Rubin, Jared (2010) Bills of exchange, interest bans and impersonal exchange in Islam and Christianity, *Explorations in Economic History*, 47, 213–227.

Schotman, Peter and Dijk, Herman K. van (1991) A Bayesian analysis of the unit root in real exchange rates, *Journal of Econometrics*, 49, 195–238.

Spufford, Peter (1986) *Handbook of Medieval Exchange*, London: Royal Historical Society.

Spufford, Peter (1991) *Money and its Use in Medieval Europe*, Cambridge: Cambridge University Press.

Taylor, Mark P. (1990) On unit roots and real exchange rates: Empirical evidence and Monte Carlo analysis, *Applied Economics*, 22, 1311–1321.

Volckart, Oliver and Wolf, Nikolaus (2006) Estimating financial integration in the Middle Ages: What can we learn from a TAR Model? *Journal of Economic History*, 66, 122–139.

Whitt, Joseph A. Jr. (1992) The long-run behaviour of the real exchange rate: A reconsideration, *Journal of Money, Credit and Banking*, 24(1), 72–82.

Wu, Yangru (1996) Are real exchange rates nonstationary? Evidence from a panel data test, *Journal of Money, Credit and Banking*, 28, 54–63.

Wu, Jyh-Lin, Tsai, Li-Ju and Chen, Show-Lin (2004) Are real exchange rates nonstationary? The Pacific Basin perspective, *Journal of Asian Economics*, 15, 425–438.

Zivot, Eric and Andrews, Donald W. K. (1992) Further evidence on the Great Crash, the Oil Price Shock, and the Unit Root Hypothesis, *Journal of Business and Economic Statistics*, 10, 251–270.

5 Local property values in fourteenth- and fifteenth-century England[1]

Margaret Yates, Anna Campbell and Mark Casson

5.1 Introduction: changing economic conditions

The changing value of property forms a central pre-occupation of economic historians and economists today. That it changed is indisputable, and so too are the long-term social and economic events which influenced the value of property. The period chosen for this study includes momentous events, especially the arrival of the Black Death in England in 1348–9 with subsequent demographic collapse and a change in the ratio of population to landholdings. A hundred years later a more capitalistic attitude towards landholding has been observed and a loosening of family ties to land. Further, we are becoming increasingly aware of the impact of market activity more generally in the medieval period, the international dimension to trade and the role of coin in the economy, all of which may have had an impact on local property values.

Previous attempts at calculating the changing value of land have been thwarted by the fragmentary nature of the historical record, even in the early modern and modern periods (Turner *et al.*, 1997). For the medieval period heroic efforts were made by historians such as Bean to calculate the relationship between purchase price recorded in isolated references across multiple estate records, and the annual value of the property. For example he calculated that the purchase price in the thirteenth century represented ten years' annual value and twenty years' by the mid-fifteenth century, but the rate varied considerably in the interim (Bean, 1991, p. 567). In a highly sophisticated analysis Bruce Campbell calculated unit values for different types of agricultural land 1270–1350 as the anticipated net profit or 'rent' that they were capable of yielding as opposed to the total capital value of the land (Campbell, 2000, pp. 345–364; Campbell and Bartley, 2006, pp. 165–195). Through a study of Inquisitions *Post Mortem* (IPMs) he was able to demonstrate a gradual decline across England in monetary value over the fourteenth century in arable, meadow and pasture land (Campbell and Bartley, 2006, p. 167). It is generally accepted, although recently contested, that IPMs become a less valuable source of evidence after the mid fourteenth century, hence the need for studies such as this one (Yates, 2012). Further, we need to be alert to the meaning of 'value' and distinguish the relationship between the purchase price, capital value and net profit or rent that could be expected from the property. We are dealing with capital value in this chapter.

The value given to a property at any point in time is influenced by a number of underlying factors, which shape the research questions in this chapter. They can be summarised as:

- *Population change* and its relationship with total landholding, and in particular the devastating mortality of the Black Death.
- Allied to this is the *amount of land* coming on to the market and available for purchase.
- The effects of *political upheavals*: the depositions of Edward II in 1327, Richard II in 1399 and Henry VI in 1461; the readeption of Henry VI in 1470 and his subsequent capture in 1471 by Edward IV whose 'reign' had been interrupted when he fled the country in 1470. Allied to these events are two periods of civil war.
- *Changes in the social distribution of landholding*, especially the build-up of small estates through the activities of the 'nouveau riche' (professional groups such as lawyers, or crown or government employees, along with the successful merchants), possibly at the expense of the small independent freeholder. These appear to be connected with the acquisition of 'lordship' as a way of being accepted into county society.
- *Agricultural changes*, involving a shift in emphasis from arable to pasture and rise of private enclosed land associated with more capitalistic attitudes towards land.
- *Changing urban fortunes*, as some towns grew while others declined, altering property values in different areas, both within counties and between them.
- *Soil and topography* as exogenous factors.

To elucidate the role of these factors, this chapter focuses on:

1. local property values, with a particular emphasis on the value of manors and the values of different types of agricultural land;
2. the composition of landholdings, as reflected in the uses of land, in particular as arable, meadow and pasture;
3. regional differences, both between counties and between urban and rural areas within the same county;
4. changes over time, especially after the Black Death; and
5. the social and occupational status and the gender of landowners.

5.2 The sources and their context

Collaborative work, such as that by Broadberry, Campbell *et al*. (2012) is capable of yielding results that the lone researcher in the dusty archive cannot contemplate, despite the advances in technology. The results presented in this chapter are the outcome of such collaborative work, involving medieval historians and economists. A previously neglected source for quantitative data, feet of fines (or fines for short), are mined for their data on predominantly lay-owned freehold property and its value.

A fine – final agreement or concord – records the conveyance of free land and is the bottom section (the foot) of a tripartite document comprising identical copies and drawn up, in Latin, by the royal clerks of the court. The fine records the terms of an agreement regarding the *inter vivos* transfer of freehold land, specifying the names of the two parties, a description of the property being conveyed, details of the different types of estates in land, and a sum of money or consideration. The language is legal and often archaic. The agreement is between a plaintiff, who is usually the recipient of the property, either as a gift or sale, while the defendant is the previous owner. When interpreting the data it is often useful to consider the plaintiff as a purchaser and the defendant as a seller.

Why hasn't this source of data been exploited before? Probably because of the fictitious nature of the suit at law that lay behind their creation. Nevertheless, the descriptions of the properties contained in these documents, while not an accurate survey in the modern sense, do describe in a recognisable manner the property being conveyed. Moreover, and it is our argument, that, while it is not the sale price, the sum of money recorded in the fine (the consideration) reflects the perceived capital value of the property at that time.

The strength of this documentary series lies in its size. Fines survive for the whole of England 1195–1834, there are literally thousands of them and many are published as abstracts in English in county record society volumes. They are, however, not without problems as their wording is formulaic and changes in the law of real property lie concealed behind the fine. That said, they remain a remarkable and easily accessible source of time series data on the composition of landholdings and their value, especially for the medieval period.

The number of fines that survive over the period 1300–1500 was investigated for four English counties: Essex, Warwickshire, Berkshire and Huntingdonshire. Essex is a large county east of London and north of the River Thames. Its county town is the old Roman fort of Colchester. The county is characterised by the heavy clay soils typical of the London area. Warwickshire is more remote from London; its county town of Warwick and commercial centre of Coventry are about one hundred miles from London, and its soil is mixed, with significant amounts of limestone. The county is centrally situated near the intersection of two Roman roads – Watling Street and the Fosse way. Berkshire is a linear county west of London along the south bank of the Thames, and includes the proto-industrial towns of Reading and Newbury, which were prominent in the cloth industry that expanded in the late fifteenth and early sixteenth century. Huntingdonshire is a small county whose county town is about sixty miles from London; like Essex, it is relatively flat, and has clay soils; it was less favoured than Essex or Berkshire as a place for wealthy Londoners to acquire country estates.

Figure 5.1 shows that the four counties reveal a broadly similar pattern of change over time, with a gradual decline in the number of fines from the end of

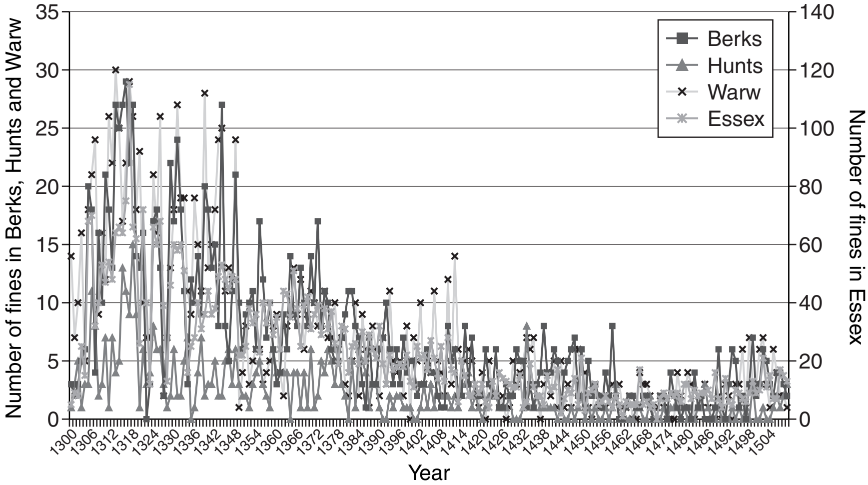

Figure 5.1 Variation of number of recorded fines by date.

the fourteenth century. Fines are particularly numerous for the period 1300–39. Because Essex is so much larger than the other counties the numbers of fines are shown on a different scale.

The main reason usually suggested for this general decline is that fines became less attractive as a legal instrument for transferring property, and in particular title to land, although they remained useful for confirming married women's title to property.

Table 5.1 reports the estimates of a VAR (vector autoregression) which analyses the number of fines in each county in each year. The VAR is a set of four regressions – one for each county – in which the number of fines in each county in a given year is related to the number of fines in the two previous years in each of the counties – the county itself and the other three counties as well (for further details of VAR regressions see Chapter 2). The regressions also contain a linear time trend, a dummy variable for the Black Death (zero for years before the Black Death and one for years afterwards) and a dummy designed to capture the effects of legal changes that occurred around 1360 (zero for years before 1360 and one for years afterwards). Each of the four regressions is estimated by ordinary least squares (OLS). To interpret the regressions results, the focus is placed on the variables that have statistically significant coefficients (as indicated by one or more asterisks in the table – the more asterisks, the greater the significance).

The results show that

- The number of fines in Essex tends to track the number of fines in Berkshire with a lag of one year;
- Huntingdonshire tracks Essex (strongly) and Berkshire (weakly);

Table 5.1 VAR regressions of the number of fines by year in four counties, 1300–1509

Explanatory variable	*Essex*	*Berkshire*	*Huntingdonshire*	*Warwickshire*
Constant	26.005***	8.554***	1.131	11.303***
	(0.000)	(0.000)	(0.143)	(0.000)
Essex lag one year	0.124	–0.002	0.042***	0.042
	(0.137)	(0.948)	(0.007)	(0.214)
Essex lag two years	0.119	0.042	0.012	–0.008
	(0.133)	(0.217)	(0.420)	(0.814)
Berkshire lag one year	1.156***	0.290***	0.073**	0.270***
	(0.000)	(0.001)	(0.050)	(0.001)
Berkshire lag two years	–0.199	0.065	0.017	0.102
	(0.340)	(0.463)	(0.666)	(0.234)
Huntingdonshire lag one year	0.605	–0.135	0.095	0.265*
	(0.121)	(0.419)	(0.190)	(0.097)
Huntingdonshire lag two years	0.435	0.125	–0.050	0.028
	(0.247)	(0.440)	(0.476)	(0.854)
Warwickshire lag one year	–0.066	0.105	0.035	–0.076
	(0.739)	(0.218)	(0.347)	(0.346)
Warwickshire lag two years	–0.305	–0.127	–0.042	–0.046
	(0.116)	(0.126)	(0.247)	(0.562)
Time	–0.081***	–0.025***	–0.006	–0.018**
	(0.000)	(0.009)	(0.116)	(0.043)
Black Death	–7.976**	–2.244	–0.389	–8.817***
	(0.033)	(0.161)	(0.575)	(0.000)
Legal change 1360	0.316	–0.841	0.388	1.953
	(0.922)	(0.545)	(0.520)	(0.142)
R^2	0.753	0.600	0.529	0.703
Adjusted R^2	0.740	0.581	0.503	0.686
F-statistic	54.484	27.116	20.038	42.160
	(0.000)	(0.000)	(0.000)	(0.000)
Durbin–Watson	2.079	2.039	1.963	2.119
Number of observations	208	208	208	208

- Warwickshire tracks Berkshire (strongly) and Huntingdonshire (weakly);
- The number of fines trends downwards over the period in all counties, although the trend is not significant in Huntingdonshire;
- The Black Death reduces the number of fines, as would be expected given the reduction in population, but not significantly in either Berkshire or Huntingdonshire.
- Legal changes around 1360 are of no significance.
- Berkshire is the only county in which the number of fines each year depends significantly on the number of fines in the same county in previous years.

The overall picture is one in which the number of fines in each county oscillates randomly about a declining determining time trend. Deviations about the trend are correlated across counties, with Berkshire tending to lead, Warwickshire to follow, with Essex and Huntingdonshire between them. The most significant transmission effects seem to be between neighbouring counties such as Berkshire and Essex (separated only by Middlesex, containing London of course), Essex and Huntingdonshire (separated only by Cambridgeshire) and Berkshire and Warwickshire (separated only by Oxfordshire). By contrast links between distant counties, such as Essex and Warwickshire, are relatively weak. It appears that shocks to the number of fines originate close to London, in Berkshire in particular, and diffuse northwards to Huntingdonshire and Warwickshire.

These results should be interpreted with caution, however. Huntingdonshire is a smaller county than the other three, and has correspondingly fewer fines; this may partly explain why results are less significant for Huntingdonshire than for other counties.

5.3 Detailed analysis of Essex and Warwickshire: methodology

The data examined in detail in this chapter relates to fines recorded in the counties of Essex and Warwickshire over the period 1300–1500. The data is subjected to sophisticated statistical analysis that can unlock patterns in the data in a previously inconceivable manner. The study builds on the author's previous work on Berkshire (Yates, 2013).

The selection of these two counties for investigation was shaped by the existence of scholarly published transcripts and translations, and by the availability of secondary literature which could assist interpretation of local issues. An adequate number of recorded fines and the existence of a continuous record over the fourteenth and fifteenth centuries were also important. Regional diversity was considered too. Warwickshire was chosen as an example from the West Midlands region; its customary land market has been studied by Dyer (1980), who also investigated agricultural change in the county. Warwickshire fines have previously been studied by Dyer (1981) and by Watkins (1997) for the Forest of Arden, while Davies and Kissock (2004) have studied Warwickshire together with other counties. Essex is known for its high level of commercialisation, a preponderance of small landholdings (Britnell, 1993) and its vigorous land market (Poos, 1991, pp. 13–16 and Table 1.1).

Over the period 1300–1500 there were 3354 fines for Essex and 1022 fines for Warwickshire, giving a total of 4376 fines for the study as a whole. The results are reported below for the two counties separately; once other counties have been examined the data can be pooled to test for heterogeneity between counties.

The way in which the data was extracted from the fines, the design of the separate fields and system of coding were all influenced by the Berkshire pilot. For statistical analysis all the data was coded into an Excel spreadsheet, with fines listed in rows and the various types of information extracted from them organised by column. The columns comprise:

- sum or money (the 'consideration') expressed in pence (£1 = 240d and 1 mark = 160d)
- date, expressed in terms of the year;
- names of the parties: the plaintiff(s) and the defendants;
- codes for gender and marital status; these codes are important because making a final concord was an important way of ensuring the transfer of good title to a married woman's property;
- codes for occupation; in particular, clergy, attorneys and tradesmen;
- geographical location, which necessitated identification of the modern place-name;
- urban or rural distinction;
- details of the constituent parts of the property as listed in the fine.

The main constituents of a property are

- Agricultural land:
    - Arable
    - Meadow
    - Pasture
    - Heath
    - Moor
- Other land and buildings:
    - Messuages (housing plots including curtilages and possibly a small yard or garden; also shops or workshops in which the owner resides)
    - Crofts (typically a cottage or small dwelling with enough land to support small-scale self-sufficient food production)
    - Gardens and orchards
- Mills:
    - Windmills
    - Water mills
    - Fulling mills (a water mill for cleaning and processing woollen cloth)

- Additional assets:
  - Manor (the landed estate with associated rights of lordship including various judicial and titular privileges)
  - Advowson (the right of nominating clergy for appointment to an ecclesiastical benefice, such as a living for a local parish priest)
  - Rent
  - Other entitlements (e.g. annual payments in kind)

Wherever possible each constituent is quantified and the resulting measurement entered in an appropriate column. Land is measured in acres, using the convention that one carucate = four virgates = eight bovates = 120 acres, so that one bovate (the acreage that can be tilled using a single-ox plough) corresponds to fifteen acres. The numbers of messuages, crofts and mills are usually recorded, and are therefore included in the analysis, but the acreage of heath and moorland is not, and so these were excluded. Independent evidence suggests that these were of limited value. The money value of rent is not always specified and so rental income is recorded as a binary dummy variable that indicates whether rent was included in the agreement or not. The numbers of manors and advowsons are usually recorded, and so they are included, but other entitlements are very varied and so they are excluded.

5.4 Changes in the use of agricultural land

The database was analysed in two stages. First, it was used to construct a profile of changing land use in the two counties, with special reference to arable, meadow and pasture. This confirmed the existence of significance structural change in agriculture over the period. The results are shown in Tables 5.2 (Essex) and 5.3 (Warwickshire).

The tables are constructed by taking the information on land use from the database and applying the conventions set out above. The analysis is conducted on a decadal basis. The left-hand columns refer to the average acreage of a given type recorded in fines for that decade; if no land of a certain type is mentioned in the fine then it is assumed to be zero. The right-hand columns refer to the average size of holdings, and refer only to holdings of positive size; thus zeros are not included in the calculations of the averages.

There are substantial differences between the two counties. In Essex the percentage of total land in arable use fell from 92 per cent in 1300 to under 64 per cent in 1500. The percentage of land used as meadow rose from 4 per cent in 1300 to over 10 per cent in 1500, while the percentage used as pasture rose from under 4 per cent to over 25 per cent in 1500. Changes gain momentum about 1440 and really accelerate in about 1480.

The average size of pastoral holdings increases temporarily in 1340 and permanently from 1380 onwards, increasing from seven acres in 1300 to seventeen acres in 1380 and fifty-six acres in 1500. The average size of arable holdings also increases, but much more modestly, from about sixty acres in 1300 to over

Table 5.2 Distribution of land between arable, meadow and pasture, together with the average size of holding of each type of land, by decade: Essex, 1300–1500

Decade	*Percentage of total land*			*Average size of holding (acres)*		
	Arable	*Meadow*	*Pasture*	*Arable*	*Meadow*	*Pasture*
1300–9	92.28	3.94	3.78	59.37	5.47	7.29
1310–9	90.73	4.48	4.68	37.17	4.82	7.85
1320–9	91.81	4.76	3.42	46.93	5.32	7.42
1330–9	92.77	3.51	3.72	52.62	4.44	7.55
1340–9	88.78	4.83	6.39	45.69	5.52	11.84
1350–9	88.35	4.63	7.02	51.84	5.52	14.49
1360–9	90.29	5.23	4.48	38.98	5.30	7.82
1370–9	92.94	3.42	3.65	57.92	4.81	9.58
1380–9	88.66	4.16	7.18	61.89	6.64	17.64
1390–9	89.29	4.42	6.29	63.64	6.96	20.21
1400–9	84.80	3.84	1.36	83.80	7.84	32.42
1410–9	86.15	6.62	7.24	67.91	9.47	23.08
1420–9	87.25	4.23	8.53	145.57	10.96	43.27
1430–9	89.21	5.16	5.64	65.21	5.29	22.67
1440–9	82.02	4.94	13.23	86.51	8.04	39.95
1450–9	81.83	5.38	12.60	98.58	8.39	80.78
1460–9	75.86	6.18	17.97	64.56	8.16	52.92
1470–9	83.53	6.28	10.19	68.53	9.21	21.99
1480–9	68.04	13.142	18.82	72.32	26.13	72.50
1490–1500	63.81	10.483	25.71	82.97	19.58	55.86
1300–1500	86.17	5.20	8.63	60.15	7.51	21.30

Sources: Authors' database derived from Fowler (1929–49); Kirk (1913–28); Reaney (1964).

Note
The average sizes of holdings is calculated from holdings with non-zero acreages of the relevant type; they are not averages across all estates that appear in the fines.

eighty acres in 1500. Meadow holding increase more markedly, though less so than pastoral holdings: from about five acres in 1300 to about twenty acres in 1500.

In Warwickshire the changes in land use are more dramatic than in Essex. The proportion of arable falls from 99 per cent in 1300 to 46 per cent in 1500, and most of this fall is accounted for by a substantial increase in pasture. Meadows also increase but more slowly than pasture.

All types of plot size increase. The biggest jumps are around 1400–19 and 1490–9. It is possible, though, that fewer of the relatively small estates would appear in the records in later periods. Average size of agricultural holding is 153.08 acres and significantly larger than Essex.

The results are a clear demonstration of agricultural change and of increase in the size of holdings. Moreover, the evidence is consistent with a process involving the early enclosure and consolidation of estates, and a growing emphasis on commercial agriculture. There are regional differences in the switch from arable to pasture and in the average size of holding.

Table 5.3 Distribution of land between arable, meadow and pasture, together with the average size of holding of each type of land, by decade: Warwickshire, 1300–1500

Decade	*Percentage of total land*			*Average size of holding (acres)*		
	Arable	*Meadow*	*Pasture*	*Arable*	*Meadow*	*Pasture*
1300–9	99.04	0.00	0.96	54.71	–	4.75
1310–19	97.62	2.40	0.00	44.47	4.40	–
1320–9	72.29	3.32	0.36	59.48	5.53	6.0
1330–9	91.43	5.49	3.08	72.81	5.19	15.20
1340–9	95.05	3.76	1.18	72.06	7.21	10.75
1350–9	88.59	6.29	5.16	68.73	8.56	40.00
1360–9	92.17	6.29	1.56	71.20	7.93	17.67
1370–9	93.29	5.38	1.33	63.35	5.29	12.67
1380–9	94.90	4.84	0.26	77.03	6.11	3.0
1390–9	95.36	4.28	0.36	61.65	6.00	6.0
1400–9	96.16	2.95	0.89	132.77	7.57	16.0
1410–19	84.09	7.89	8.02	122.11	14.59	65.20
1420–9	79.45	8.94	11.61	100.67	17.00	36.78
1430–9	72.09	10.94	16.97	101.35	12.15	51.14
1440–9	87.08	5.41	7.51	114.62	10.08	42.00
1450–9	70.53	12.95	16.52	126.67	18.50	35.60
1460–9	73.68	5.26	21.05	77.78	8.33	200.0
1470–9	67.35	16.86	15.79	75.50	15.75	14.75
1480–9	77.92	7.61	14.47	76.75	7.5	16.29
1490–1500	46.20	10.02	43.78	176.76	40.26	234.47
1300–1500	82.94	6.20	10.86	77.48	10.73	64.87

Sources: authors' database derived from Drucker (1943); Stokes and Drucker (1939).

Note
The average sizes of holdings are calculated from holdings with non-zero acreages of the relevant type; they are not averages across all estates that appear in the fines.

These results are broadly consistent with the generally accepted view that links agricultural change to demographic change and the need to feed the population. This view prioritises agriculture as a source of income and sees, particularly in the fifteenth century, a move to more capitalistic attitudes towards agricultural land. The 'story' runs that when population is high the demand for grain is also high as it is a staple in the diet for bread and ale, and so land is put to labour-intensive arable cultivation. When demand is reduced, particularly after the devastating mortality of the Black Death, there is a move towards pastoral husbandry in response to both an increased demand for meat as standards of living rose, and a desire by farmers to switch out of labour-intensive cultivation as wages increased due to labour scarcity. At the same time the demand for wool from both home and export markets was expanding. The lack of pressure on landholdings allowed for engrossment and investment in enclosure and an increase in the size of many properties, many of which were farmed for profit in order to produce a surplus for sale.

5.5 Changes in the value of property

The second stage involved analysing changes in the value of property. This is difficult, because, unlike IPMs, fines do not value individual items separately. The value of individual items must be inferred from the value of the whole. It is possible, in principle, to 'reverse out' from a collection of fines the values that would need to be imputed to individual items if the values of these different items were to add up to the value of the fine.

This approach assumes that the same item has the same value in comparable fines. It is not necessary, however, to assume that items have the same value in all locations, or at all times, or even that they have the same value when the parties to the agreement differ in occupation or gender. All that is necessary is to assume that the same item in the same location transferred between similar parties would have the same value when it appears in more than one fine. This is the approach that is adopted here.

The results of using this approach can be calibrated against evidence from other sources, notably IPMs, and the degree of concurrence can be assessed. It turns out that discrepancies are relatively few and that agreement is high. This may be taken as an indirect validation of the exercise. Furthermore, the consistency with which the considerations of fines appear to have been calculated suggests that, despite their archaic language, their occasional ambiguities and the fictional nature of some of the transactions, the value of the consideration quoted in a fine, can be taken, with suitable qualifications, as an indicator of the economic value, and specifically the capital value, of a portfolio of assets. Because of the huge number of fines available, this method of calculating values for individual types of property can be used to unlock a vast store on information on medieval property values that has never been exploited systematically for this purpose before.

To implement this approach, the considerations relating to each county were examined statistically on the assumption that they reflected the capital values of the assets involved in the notional transactions. The total consideration (measured in pence) in each fine was resolved into a series of components, such that when these components were added together they gave the total value of the fine. Each component reflected the value of a particular type of asset. Because it is impossible to choose imputations of asset values that exactly account for the value of each fine, allowance was made for unobserved omitted factors which may create random noise in the value of a fine. This approach was implemented using regression analysis.

The classification of assets was described above. For statistical purposes it is useful to focus on types of assets that appear in a number of different fines. It is also useful to know the quantity of each type, e.g. the acreage of land, the number of mills, the number of messuages, and so on. The assets included in this exercise are those on which good information is available. Other assets, such as moor and heath, for which acreage is not usually available, are excluded. Sometimes a fine refers to a share in some asset, and when this occurs the share is recorded as a fractional amount of the relevant asset, e.g. a half share in a mill is recorded as 0.5 units.

As indicated above, a notional value is associated with each type of asset; e.g. ten acres of arable land notionally priced at thirty pence per acre have a value of $10 \times 30 = 300$ pence. The values are assumed to be the same for all units of a given type of asset in a given county, e.g. all pasture in Essex has the same price per acre, and all pasture in Warwickshire has the same price too, although this may differ from the price in Essex.

The notional values are unknown, in the sense that they are not recorded in the fines, but they can be inferred from the evidence. If, for example, there were ten types of asset then there would be ten notional values. Suppose that there were 100 recorded fines. Then there would be 100 data points from which to determine ten unknown values. In this case there are many more data points than there are prices to estimate, and so estimation is straightforward. It can be effected by running a linear regression of the consideration (the dependent variable) on the quantities of each type of asset mentioned in the fine (the independent or explanatory variables). The coefficients in this regression correspond to the unknown asset values. This method is well established in econometrics; it represents a special type of 'hedonic regression'. Estimates of the coefficients can be compared with estimates of asset values obtained from other sources of information.

To allow for excluded assets, a constant term is included in the regression; the evidence suggests that the value of the constant is quite substantial.

Another refinement is to allow for changes in valuation over time. Such changes could be the result of changes in the legal and administrative framework, and also in the kinds of transactions being recorded. It is also likely, however, that change over time will reflect changing economic conditions which alter asset values. The analysis allows for changes over time that affect both asset values in general, and the values of certain specific assets in particular.

Ideally, change over time would be analysed on an annual or decadal basis. There are, however, insufficient observations to do this effectively. This applies with particular force to changes in the valuation of particular types of assets. Suppose for example, that in 1000 fines mills appear in only 100 cases, which means that over 200 years there is on average only one mill occurring every two years or five mills in any decade. This is insufficient information to give reliable information on the changing value of mills. For this reason change over time is analysed using just four fifty-year periods. One of these periods must be used as a control, to form a basis from which changes in the other periods are measured, and following the usual conventions the first period is adopted as the control in this study. Even with fifty-year periods, some of the measured changes are quite sensitive to the nature of the fines recorded during the period. This applies particularly to the final period, 1450–99, when the number of fines diminished and the average size of asset portfolios increased.

As already noted, it is not only the nature of the assets that may influence the value of the consideration, but the nature of the parties involved as well.

Different types of plaintiff and defendant have been identified, according to marital status and occupation. Once again, there is an issue of small numbers in some cases, e.g. there are insufficient instances of attorneys acting as defendants in Essex for this category to be included in the analysis.

Personal factors are assumed to have additive effects on the value of the consideration. While addition makes sense for calculating the total value of an asset portfolio from its components, however, it is less appropriate when adjusting the value of a portfolio for the influence of personal factors. It could be argued that the impact of these personal factors should be scaled by the size of the asset portfolio. This makes relatively little difference to the results, however, and so in the interests of simplicity the issue is not considered further in this chapter.

It should be noted that considerations are often rounded to 'price points'; small sums are typically multiples of a mark or 160 pence, while large sums are multiples of 1600 pence, or even 16,000 pence. It is fairly clear, therefore, that the valuations are not necessarily exact. It would be possible to address this problem by treating the fine as a categorical rather than continuous variable, but this approach has not been pursued in this study.

Heteroskedasticity is also a potential concern and, in response, robust standard errors have been used throughout. Furthermore, visual examination of the data identified two fines in Essex and one in Warwickshire where the valuation was huge – more than twice that of any other fine. A high proportion of the explained variance in each case was accounted for by placing the estimated regression plane close to the outlier, and it was clear that these observations were exerting a considerable influence on the estimated prices of the assets included in the relevant portfolios. For this reason these cases were dropped, and the results reported below relate to the remaining cases only.

The regression results are presented in Table 5.4. The explanatory variables are listed in the left-hand column. The first two columns of numbers relate to Essex and the second two to Warwickshire. For each county, column A reports results in which the time factor is applied independently of the composition of the asset portfolio, while column B applies the time factor to selected assets individually. The assets selected for detailed time analysis are ones which are important for the historiography of the period and which were included in a high proportion of actual asset portfolios. The premia quoted for values in 1400–49 and 1450–99 are based on a comparison with the period 1300–49 and not with the immediately preceding period.

Each cell reports the estimated coefficient from an ordinary least squares regression, with the probability value (or significance level) below it. Three levels of significance are identified using asterisks: 1 per cent (***), 5 per cent (**) and 10 per cent (*). A dash indicates that the relevant variable was not included in the regression. In some cases quite large coefficient values appear insignificant; this is usually because there are few relevant positive observations, so that the standard error of the coefficient is just as large as the coefficient itself.

The results are discussed in the order they appear when reading down the table.

Messuages are of greater value in Warwickshire (average value 729 pence per messuage) than in Essex (average value 183 pence per messuage). Values in the period 1300–50 are very low (even negative in the case of Essex), but increase substantially in 1350–99. Thereafter experiences in the two counties diverge: in Warwickshire messuages retain their higher value in 1400–49, and witness a further increase in 1450–99, whilst in Essex values slump again in 1400–49 and then partially recover to 1350–99 values (but not significantly) in 1450–99.

Crofts are of negligible value, but miscellaneous buildings are of considerable value in Warwickshire and of small (but nevertheless significant) value in Essex.

Arable land in Essex (valued at 29 pence per acre) is worth more than twice arable land in Warwickshire (valued at 13 pence per acre) over the whole period. Its value does not vary significantly over time, except in Warwickshire in 1350–99, when it increases significantly before declining again afterwards.

Pasture is also more valuable in Essex (valued at 35 pence per acre) than in Warwickshire (valued at 12 pence per acre). The value does not vary significantly in Essex. In Warwickshire pasture has a relatively high value in 1300–49 (higher than in Essex) but then declines permanently (with 10 per cent significance) to a very low value after 1350.

Meadow is much more valuable than either arable or pasture. It is also more valuable in Warwickshire (146 pence per acre) than in Essex (103 pence per acre) – unlike arable or pasture. Its value in Essex is relatively stable (there are no significant changes), but in Warwickshire its value tends to decline, especially in the period 1450–99.

On the whole the value of land appears to be more volatile in Warwickshire than in Essex.

Woodland is of negligible value (small and insignificant coefficients in both counties, and negative in Warwickshire).

Manors, on the other hand, are of very considerable value, as expected. They are more valuable in Essex (average value 16,138 pence) than in Warwickshire (11,494 pence). Warwickshire manors show a dramatic increase in value in 1350–99, but then fall back again from 1400 onwards, with a modest but insignificant recovery in 1450–99. In Essex almost the exact opposite occurs. After an insignificant fall in 1350–99, values rise substantially in 1400–49 and quite enormously in 1450–99 (both increases are significant at 1 per cent).

Advowsons are also valuable, though less so than manors. Like manors, they are more valuable in Essex (4614 pence) than in Warwickshire (2294 pence), and reveal a sustained increase in value in Essex, but not in Warwickshire. Warwickshire experiences a significant boom in the value of advowsons 1350–99, which mirrors the increase in the value of Warwickshire manors at this time, but this is not sustained, and there is a substantial (though marginally insignificant) decline in value in 1450–99. In Essex, on the other hand, the value of advowsons increases (though not significantly) in 1350–99; the increase is sustained 1400–49, and then there is a remarkable increase (significant at 5 per cent) in 1450–99. This increase mirrors the dramatic increase in the value of Essex manors at this time.

Table 5.4 Analysis of the factors affecting the value of the consideration paid in a fine

	Essex		*Warwickshire*	
	A	*B*	*A*	*B*
Constant	4731.97*** (0.000)	5975.59*** (0.000)	3738.23*** (0.000)	4758.70*** (0.000)
Messuages (number)	183.28 (0.447)	–343.86 (0.161)	729.42*** (0.000)	30.22 (0.931)
Messuage premium 1350–99	–	775.40** (0.031)	–	616.11* (0.088)
Messuage premium 1400–49	–	–3.74 (0.993)	–	738.42** (0.041)
Messuage premium 1450–99	–	529.79 (0.352)	–	1385.78*** (0.003)
Croft	–257.70 (0.608)	–230.27 (0.718)	215.50 (0.265)	186.45 (0.520)
Miscellaneous buildings	60.253** (0.019)	55.03 (0.140)	1398.60*** (0.009)	1417.24*** (0.005)
Arable land (acres)	29.27*** (0.000)	33.32 (0.000)	13.21** (0.016)	12.59* (0.065)
Arable land premium 1350–99	–	–10.65 (0.422)	–	15.11* (0.086)
Arable land premium 1400–49	–	–1.030 (0.931)	–	1.54 (0.874)
Arable land premium 1450–99	–	1.81 (0.907)	–	–1.99 (0.914)
Pasture (acres)	35.36** (0.036)	103.92 (0.275)	12.49 (0.521)	199.31* (0.050)
Pasture premium 1350–99	–	3.313 (0.975)	–	–202.882* (0.100)
Pasture premium 1400–49	–	–85.09 (0.382)	–	–194.53* (0.086)
Pasture premium 1450–99	–	–105.93 (0.284)	–	–173.11* (0.092)
Meadow (acres)	103.49** (0.018)	127.00 (0.207)	146.28** (0.013)	316.49** (0.032)
Meadow premium 1350–99	–	116.46 (0.416)	–	–95.80 (0.563)
Meadow premium 1400–49	–	77.64 (0.557)	–	–145.82 (0.422)
Meadow premium 1450–99	–	–49.88 (0.669)	–	–278.22** (0.045)
Woodland (acres)	35.82 (0.123)	29.35 (0.243)	–40.27 (0.317)	9.86 (0.840)
Manors (number)	16137.68*** (0.000)	11264.17*** (0.000)	11494.19*** (0.000)	9243.10*** (.001)
Manor premium 1350–99	–	–837.75 (0.681)	–	10380.00*** (0.004)
Manor premium 1400–49	–	8055.36*** (0.002)	–	58.36 (0.985)
Manor premium 1450–99	–	16108.16*** (0.000)	–	2276.92 (0.522)
Advowsons (number)	4614.28** (0.029)	3103.28 (0.113)	2294.05 (0.231)	2417.36 (0.301)
Advowson premium 1350–99	–	4563.95 (0.128)	–	6988.67** (0.016)
Advowson premium 1400–49	–	3784.65 (0.330)	–	–153.44 (0.963)
Advowson premium 1450–99	–	17096.13** (0.041)	–	–20255.66 (0.105)
Mills (number)	1197.27 (0.373)	732.25 (0.563)	3916.09*** (0.001)	4023.58*** (0.004)

Rent (pence)	5.87*** (0.000)	4.151*** (0.002)	4.32** (0.012)	3.90** (0.031)
Town location only	–1419.46*** (0.000)	–1913.37*** (0.000)	–2107.77*** (0.000)	–2734.73*** (0.000)
Town location premium 1350–99	–	–157.84 (0.702)	–	1595.60*** (0.008)
Town location premium 1400–49	–	2224.24*** (0.009)	–	2895.00*** (0.001)
Town location premium 1450–99	–	3102.24*** (0.001)	–	–102.41 (0.922)
Town and rural location	1220.31 (0.109)	735.35 (0.546)	476.93 (0.635)	–2324.85** (0.021)
Town and rural premium 1350–99	–	119.17 (0.939)	–	3786.21** (0.015)
Town and rural premium 1400–49	–	465.81 (0.793)	–	4433.14 (0.637)
Town and rural premium 1450–99	–	3573.34 (0.386)	–	–1681.05 (0.637)
Plaintiff married man	128.04 (0.932)	–874.91 (0.576)	1510.87 (0.317)	1589.92 (0.269)
Plaintiff married woman	–550.12 (0.715)	271.13 (0.864)	–1846.67 (0.241)	–1.614.92 (0.287)
Plaintiff single woman	–1022.68 (0.715)	–1629.89** (0.033)	–538.69 (0.487)	–611.72 (0.406)
Plaintiff widow	4131.01** (0.024)	4220.70** (0.017)	38.28 (0.984)	–1130.62 (0.432)
Plaintiff clergy	1255.81*** (0.001)	1254.92 (0.003)	234.58 (0.631)	224.27 (0.592)
Plaintiff attorney	511.25 (0.187)	412.91 (0.267)	400.69 (0.679)	73.21 (0.937)
Plaintiff trade	97.06 (0.782)	755.66 (0.038)	–530.11 (0.230)	–567.75 (0.186)
Defendant married man	–7031.47*** (0.000)	–7.154*** (0.000)	–1385.41* (0.093)	–1181.86 (0.140)
Defendant married woman	5632.71*** (0.003)	5717.85***(0.001)	549.41 (0.338)	683.79 (0.221)
Defendant single woman	2284.51 (0.156)	2951.73 (0.135)	–1772.27* (0.082)	–1881.93* (0.060)
Defendant widow	–948.11 (0.615)	90.16 (0.958)	1022.99 (0.400)	666.41 (0.635)
Defendant clergy	–2098.24** (0.03)	–1370.26* (0.090)	451.23 (0.774)	198.99 (0.872)
Defendant attorney	–	–	919.90 (0.514)	789.82 (0.574)
Defendant trade	5785.93*** (0.003)	5980.16*** (0.000)	–1484.29 (0.264)	–710.50 (0.529)
Dummy variable for 1350–40	804.79*** (0.002)	–	3511.76*** (0.000)	–
Dummy variable for 1400–49	1692.74*** (0.001)	–	2700.65*** (0.000)	–
Dummy variable for 1450–99	6360.10*** (0.000)	–	2258.27*** (0.001)	–
R^2	0.595	0.632	0.563	0.613
Adjusted R^2	0.591	0.626	0.550	0.593
F statistic	168.28 (0.000)	113.26 (0.000)	42.461 (0.000)	30.082 (0.000)
Number of observations	3352	3352	1020	1020

Sources: see Tables 5.2 and 5.3.

Notes

Arable land is calculated as a total acreage by counting bovates as 15 acres, virgates as 30 acres and carucates as 120 acres. Control variables are rural location, single men and time-dependent variables relating to 1300–49. Outliers have been excluded in cases where the value of the dependant variable is double, or more than double, that of any other case. There are two exclusions in Essex, and one in Warwickshire.

Mills are much more valuable in Warwickshire than in Essex. The value of a mill is 3916 pence in Warwickshire and only 1197 pence in Essex. This is difficult to explain, given that price of arable land is higher in Essex and that manors are more valuable too. One possibility is that there were more mills in Essex, or that transport of grain was easier, and this made mills more competitive and so reduced a mill-owner's monopoly power.

Rent is valued at 5.87 pence in Essex and 4.32 pence in Warwickshire. If rent income is simply valued as income, rather than as an indicator of power and status, then these figures suggest that one pence in perpetuity is valued at 5.87 pence in Essex, equivalent to an interest rate of $(1/5.87) \times 100 = 17$ per cent, and to $(1/4.32) \times 100 = 23$ per cent in Warwickshire. These values are rather high, and may perhaps be explained by a perception that rental income was not secure and was liable to default.

The results for location are unambiguous: over the period 1300–1500 as a whole, town location reduces valuations substantially. There are significant differences between counties and over time, however. Consider first the coefficients on the 'town only' variables that capture asset portfolios located entirely in towns. They show that in Warwickshire the negative impact of town location is strongest at the beginning and end of the period (1300–49 and 1450–99), but small or even negligible in the middle of the period (1400–99); towns appear to have recovered parity with the countryside by 1400–49, but this is not sustained thereafter. In Essex, on the other hand, the situation has been reversed by the end of the period, with towns staging a strong recovery from 1400 onwards (significant at 1 per cent).

The results for town and rural locations, which relate to geographically distributed asset portfolios, are less clear cut, as would be expected, but they are consistent with the interpretation above. In Essex there are no significant effects, suggesting an overall parity between town and country, while in Warwickshire there is a significant improvement in the position of town and country portfolio in the period 1350–99 which is not sustained.

The gender and marital status of the buyer (plaintiff) and seller (defendant) have a significant impact on valuations, but there are marked differences between the counties, suggesting that local context is important where status are concerned. Status appears to be more important in Essex than in Warwickshire.

In Warwickshire the marital status of the plaintiff has no significant impact on the value of the consideration once the asset composition of the portfolio is taken into account. So far as defendants are concerned, only single women experience significant impacts linked to status; the value of the consideration is lowered. Occupation and trade have no effects on considerations; although some of the coefficients are quite large, none of them is significant.

In Essex the situation is very different. Among the plaintiffs, widows are associated with very high considerations (significant at 5 per cent), and regression B suggests that single women are associated with low considerations (also significant at 5 per cent). Where defendants are concerned, there is a clear asymmetry between married men (associated with a low consideration) and married women (associated with a high consideration) – both at 1 per cent significance. The status of clergyman is important for both plaintiffs and defendants as they are part of a wider trend in the creation of trusts or feoffments as strategies for managing inheritance and the descent of property; as plaintiffs they are associated with high considerations (at 1 per cent significance in regression A) and as defendants with low considerations (at 5 per cent significance in regression A and 10 per cent significance in regression B). By contrast, defendants in trade are associated with very high considerations (significant at 1 per cent in both regressions).

The dummy variables that appear at the end of the table apply to regressions A only; their coefficients reflect the impact of time on asset valuations as a whole, rather than on the valuation of any specific type of asset. The dummy variables highlight both similarities and differences between the two counties. They reveal the period 1300–49 as a period of relatively low asset values in both counties. From 1350 onwards values are significantly higher than they were before (at 1 per cent significance in each period for both counties). The trend is very different, however. Warwickshire experiences a substantial improvement after 1350, but this is not sustained. Values remain higher than before 1350 but they decline steadily after 1400. In Essex the rise in 1350 is much more modest than in Warwickshire, but the upward trajectory is not merely sustained – it accelerates substantially, with a dramatic increase in values in 1450.

Before drawing conclusions from these results it is appropriate to validate them by comparing them with mainstream historiography. As noted above, the valuations of property derived from fines should be interpreted as capital values. These capital values can be compared with the annual net profit or ‘rent’ which are recorded in IPMs. The IPM data relates to the period 1270–1349, and so it is appropriate to focus on a comparable period for fines, which was chosen as 1300–39. The IPM data averages across locations, and does not adjust for the personal characteristics of owners, and so comparable estimates from fines were derived by re-estimating the regression using only fines from the period 1300–39 and excluding location and personal variables. The coefficients obtained for the capital values of the three main types of agricultural land were then compared with corresponding ‘rent’ values from IPMs. The results are shown in Table 5.5. The rows report estimated values for arable, meadow and pasture; the three left-hand columns report ‘rent’ values derived from IPMs and the right-hand columns capital values derived from fines. The valuation of pasture is somewhat erratic in both sources. The values of both arable and meadow are easily reconciled, however, by supposing that rents are capitalised at approximately 10 per cent per annum, or equivalently that a fine corresponds to ‘ten years’ purchase’. It is interesting to note that this corresponds closely with Bean’s (1991, p. 567) assessment that purchase price in the thirteenth and early fourteenth century represented ten years’ annual value.

Another way of validating the results is to examine the decadal variation of valuations as a whole. In Table 5.4 time variation was analysed only over fifty-year intervals because there is insufficient data to estimate decadal variation

Table 5.5 Comparative values in pence of agricultural land

Asset	*IPM 1270–1349 Range: Campbell*	*IPM 1270–1349 Mode: Campbell*	*IPM 1270–1349 Mean: Campbell*	*IPM 1270–1339 Mean: CMH*	*Fines 1300–39 Mean: Essex*	*Fines 1300–39 Mean: Warwickshire*
Arable per acre	1d–36d	4d	4.8d	4.9d	36d (3.6d)	27d (2.7d)
Meadow per acre	3d–80d	12d	18.6d	21.2	158 (15.8d)	369 (36.9d)
Pasture per acre	fraction of 1d–24d	less than 2d	3.4d	7.8	–6d (–0.6d)	157d (15.7d)

Sources: Campbell and Bartley (2006, pp. 168–169); Centre for Metropolitan History (CMH) (2009); Fowler (1929–49); Kirk (1913–28); Reaney (1964); Stokes and Drucker (1939).

separately for each type of property. There is, however, sufficient data to analyse decadal variation in the value of the property as a whole. The procedure was to eliminate all the variables involving fifty-year sub-periods and to add a set of decadal dummies, using the period 1300–9 as the control.

The results are summarised graphically in Figure 5.2. By construction, the initial base level is the same for both counties. It should be noted that the smaller deviations from the base line are statistically insignificant (particularly in Warwickshire, where the number of observations is smaller). The figure shows that Essex values rose faster than Warwickshire values over the period as a whole. There was a peak in the 1360s in Warwickshire and in the 1380s in Essex, before a sharp decline that 'bottomed out' in the 1400s in Warwickshire and the 1410s in Essex. There was a dramatic rise in the 1420s and 1430s, which was then sustained in Essex and reversed in Warwickshire. By the 1490s Essex values were very high compared to 1300–9, while in Warwickshire the values were virtually unchanged. When interpreting these results it is important to remember that the value of money changed over this period (see Chapter 3), and that the nature of the fines also changed, with fines on average reducing in number and increasing in value. The regression analysis shows that part of the increase in value in the second half of the fifteenth century arose from changes in the size and composition asset portfolios, but that, at least in Essex, part of the change is also accounted for by changes in the value of individual items, such as manors.

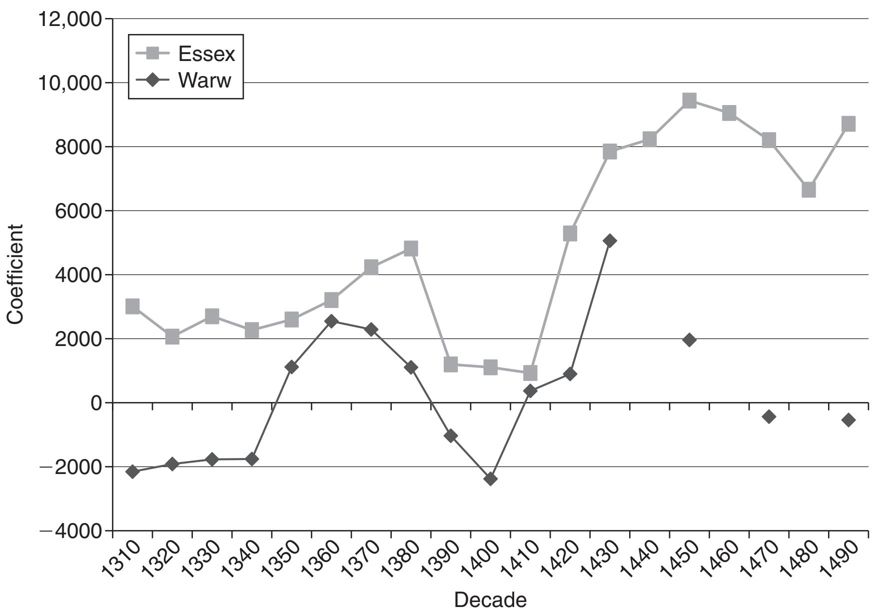

Figure 5.2 Changes in the value of property by decade.

5.6 Conclusion

Overall the results suggest that after a period of recovery in 1350–99 from the adversities of the period 1300–50, the fortunes of Essex and Warwickshire diverged at an accelerating rate. This is reflected primarily in the values of manors, advowsons and urban properties. It is not, however, reflected in the values of arable, pasture or meadow to anything like the same extent. Although there are substantial differences between Essex and Warwickshire in the values of particular types of land, the relative values of different types of land do not change significantly over time in either county. Yet the evidence from the fines themselves, reported in Tables 5.2. and 5.3, shows that land in both counties was increasingly being diverted from arable into pasture – almost certainly in response to developments in the wool and cloth trades at this time along with other factors explained above. It is important to recognise, however, that where land is easily substituted between different uses, a very small adjustment in relative prices can lead to a very large adjustment in the use of land. This suggests that it is not changing land values associated with the reallocation of land that was driving increased valuations in Essex, but rather the demand for manors and advowsons, and increasing interest in the commercial use of urban properties. Land values are higher, on average, in Essex than in Warwickshire, probably because of easier access to the London market, but these values do not change sufficiently to explain the rising valuations of manors, advowsons and urban property. These are best explained as the consequence of mercantile affluence in towns and in the countryside the purchase of manors and advowsons by those with aspirations to enter ‘gentle’ society.

Rising prosperity in Essex in the fifteenth century is usually attributed to its proximity to London, which was emerging as the commercial hub of both the national and export trade in cloth. This was linked to the expansion of cloth production in towns and villages. The evidence from the statistical analysis supports the view that the changing values of assets was driven by strategies of investment by wealthy individuals, especially the growing numbers of professionals such as royal or government employees, lawyers or merchants and their desire to acquire prestigious country estates. It was the rising affluence of successful individuals and their desire for influence in county affairs, rather than the restructuring of agriculture per se, that was the main influence on changing asset values in the fifteenth century. It explains both the different levels of valuation in Essex and Warwickshire, and the local trends in these valuations too.

Note

1 We are indebted to the Economic History Society for the award of a grant to fund the work of data entry by Anna Campbell for this project.

References

Manuscript feet of fines, The National Archive (TNA), CP25/1 with research guide at www.nationalarchives.gov.uk/records/research-guides/land-conveyance-feet-of-fines.htm A list of county series can be found at www.medievalgenealogy.org.uk/fines/counties.shtml including bibliography and links to images of manuscript fines.

Bean, J.M.W. (1991) Landlords, in Edward Miller ed., *The Agrarian History of England and Wales vol. III 1348–1500*, Cambridge, Cambridge University Press, 526–586.

Britnell, Richard H. (1993) *The Commercialisation of English Society, 1000–1500*, Cambridge: Cambridge University Press.

Broadberry, Stephen, Bruce M.S. Campbell, Alexander Klein, Mark Overton and Bas van Leeuwen (2012) British Economic Growth, 1270–1870, working paper at www2.lse.ac.uk/economicHistory/seminars/ModernAndComparative/papers2011–12/Papers/Broadberry.pdf.

Campbell, B.M.S. (2000) *English Seigniorial Agriculture, 1250–1450*, Cambridge: Cambridge University Press.

Campbell, B.M.S. and K. Bartley (2006) *England on the Eve of the Black Death. An Atlas of Lay Lordship, Land and Wealth, 1300–49*, Manchester, Manchester University Press.

Centre for Metropolitan History (2009) Inquisitions Post Mortem (IPM) database compiled by the 'Feeding the City (I): London's impact on the agrarian economy of southern England, *c.*1250–1350' project funded by the Leverhulme Trust (F.264C).

Davies, M. and J. Kissock (2004) The feet of fines, the land market and the English agricultural crisis of 1315 to 1322, *Journal of Historical Geography*, 30, 215–230.

Drucker, L. (ed.) (1943) *Warwickshire Feet of Fines. Volume 3. 19 Ed. III (1345)-24 Hen. VII (1509)*, Warwick: Dugdale Society, Vol. 18.

Dyer, Christopher (1980) *Lords and Peasants in a Changing Society: The estates of the bishopric of Worcester, 680–1540*, Cambridge: Cambridge University Press.

Dyer, Christopher (1981) Warwickshire Farming, 1349–c.1520, Warwick: Dugdale Society, *Dugdale Society Occasional Papers* 27.

Fowler, R.C. (ed.) (1929–49) *Feet of Fines for Essex, Vol. III, 1327–1422*, Colchester: Essex Archaeological Society.

Kirk, E.F. (ed.) (1913–28) *Feet of Fines for Essex, Vol. II, 1272–1326*, Colchester: Essex Archaeological Society.

Poos, Lawrence R. (1991) *A Rural Society after the Black Death, Essex 1350–1525*, Cambridge: Cambridge University Press.

Reaney, P.H. (ed.) (1964) *Feet of Fines for Essex, Vol. IV, 1423–1547*, Colchester: Essex Archaeological Society.

Stokes, E. and L. Drucker (eds) (1939) *Warwickshire Feet of Fines, Vol. II. 13 Ed. I (1284)-18 Ed. III (1345)*, Warwick: Dugdale Society, Vol. 15.

Turner, M.E., J.V. Beckett and B. Afton (1997) *Agricultural Rent in England, 1690–1914*, Cambridge: Cambridge University Press.

Watkins, A. (1997) Landowners and their estates in the Forest of Arden in the fifteenth century, *Agricultural History Review*, 45, 18–33.

Yates, Margaret (2012) Descriptions of land, in Michael A. Hicks (ed.) *The Fifteenth-century Inquisitions Post Mortem: A Companion*, Woodbridge: Boydell, 145–154.

Yates, Margaret (2013) The market in freehold land: The evidence of feet of fines, 1300–1509, *Economic History Review*, 66(2), 579–600.

6 Visual analytics for large-scale actor networks

A case study of Liverpool, 1750–1810

John Haggerty and Sheryllynne Haggerty

6.1 Introduction

This chapter outlines some key aspects and advantages of using Visual Analytics (VA) for the analysis of large data sets. VA incorporates a wide range of visualisation and statistical techniques to provide an explicitly exploratory interaction with researchers' data. VA does not provide answers to research questions, so much as provide further research questions through visualising and measuring the users' data in a wide variety of ways. VA may be used for small data sets, but it is particularly useful for large data sets, as outlined here. The context for this chapter is Liverpool's slave trade and related associational networks, 1750–1810, but VA can be used for the analysis of any actor (people) networks, whatever the context and time period. The sections of this chapter are as follows. First the use of networks as a framework of analysis within history is outlined, especially where this has included the interdisciplinary adoption of socio-economic theory. Then follows an outline of VA, what it can do, and a brief consideration of the benefits of, and issues with, its use. After introducing the case study to demonstrate this methodology, it then describes two different visualisation techniques for the analysis of actor networks: static network analysis and the measurement of multiple networks via clustering and centrality measurement. Finally there is a brief conclusion.

6.2 Networks in history

Networks have long been used in history as an analytical tool to explore actor or people networks. This is particularly true within the context of commercial relations, where they have often been seen as a simple economic good. This is due to the fact that many historians concentrated on familial, ethnic and religious ties, which were seen to promote trust and reduce information and transaction costs (Davidoff and Hall, 1987; Mathias, 2000; Trivellato, 2009; Prior and Kirby, 1993; Walvin, 1997). More recently, however, historians have questioned this rather positive view of networks, and have looked at how and why networks failed to fulfil their purpose. Certainly the family network was not always reliable and often incurred obligations that were difficult to get out of or re-negotiate (Haggerty, 2011), and ethic networks can be seen as having restricted choices (Tilly, 1995). Moreover, networks could be difficult and time consuming to develop and maintain (Hancock, 2005; Popp, 2007). While metropolitan business networks are usually seen as efficient (Casson, 2003), their members can form cliques that become opaque and even corrupt (Crumplin, 2007). Work has also been conducted on institutional networks such as town guilds, but these too have often been found wanting, being seen as backward looking and restricting progress (Ogilvie, 2004; Rosenband, 1999). Often they were not used for their alleged purpose in any case (Goddard, 2013). This is of course not to say that the network as an organisational form was, and is, not useful, but that historians are developing a more balanced view of their use and abuse.

A crucial part of this development has been the adoption of socio-economic theory. This has promoted a far more nuanced view of networks. Perhaps the most widely adopted concept in this regard is Granovetter's 'strength of weak ties' (1973). The idea is not that all weak ties are good or useful, but those that provide a bridge to other networks potentially provide new and improved access to information, capital and, indeed, further networks. Such bridges have the potential to provide crucial information or access where a structural hole (Burt, 2004) exists in a network which is restricting progress. Not all ties within a network are equally valuable or useful; some people are more important than others. Therefore, power relationships come into play. That power is of course relative (Bonacich, 1987), and a person who is powerful in one network may not be so in another. This means that many networks are in existence and use at the same time (Freeman, 1978–9) and people may well play different roles in each of them. Indeed, they may use different networks for different purposes, which may or may not cause friction (Lawler and Yoon, 1996). In contrast, some networks can become so staid or inactive that they produce negative social capital (Portes, 1998). People therefore move in and out of networks for a variety of reasons, making networks dynamic, arguably with a life cycle similar to industrial or business clusters (Haggerty and Haggerty, 2011; Swann and Prevezer, 1998).

Perhaps most importantly, adopting socio-economic theory has forced historians to think about exactly what a network is, and what they are for (Duguid, 2004; Haggerty, 2012; Hancock, 2005). Depending on the context of course, these definitions change. A network may comprise merely a number of actors with specific types of connections (Smith-Doerr and Powell, 2005), a community of people who share common attributes (Rauch, 2001) and, within the commercial context, a set of people who provide high-trust linkages (Casson, 2003). In terms of purpose, a child wants its family network to provide love and general support (Renzulli *et al.*, 2000), a teenager wants their networks to provide cultural capital (Bourdieu, 2001), while business people want their networks to provide new opportunities for profit, by providing reliable information, reducing risk and promoting trust. In particular, business networks should promote the prompt transfer of information, confer status or legitimacy, and afford lower information and transaction costs, by enforcing contracts and reducing information asymmetries (Podolny and Page, 1998; Rauch, 2001). Historians are therefore thinking about the complexity of networks in general, their positives and negatives, and what they are for in the first place. This chapter argues that historians (and others) can further refine and develop their analysis of networks through the adoption of VA, which uses techniques taken from mathematics, computer science and social science.

6.3 Visual analytics

VA is an exciting way of exploring large-scale actor networks, or indeed any large data set. It incorporates a wide range of visualisation and statistical techniques to provide an explicitly exploratory interaction with a researcher's data (Perer and Schneiderman, 2009). VA does not provide answers to research questions, so much as provide further research questions through visualising and measuring the data in a wide variety of ways. Figure 6.1 illustrates different possible views of network data in *Matrixify* (Haggerty and Haggerty, 2012), ranging from temporal views to illustrate the effect of exogenous and endogenous events, through to node sizing based on relational information to network statistics. Users are encouraged to explore the different views, and therefore their data, using such tools.

By visualising and measuring their data in these ways, historians are forced to ask new questions of their data. In terms of networks, for example, the questions might include:

- How active were actors in the network(s) over time?
- Were actors using a variety of networking opportunities at any one time?
- Did particular groups/cliques dominate the network(s)?
- What were the effects of endogenous and exogenous events and how did actors respond to them?
- How did shifting actor engagement impact access to information and social, financial and human capital?
- What were the relationships between various subnets within the network?

Central to VA is that answering these questions usually entails revisiting the primary sources iteratively with these new questions, promoting new research in turn. This may well include running new visualisations and measurements with an enhanced data set.

VA is particularly good for large-scale data sets because they often comprise information that is both structured or tangible (e.g. nominal data from membership lists or parish records, or a list of acquaintances from a letter book or diary) and unstructured or intangible data (e.g. the relationship between those members or acquaintances). Structured or tangible data is often quite easy for historians to represent through traditional methods such as graphs or tables. Indeed, unstructured data from qualitative sources such as letters or diaries for example, is usually pre-

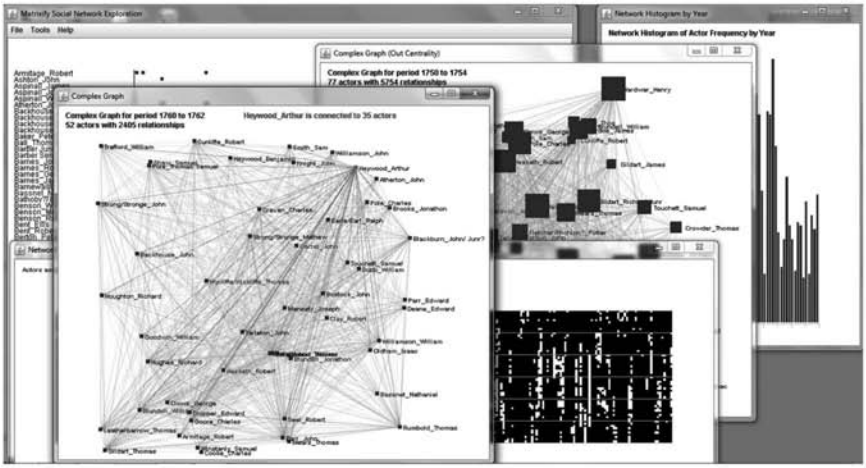

Figure 6.1 Possible visualisations using VA (source: Haggerty and Haggerty, 2012).

sented in textual form. However, this is not possible where there is no matching qualitative data for the quantitative sources, or conversely, where there are vast amounts of data. However, VA can measure the relationships of actors from quantitative data – thereby providing the 'intangible' data from the tangible; for example, major or central players can be identified, as can the relationships between them. This is achieved through a variety of visualisation and statistical techniques including, but not restricted to: network clusters, histograms, network graphs, centrality measures, network density and regression analysis. VA can cope easily with the representation of a variety of data at one time, but importantly, can measure complicated relationships within that data over time. Many tools are already freely available online for this type of analysis. These include Gephi (2013), Gretl (2013), Pajek (2013), SocNetV (2013) and Visone (2013).[1]

As mentioned above, almost any type of data set can be interpreted using VA. The data simply needs be to be formatted in a certain way. For example, in actor network analysis this is in a simple adjacency matrix format, noting a 0 (zero) where no relationship exists and a 1 (one) where one does exist. For most actor networks this is achieved simply by listing all the actors along both the horizontal and vertical axes and then in the matrix entering a zero where they do not know each other, and a one where they do. Some software requires a text file input (e.g. using Notepad) whereby the actors are numbered and the relationships identified by those numbers (see examples below). It is also possible to complicate these visualisations using partitions or clusters (terminology is dependent on the software used) to assign 'value' judgements to the actors. This might include the type of relationship (familial, ethnic, credit), or other values such as geography, gender, occupation, etc. An example of using such software is presented in Table 6.1 below.

Table 6.1 Adjacency matrix for Figure 6.2

	Point_1	*Point_2*	*Point_3*	*Point_4*	*Point_5*
Point_1	–	1	0	0	0
Point_2	1	–	1	0	0
Point_3	0	0	–	1	0
Point_4	1	1	0	–	0
Point_5	0	0	1	0	–

The adjacency matrix forms the basis of a graph. A *graph* (G) comprises a set of *vertices* (or *points*, *V*) (the actors) which are connected to other points through *edges* (or *arcs*, *E*) (the relationships between the actors); or $G=(V, E)$. *V* and *E* are taken to be finite and a vertex can exist within a graph and not have any associated edges, i.e. they are unconnected to other vertices. Two vertices that are directly connected by an edge are said to be *adjacent*. The number of other vertices to which any given vertex is adjacent is called the *degree* of that point. The *distance* between vertices is calculated by the number of edges in a path. The shortest paths linking pairs of vertices (as there may be many paths linking vertex pairs (actors) within a network) are called *geodesics*. Vertices falling on the geodesics between a given pair of vertices stand *between* these points. The *order* of a graph is the number of vertices while its *size* is expressed as the number of edges. Conceptualising networks in this way enables us to visualise, quantify and measure relationships between actors as well as the network institution, as will be demonstrated in our case study. A simple graph, based on Table 6.1, is shown in Figure 6.2, in which Points 1 to 5 represent the actors.

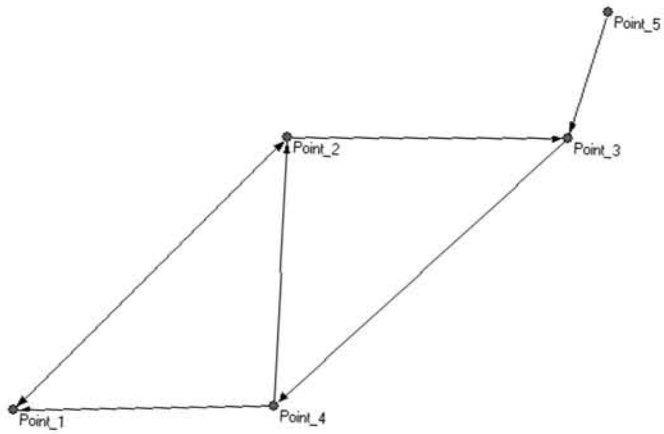

Figure 6.2 Simple graph (source: authors' own drawing).

The input in a text editor to produce this graph in Pajek is as follows:

```
/* DJLV1-PAJEK */

*Vertices 5
1 "Point_1"
2 "Point_2"
3 "Point_3"
4 "Point_4"
5 "Point_5"
*Arcslist
1 2
2 1 3
3 4
4 1 2
5 3
```

Once a graph has been visualised, it can be laid out and measured in a number of ways, as shown in Figure 6.1. This chapter however, outlines two in detail. The researcher may lay out the graph manually to present a static view of the network, as shown in section 6.5 below.[2] Alternatively, the researcher can cluster actors (vertices) representing multiple networks based on statistical analysis to identify the relationships *between* actors and *between* networks, as demonstrated in section 6.6.

6.4 The case study: Liverpool's slave trade and related associational networks, 1750–1810

Liverpool provides an excellent prism through which to consider how British merchants managed an increasingly global economy, but one in which there were also many wars (1756–63, 1775–83, 1793–1815) and various credit crises (1763, 1772, 1793). The port was well connected to the hinterland and international markets, and fulfilled various functions as an entrepôt, insurance and financial centre (Price, 1996). Liverpool benefitted greatly from Britain's Atlantic trade and was hailed as the second city of the realm during this period. Liverpool was also the leader in the slave trade after 1750, traded with Europe, Nova Scotia, throughout the thirteen colonies/states and had excellent connections with the British West Indies as well as a vibrant coastal trade (Haggerty, 2012). Expanding trade meant that imports increased from 14,600 tons in 1709 to 450,000 tons in 1800 (Marriner, 1982). This success encouraged in-migration and Liverpool's population grew from 6,500 in 1708 (Lawton, 1953) to 77,653 in 1801(BPP, 1968). The town also had a variety of formal and informal institutions as befitted its status, some of which are analysed here (Wilson, 2008).

Liverpool was the leader in one of the most important trades in the Atlantic world, the slave trade, which gives the commercial networks of those involved in it particular importance. Several sets of records are used here to form the case study of Liverpool's slave trade and associated networks 1750–1810. Sales ledgers, an online database and manuscripts have been used for the static network analysis. For the clusters and centrality measures, the membership records of several formal trade and cultural associations are used. These are detailed further below. The following introduction to VA methods therefore shows that a variety of records can be used for VA. In particular, it highlights the way in which nominal data (structured or tangible data) can be used to measure unstructured (qualitative or intangible) data.

6.5 Static network analysis

A static network analysis (or graph) is one which captures a network at a particular point in time, or collates data over a period of time, and then presents it as a snapshot of that period. This section demonstrates what such an analysis can reveal about that network, and what questions it might raise. This type of simple network can be drawn in Gephi, Pajek, SocNetV or Visone.

Samuel Rainford was a merchant based in Kingston, Jamaica, between 1774 and 1798.[3] Jamaica was a good choice, being the leading British West-India island at this time. Having emigrated from Liverpool, he also had good contacts in the leading British slave trade port, specifically Jonathon Blundell senior. Although his merchant house (he was in business with his brother, Robert) dealt in a wide variety of goods, it also dealt with a significant percentage of slave sales in Kingston. The slave sales ledger of the business, which covers the period 1779–93 and lists thirty-seven slave voyages, was used to construct Rainford's networks in Jamaica. The ledger listed all the slave ships that the house acted for, the number of slaves sold, and to whom. The slave trade database online (at www.slavevoyages.org) was used to form the Liverpool end of his networks. This database outlines the majority of slave trade voyages undertaken across the Atlantic in the early modern period; among other information it lists the names of the vessel, place of departure and port where the ships disembarked the slaves. It also lists all the owners of the slave ship. The ships mentioned in the slave sales ledger were cross-referenced with the slave trade database, the first listed owner was taken to be the ship's husband, and this person and the ship's captain were taken to have direct relationships with Rainford either by letter or in person. The remaining owners were taken to be indirect relationships with whom Rainford had no direct contact. Together these formed the quantitative data for the analysis of Rainford's slave trade networks over the period 1774–98. Figure 6.3 shows these networks as a static graph. This visualisation was created using Pajek. Actors have been laid out geographically, with the Jamaican networks in the left-bottom corner, and the Liverpool ones in the top right.

It is immediately obvious that Rainford had far more contacts in Jamaica than in Liverpool (over 350). However, these relationships appear to be of equal importance to Rainford, forming a 'star' network. In contrast, the networks in Liverpool are more complicated and with various subnets, so it makes sense to focus on them. The actors highlighted in grey are ships' captains who had direct

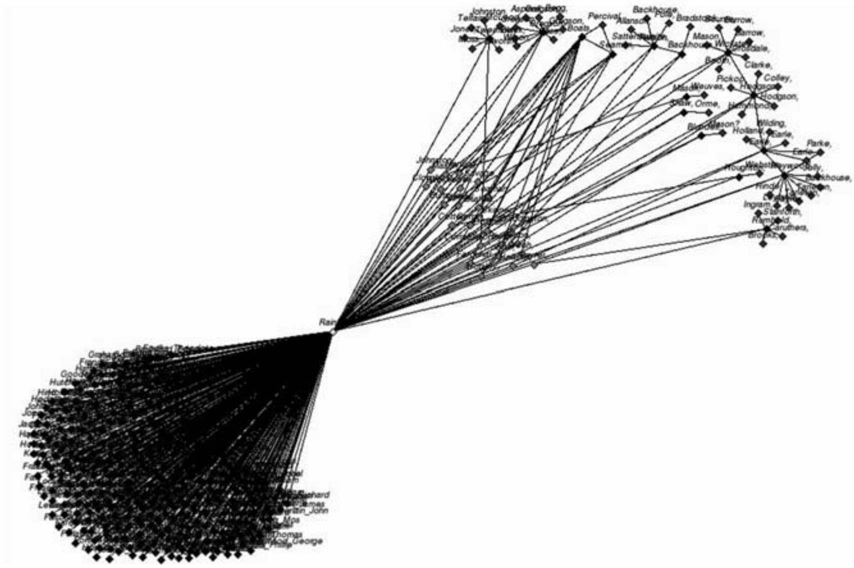

Figure 6.3 Rainford's slave trade networks, 1774–98 (source: adapted from Haggerty and Haggerty, 2010, Figure 1).

relationships with Rainford which could be classed as strong ties, and those in black are the ships' investors, combining both direct and indirect relationships, or strong and weak ties. One question should be: Why are the groups of actors the size they are? In the commercial literature (Doerflinger, 1986; Hancock, 1995; Morgan, 1993) merchant groups or partnerships usually comprised two or three actors. However, the slave trade was known as a particularly risky trade (Haggerty, 2010), and the figure clearly highlights the fact that Liverpool merchants pooled in larger groups to bring together a wider breadth of human and financial capital, in order to combat that higher risk.[4] Some of the subnets are interconnected, others discrete. Clearly some actors chose to work in isolation, whereas others actively pooled their human and financial capital. Following up the names of some of the individuals in the historiography shows that many leading Liverpool mercantile families were in this network (e.g. the Tarletons, the Backhouses and the Earles), demonstrating that Rainford was well connected. Many of these people were on the Town Council and the African Committee of Merchants Trading to Africa from Liverpool, which will be discussed in section 6.6.

The actor in Figure 6.3 with the most ties is William Boats, followed by John Gregson, Thomas Earle and then Benjamin Heywood. William Boats is particularly interesting because he illustrates the way that career progression of ships' captains in this trade served to retain human capital within the network (Behrendt, 2007). However, Boats is also shown as quite an independent actor, working in a relatively small investment group with Thomas Seaman and James Percival, mostly as lead investor (re-visiting the online slave trade database revealed that Seaman only acted as lead investor occasionally between 1788 and 1790). Boats may have been choosing to increase his profits, but at the same time he was heightening his risk. It is also possible that his preferences were to trust only a few people, or perhaps that others did not want to work with him for some reason.[5] In contrast to larger investment groups, Boats invested in several voyages each year, and he was therefore connected to several ships' captains. A return to the slave sales ledger of Rainford also highlights another reason for Boats' importance; after 1787 he was the only merchant from Liverpool to send slaves for sale to Rainford. The reasons for this needed further research, as the nominal records do not explain this. The manuscripts show this was due to a breakdown in relations between Rainford and a key Liverpool contact, Jonathon Blundell. Blundell used to invest heavily in the slave trade and no doubt had introduced Rainford to his slave trade contacts. However, as Blundell was no longer investing in the slave trade during the period studied here, he is absent from the figure. The manuscripts showed that despite this, he was central to Rainford's business and remained well connected to those still involved in the slave trade. It would appear that sometime before 1787 he used his influence to get slave traders (apart from Boats, Seaman and Percival) to withdraw their connections with Rainford.

Thereafter, Blundell remained a strong tie, but a negative one. This highlights one of the shortcomings of VA, but at the same time demonstrates the iterative nature of its use with the sources.

The other actors highlighted as important are John Gregson, Thomas Earle and Benjamin Heywood.[6] Gregson organised one voyage alone in this period and another with a much larger group, including three members of his family. Earle and Heywood both used a large number of other investors in line with the general pattern of slave traders in Liverpool, including each other, Thomas Parke and William Earle junior in an interlinking pattern of investors. Clearly these merchants used their family ties, but also a wider range of strong and weak ties for their investment activity. Maybe their wealth gave them access to better networks. Certainly they were active in various formal networking institutions at this time as will be demonstrated in section 6.6. However, Boats was also present in the same institutions, and yet managed to appear less well connected to other merchants in this study. Perhaps the fact that Gregson, Earle and Heywood were from wealthy, well-established families, while Boats was notoriously an orphan, counted over and above his skills and knowledge (Wilson, 2008).

This static network analysis of Rainford's networks shows four advantages of using such an approach: it provides a visual representation of his networks as outlined in (some of) the historical records; it identifies the merchant groups and the various subnets in which they invested; it highlights the most well-connected merchants and those who had control over investments and information at this time; and the key facilitators of slave trade voyages connected with Rainford have been highlighted. However, this approach is not without its limitations: the network is still source-centric with Rainford shown as central; linked to this the fact that power relationships are not identified; it does not show change over time; and key actors may be absent. To resolve these issues, qualitative data is still required to help analyse the network. However, this is an explicit point of VA. Using this methodology clearly highlights new avenues for research, even if contextual information about the actors is initially missing.

6.6 Analysing multiple networks using clusters and centrality

As noted above, there are usually multiple networks in action at any one time and they are often interconnected. Using clusters highlights the interconnection between various networks, and centrality measures the relationships between those actors and, indeed, the relationship or relative power between the various subnets.

There are a variety of centrality measures, depending on what the researcher wants to ask of the data as part of the exploratory process. Therefore, no one measurement is ‘correct’. There are four main centrality measures: *out-degree centrality* (how many contacts any actor has access to); *in-degree centrality* (how many people want to contact any actor); *closeness centrality* (how close any actor is to other actors (the least vertices); and *betweeness* centrality (how much control any actor has). Here we are going to demonstrate *betweeness* centrality.[7] These calculations are performed in SocNetV (although other tools such as Gephi also analyse networks using centrality measures). Once the main matrix is formed, separate clusters or partitions are created; in this case the ‘value’ judgement is the association to which they are affiliated. This can be done within Pajek or Visone.

Betweeness centrality measures potential points of control within the network. This measure recognises that communication flow within a given network does not always rely on adjacent vertices (actors), but geodesics (shortest paths between actors). Actors falling on these geodesics are influential in that they act as a chokepoint of information in a network due to their relative network position and their ability to facilitate communications and contacts. In effect, an actor may choose whether or not to share the information they have with other actors within the network. An actor betweeness value falls between zero (as in a circular network) and one (as in a star network). As Wasserman and Faust (1994) note, a path from vertex j to vertex k takes a particular route and it is assumed that all lines have an equal weight and that communications takes the geodesic. It is also assumed that when there are more than one geodesic between j and k, all geodesics are equally likely to be used. If g_{jk} are the number of geodesics linking the two vertices then the probability of using any of them is $1/g_{jk}$. There is a probability that a distinct vertex, i, is involved in the communication between j and k, then the number of geodesics linking j and k and involving i is $g_{jk}(n_i)$. This probability is estimated by $g_{jk}(n_i)/g_{jk}$. The actor betweeness index for

n_i is the sum of these estimated probabilities over all pairs of vertices not including the ith vertex, i.e.

$$C_B(n_i) = \sum_{j<k} g_{jk}(n_i) / g_{jk}$$

for i distinct from j and k. This index is a sum of probabilities with a minimum of zero when n_i falls on no geodesics and maximum $(g-1)(g-2)/2$ which is the number of pairs of vertices not including n_i. This is standardised as:

$$C'_B(n_i) = C_B(n_i)/[(g-1)(g-2)/2]$$

This standardised computation allows for a value between zero and one (Wasserman and Faust, 1994).

This section demonstrates how cluster analysis and betweeness can be used to study a number of interconnecting networks. Here we use Liverpool's metropolitan business networks 1750–1810. Such networks are usually argued to be efficient (Casson, 2003), but to also play an important institutional role, as part of the development of a town and indeed the wider economy (Tullock, 1997). They are also thought to confer status and legitimacy on their members (Shearmur and Klein, 1997). Indeed, many actors enjoy belonging to the institution itself, independently of the business opportunities it may offer (Lawler and Yoon, 1996). It is not certain that institutions were always used for their alleged purpose (Goddard, 2013). They therefore make an interesting point of study.

Here we provide an analysis of the membership records of four institutions: the Town Council; the Committee of Merchants Trading to Africa from Liverpool (African Committee); the Library/Lyceum (Library); and Mock Corporation of Sephton (a drinking club). These represent a variety of formal, trade, cultural and social institutions. The Town Council and the African Committee were also dominated by slave traders and so interlink with our general analysis of Liverpool's slave trade networks. One hundred per cent samples of membership attendance were taken (where the data was extant) for the period 1750–1810. This created a database of a total of 1,700 actors with approximately 210,000 relationships over a sixty-year period (see Haggerty and Haggerty, 2011, for a fuller discussion of this case study and results). Here we discuss two ten-year periods, comparing 1770–9 and 1790–9 using cluster and betweeness centrality. This analysis incorporates 628 actors with 74,000 relationships.

Figure 6.4 shows clearly the levels of cross-associational membership in Liverpool in the 1770s and 1790s. In the 1770s, there was a reasonable degree of cross-associational membership, although no one actor belonged to all four institutions. However, Thomas Rumbold and John Blackburne were members of the Town Council, African Committee and Lyceum, while Richard Hughes belonged to the Town Council, African Committee and the Mock Corporation of Sephton. It is clear, however, that the highest cross-associational relationship was between the Town Council and African Committee. A review of the literature shows that both were dominated by slave traders (Power, 1997; Sanderson, 1977). This

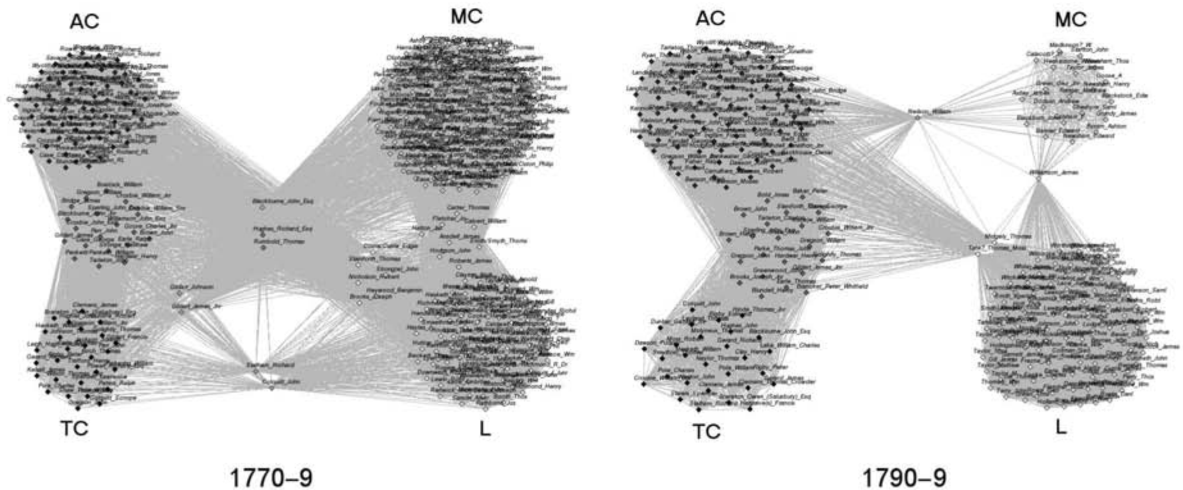

Figure 6.4 Cluster analysis of the actors involved in these institutions during the 1770s and 1790s (source: reprinted from *Explorations in Economic History*, 48, John Haggerty and Sheryllynne Haggerty, 'The Life Cycle of a Metropolitan Business Network: Liverpool 1750–1810', 189–206 (2011), with permission from Elsevier).

Key
AC = African Committee; TC = Town Council; L = Library; MC = Mock Corporation of Sephton.

demonstrates that there was a strong relationship between these two institutions, and may suggest a clique or opaque network. The second highest cross-institutional membership was that between the Library and the Mock Corporation of Sephton, suggesting a strong relationship between the two. This could point to factions or rifts between the associations, the Town Council being dominated by Anglicans, and the Library by non-conformists (Stobart, 2000).

While there were a few actors who were on either both the Lyceum or Mock Corporation and/or the Town Council and African Committee in the 1770s, by the 1790s this relationship had diminished startlingly. By the later period, not one actor in the Library was also on the Town Council. Only two actors, Thomas Midgely and Thomas Tate, belonged to both the African Committee and the Lyceum; only one actor, William Neilson (and possibly also in his partnership Neilson & Heathcote) belonged to the African Committee and the Mock Corporation; and only one actor, James Williamson, belonged to both the Lyceum and Mock Corporation. Even allowing for the fact that the extant data for the Mock Corporation of Sephton ends in 1792, the trend is clear. The cultural and drinking associations were no longer talking to one another, or the more commercial-based ones. This is in stark contrast to the continuing relationship between the Town Council and the African Committee. This comparison clearly shows the changing nature of these relationships over time. This could suggest the further entrenchment of networks and cliques using strong ties rather than weak ties. This means that it was far less likely that information about new opportunities and credit were being spread around the network. It could also suggest that negative social capital was at work here, which could be problematic for both the individual actors involved and the wider trading community. Indeed, it has been suggested that this was the reason for the Liverpool merchants' slow reaction to early attacks on the slave trade in the 1780s (Haggerty and Haggerty, 2011; Sanderson, 1972).

As mentioned above, it is possible to measure the relationships between these actors using betweeness centrality (among others). Figure 6.5 shows the betweeness centrality analysis of the four institutions, comparing the same two decades, 1770–9 and 1790–9.

In the 1770s those actors at the centre highlight the importance of cross-associational membership. Although as a subnet the Town Council/African Committee appears to have the best betweeness, certain individuals had far more control over information. These included Richard Hughes and Thomas Rumbold who were on the Town Council, African Committee and the Library, and John Blackburne, who was on the Town Council, African Committee and the Mock Corporation of Sephton. Johnson Gildart and James Gildart held slightly less strong positions being on the Town Council and Mock Corporation of Sephton, and James Hatton and John Hodgson trailed slightly, being on the Library and the Mock Corporation of Sephton. Clearly, although the Town Council was a relatively important networking institution, those actors who belonged to several institutions had more control, and the drinking and cultural institutions still played an important part in the transference of information. The suggestion above that

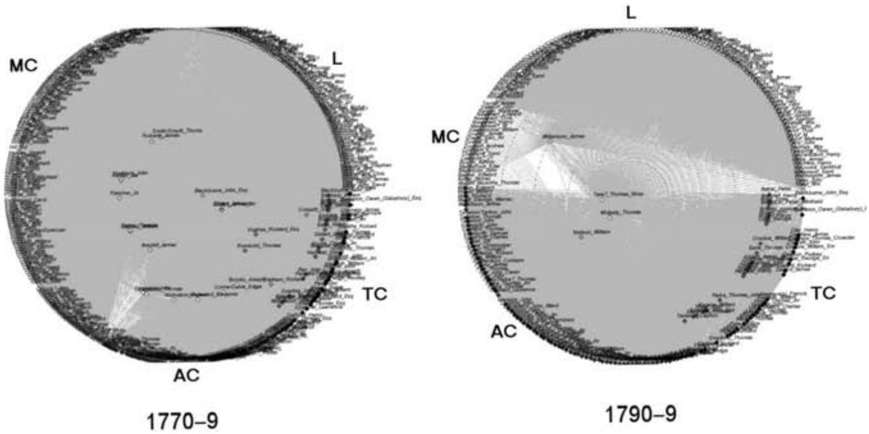

Figure 6.5 Betweeness analysis of Liverpool associational networks 1770–9 and 1790–9 (source: reprinted from *Explorations in Economic History*, 48, John Haggerty and Sheryllynne Haggerty, 'The Life Cycle of a Metropolitan Business Network: Liverpool 1750–1810', 189–206 (2011), with permission from Elsevier).

Key
AC, African Committee; TC, Town Council; L, Library; MC; Mock Corporation of Sephton.

less cross-institutional relationships in the 1790s may be problematic is confirmed by the betweeness analysis. Even fewer people had any control (are nearer to the centre). Thomas Midgley and Thomas Tate are shown as important actors, belonging to the African Committee and the Library as is William Neilson, who was on the African Committee and the Mock Corporation of Sephton. James Williamson, on the Mock Corporation of Sephton and the Library, came a poor fourth. Therefore, the use of cultural and drinking institutions was not only in decline by the elite mercantile group on the Town Council and African Committee, but all these institutions were in general becoming more isolated. Figure 6.5 shows that not only was cross-institutional membership in decline, but that this dramatically reduced the ability of the number of individual actors to exert control within the networks, and indeed the strength of that control. Therefore, although the Town Council/African Committee as a subnet appears to have had more control (it has moved inwards), in fact the general lack of cross-institutional membership was harmful overall. As mentioned above, negative social capital could account for a failure to react to exogenous events counter to Liverpool's interest in the slave trade such as Dolben's Bill in 1788 (Sanderson, 1972). It is certainly hard to explain the abolitionist William Roscoe's election as an MP in 1806 (Wilson, 2008). While it has therefore been argued that these bonding networks were good for a few of the actors investing heavily in the slave trade (McDade, 2011), it is likely that they were harmful for the remainder of Liverpool's slave trading community.

Using clusters and betweeness to analyse Liverpool's multiple associational membership demonstrates three advantages of such an approach: the extent to which individuals are involved in cross-associational membership are highlighted; those individuals and subnets that had the most control (betweeness) over information are clearly shown; and importantly, comparing clusters and betweeness over time demonstrates the changing relationships within and between these networks. This forces the researcher to question why certain networks were being joined and used at particular times, in the short and long term, and to think about the effect of these changing relationships for not only the individuals involved, but also the wider community. As with static network analysis, this does not provide the whole story, nor is it meant to. Ideally this analysis should be followed up in the minutes of relevant associations and the personal manuscripts of the key actors, where they are extant. Once again, this analysis identifies the key actors and subnets, and thereby highlights new avenues for research.

6.7 Conclusion

Using VA for large data sets is an explicitly exploratory exercise. VA is not meant to provide answers, but to work in an iterative way with the researcher and his or her sources, in order to promote a more nuanced analysis of networks. Here two techniques for the analysis of large actor networks have been highlighted, but VA can visualise and measure data in far more ways, and for a wide variety of purposes.

The case study of Liverpool's slave trade and related associational networks has shown that these techniques raise important questions about a researcher's networks. How and why are they the way they are? Who is important and who is not? How and why do people use their networks – especially institutional ones? How and why do they change over time? These questions remind us that we cannot take actor networks as a simple positive economic good. Historians need to question what their networks are for in the first place and, indeed, what constitutes a network. VA provides enormous help in providing some initial answers to these questions, but more importantly, in raising new questions of the researcher's data in turn.

Using VA can be seen as an important development in an increasingly interdisciplinary environment, especially where the analysis of actor networks is concerned. VA incorporates a number of tools which require a varying degree of technical ability. Simple static network visualisations as in section 6.5 above are easy to use and a simple introduction to this methodology. The historian only needs to format their data into a simple matrix or formatted text file in order to start using it. Furthermore, there is now a variety of large data sets available online which could also be imported for use with VA. Here we used the slave trade database online (at www.slavevoyages.org), but there are others available which could easily be used. This could include The London Lives Project, (www.londonlives.org/index.jsp), William Godwin's Diaries (http://godwindiary.

bodleian.ox.ac.uk/index2.html), The Humphrey Davy Letters (www.davy-letters.org.uk/), and so on. VA therefore provides exciting and fun visualisations for both small and large data sets. More significantly, it provides an important iterative and explicitly exploratory tool for the research of actor networks.

Notes

1 Some of this software requires more technical ability than others. An easy starting point however for basic visualisations is Pajek or Visone.
2 This could possibly be overlaid on a map using GIS.
3 For a fuller analysis of the methodology and findings related to Rainford's networks see Haggerty and Haggerty, 2010.
4 This pattern was not evident in Bristol (McDade, 2011).
5 As yet no extant manuscripts of Boats' mercantile or private life have been found with which to flesh out his personality and other business dealings.
6 It should be noted that only their voyages connected with Rainford are highlighted here. These actors were involved in many other voyages during this period, and both before and after it.
7 For a fuller discussion of all these measures and their application to Liverpool's institutional networks see Haggerty and Haggerty, 2011.

References

Note: Primary sources are excluded as they are all cited in the authors' papers listed below.

Behrendt, Stephen D. (2007) Human capital in the British slave trade, in David Richardson, Suzanne Schwarz and Anthony Tibbles (eds), *Liverpool and Transatlantic Slavery*, Liverpool: Liverpool University Press, pp. 66–98.

Bonacich, Phillip (1987) Power and centrality: A family of measures, *American Journal of Sociology*, 92(2), March, 1170–82.

Bourdieu, Pierre (2001) The forms of capital, in Mark Granovetter and Richard Swedberg (eds), *The Sociology of Economic Lif*e, 2nd edn, Cambridge, MA: Westview Press, pp. 96–111.

British Parliamentary Papers, Census Reports: Abstracts of the Answers and Returns, Enumeration [1801] (London: Cass, 1968).

Burt, Ronald D. (2004) Structural holes and good ideas, *American Journal of Sociology*, 10(2), 349–99.

Casson, Mark C. (2003) An economic approach to regional business networks, in John F. Wilson and Andrew Popp (eds), *Industrial Clusters and Regional Business Networks in England 1750–1970*, Aldershot: Ashgate, pp. 19–43.

Crumplin, Tim E. (2007) Opaque networks: Business and community in the Isle of Man, 1840–1900, *Business History*, 49(6), November, 780–801.

Davidoff, Leonore and Catharine Hall (1987) *Family Fortunes: Men and Women of the English Middle Class 1780–1850*, London: Hutchinson.

Doerflinger, Thomas M. (1986) *A Vigorous Spirit of Enterprise: Merchants and Economic Development in Revolutionary Philadelphia*, Williamsburg, VA: University of North Carolina Press.

Duguid, Paul (2004) Networks and knowledge: The beginning and end of the port commodity chain, 1703–1860, *Business History*, 79, Autumn, 492–526.

Freeman, Linton C. (1978–9) Centrality in social networks: Conceptual clarification, *Social Networks*, 1, 215–39.

Goddard, Richard (2013) Medieval business networks: St Mary's Guild and the Borough Court in Later Medieval Nottingham, *Urban History*, 40(1), February, 3–27.

Granovetter, Mark S. (1973) The strength of weak ties, *American Journal of Sociology*, 78(6), May, 1360–80.

Haggerty, John and Sheryllynne Haggerty (2010) Visual Analytics of an eighteenth-century network, *Enterprise and Society*, 11(1), 1–25.

Haggerty, John and Sheryllynne Haggerty (2011) The life cycle of a metropolitan business network: Liverpool 1750–1810, *Explorations in Economic History*, 48(2), 189–206.

Haggerty, John and Sheryllynne Haggerty (2012) Investigating networks over time: Matrixify, Digital Humantities Congress, Sheffield, 2012, available at www.sheffield.ac.uk/hri/dhcpres.

Haggerty, Sheryllynne (2011) 'You promise well and perform as badly': The failure of the 'implicit contract of family' in the Scottish Atlantic, *International Journal of Maritime History*, 23(2), 267–82.

Haggerty, Sheryllynne (2012) *'Merely for Money'? Business Culture in the British Atlantic, 1750–1815*, Liverpool: Liverpool University Press.

Hancock, David (1995) *Citizens of the World: London Merchants and the Integration of the British Atlantic Community, 1735–1785*, Cambridge: Cambridge University Press.

Hancock, David (2005) The trouble with networks: Managing the Scots' early-modern Madeira trade, *Business History Review*, 79, Special Edition on Networks in the Trade in Alcohol, Autumn, 467–91.

Lawler, Edward J. and Yeongkoo Yoon (1996) Commitment in exchange relations: Test of a theory of relational cohesion, *American Sociological Review*, 61(1), February, 89–108.

Lawton, R. (1953) Genesis of population, in W. Smith (ed.), *A Scientific Survey of Merseyside*, Liverpool: University of Liverpool Press for the British Association, pp. 120–31.

Marriner, S. (1982) *The Economic and Social Development of Merseyside, 1750–1960*, London: Croom Helm.

Mathias, Peter (2000) Risk, credit and kinship in early modern enterprise, in John J. McCusker and Kenneth Morgan (eds), *The Early-Modern Atlantic Economy*, Cambridge: Cambridge University Press, pp. 15–35.

McDade, Katie (2011) Bristol and Liverpool Merchants' Slave Trade Networks, Unpublished PhD thesis, University of Nottingham.

Morgan, Kenneth (1993) *Bristol & the Atlantic Trade in the Eighteenth Century*, Cambridge: Cambridge University Press.

Ogilvie, Sheilagh (2004) Guilds, efficiency, and social capital: Evidence from German proto-industry, *Economic History Review*, 57(2), 286–333.

Perer, A. and B. Schneiderman (2009) Integrating statistics and visualisations for exploratory power: From long-term case studies to design guidelines, *IEEE Computer Graphics and Applications*, May/June, 39–51.

Podolny, Joel M. and Karen L. Page (1988) Network forms of organization, *Annual Review of Sociology*, 24, 57–76.

Popp, Andrew (2007) Building the market: John Shaw of Wolverhampton and commercial travelling in early nineteenth-century England, *Business History*, 49(3), May, 321–47.

Portes, Alejandro (1998) Social capital: Its origins and applications in modern sociology, *Annual Review of Sociology*, 24, 1–24.

Power, M. J. (1997) Councillors and commerce in Liverpool, 1650–1750, *Urban History*, 24(3), 301–23.

Price, Jacob M. (1974) Economic function and the growth of American port towns in the eighteenth century, *Perspectives in American History*, 8, 121–86.

Prior, Ann, and Maurice Kirby (1993) The Society of Friends and the family firm, *Business History*, 35(4), 66–85.

Rauch, James E. (2001) Business and social networks in international trade, *Journal of Economic Literature*, 39, December, 1177–1203.

Renzulli, Linda A., Howard Aldrich and James Moody (2000) Family matters: Gender, networks, and entrepreneurial outcomes, *Social Forces*, 79(2), December, 523–46.

Rosenband, L. N. (1999) Social capital in the early Industrial Revolution, *Journal of Interdisciplinary History*, 29(3), 435–57.

Sanderson, F. E. (1972) The structure of politics in Liverpool 1780–1807, *Transactions of the Historic Society of Lancashire and Cheshire*, 127, 65–88.

Shearmur, Jeremy and Daniel B. Klein (1997) Good conduct in the Great Society: Adam Smith and the role of reputation, in Daniel B. Klein (ed.), *Reputation: Studies in the Voluntary Elicitation of Good Conduct*, Ann Arbor: University of Michigan Press, pp. 29–46.

Smith-Doerr, Laurel and Walter W. Powell (2005) Networks and economic life, in Neil J. Smelser and Richard Swedberg (eds), *The Handbook of Economic Sociology*, second edition, New York and Oxford: Princeton University Press, pp. 379–402.

Stobart, J. (2000) Culture versus commerce: Societies and spaces for elites in eighteenth-century Liverpool, *Journal of Historical Geography*, 28(4), 471–85.

Swann, G. M. P. and M. Prevezer (1998) Introduction, in G. M. P. Swann, M. Prevezer and D. Stout (eds), *The Dynamics of Industrial Clustering: International Comparisons in Computing and Biotechnology*, Oxford: Oxford University Press, pp. 1–12.

Tilly, Charles (1995) Transplanted networks, in Virginia Yans-McLaughlin (ed.), *Immigration Reconsidered: History, Sociology, and Politics*, Oxford: Oxford University Press, pp. 79–95.

Trivellato, Francesca (2009) *The Familiarity of Strangers: The Sephardic Diaspora, Livorno, and Cross Cultural Trade in the Early Modern Period*, New Haven, CN: Yale University Press.

Tullock, Gordon (1997) Adam Smith and the Prisoners' Dilemma, in Daniel B. Klein (ed.), *Reputation: Studies in the Voluntary Elicitation of Good Conduct*, Ann Arbor: University of Michigan Press, pp. 21–28.

Walvin, James (1997) *The Quakers: Money and Morals*, London: John Murray.

Wasserman, S. and K. Faust (1994) *Social Network Analysis Methods and Applications*, Cambridge: Cambridge University Press, pp. 189–191.

Wilson, A. (2008) *William Roscoe: Commerce and Culture*, Liverpool: Liverpool University Press.

Websites

Gephi (2013), https://gephi.org/, accessed 2 April 2013.

Pajek (2013), http://pajek.imfm.si/doku.php?id=pajek, accessed 2 April 2013.

Gretl, Cottrell, A. and R. Lucchetti (2013), GRETL, http://gretl.sourceforge.ne/, accessed 2 April 2013.

SocNetV, Kalamaras, D. V. (2013), SocNetV, http://socnetv.sourceforge.net/index.html, accessed 2 April 2013.

Transatlantic Slave Trade Database, www.slavevoyages.org/tast/index.faces, accessed 2 April 2013.

Visone (2013), http://visone.info/, accessed 2 April 2013.

7 Railways and local population growth

Northamptonshire and Rutland, 1801–91

Mark Casson, A.E.M. Satchell, Leigh Shaw-Taylor and E.A. Wrigley

7.1 Introduction

It is widely believed that the arrival of a railway in a town or village will boost local population. Local historical case studies provide numerous instances of how the arrival of railway stimulated local industry and population. On the other hand, there are instances where the arrival of a railway seems to have had no effect on the local community whatsoever. Railway promoters saw no need to build a local station where they perceived little traffic potential, and this expectation was often self-fulfilling. Indeed, records suggest that in nineteenth-century England and Wales railway investments sometimes reduced population rather than increased it – small towns served by a railway went into decline as trade and production switched to larger centres at either end of the line; opening a station made it easier for local trade to move away from the town. Local conditions were therefore an important mediating influence on the impacts of railways.

The connection between railways and population growth can be examined through either railway openings or railway closures. Gregory and Henneberg (2010) have used a GIS approach to analyse the impact of railway construction on local population growth, while Patmore (1966) and Kennedy (2013) have analysed the impact of closures on population decline and social deprivation. This chapter focuses on railway openings and adopts a statistical approach grounded in economic theory.

The investigation focuses on the towns and villages of the South-Midlands English counties of Northamptonshire and Rutland in the nineteenth century. The basic unit of analysis is the parish; 291 parishes in Northamptonshire and sixty-one in Rutland. The study examines how far differential parish population growth can be explained by differential access to railways and how far it was due to other factors. The key data is decadal estimates of parish population 1801–91 derived from the Census of Population, supplemented by information from a variety of other sources described below.

The chapter develops and tests an equilibrium theory of population change. At any given time, it asserts, there is an equilibrium distribution of population across the parishes which reflects economic and social conditions prevailing at the time. The distribution is expressed statistically as a variation across parishes in the density of population – i.e. population per acre. The equilibrium population depends upon a number of parish characteristics, some of which, like access to rivers and type of soil, are constant over time, and others, like access to railways, vary over time. As the time–dependent factors vary so the equilibrium distribution of population changes, and this is reflected in the differential growth of parish populations. Parishes that benefit economically from changes tend to grow while those that suffer tend to decline.

There are problems with this approach, however. The first concerns the direction of causation. Just as population adjusts to railway building, so railway building may adjust to population change. If the statistics show that local population increased at the time a railway was constructed, the explanation could be either that the building of the railway stimulated population growth or that railway promoters were targeting areas of high population growth in which to build a railway. This issue can be addressed by recognising the importance of lags: if railway building stimulates population growth then population growth will generally follow later, whereas if population growth attracts railways then population growth will generally occur earlier.

Another problem is that proximity to a railway line does not guarantee convenient access to a station. If access to a station is more important for sustained population growth than mere proximity to a railway then station-building needs to be considered explicitly as a factor in population growth. However, while a railway can be built without a station, building a station without a railway usually makes little sense; the relation between railway building and station building therefore needs to be analysed with care if misleading inferences are to be avoided.

7.2 General principles of modelling population change

Census data allocates people to place by residence rather than by place of work. Economic forces, however, often impinge most directly on places of work. It is therefore necessary to analyse the relationship between place of residence and place of work.

Places of work are influenced by natural resource endowments, such as soils that are conducive to growing crops in popular demand, or mineral deposits useful to manufacturing industry. If the prices of certain crops increase relative to others then land may switch out of one form of cultivation into another. This will change local demands for agricultural labour, which in turn may affect residential patterns.

Cathedrals, abbeys and county courts are important sources of employment in counties where alternatives to agricultural employment are few. The professionals employed in such institutions constitute a high-status elite whose consumption may provide a range of work for local craft and service workers. Country houses are another important source of labour demand, e.g. landscape gardening, hunting and shooting, housing art collections and establishing model farms. In nineteenth-century Northamptonshire shoe-making was an important industry, involving both factory work and out-work.

Competition in the land market will allocate land between alternative uses, including different types of work (agriculture, manufacturing, services, etc.), different types of leisure activities (gardening, hunting, etc.) and different types of residence (houses, flats, workhouses, etc.). This suggests that at any given time there will be an equilibrium in the allocation of land to alternative uses.

If residential accommodation is scarce in one village, e.g. a 'closed' village where a dominant local landowner is opposed to housing development, then workers and their families may live in a neighbouring 'open' village instead. Differences between parishes in the implementation of poor relief may encourage marginal groups – the unemployed, gypsies, vagrants, etc. – to agglomerate in certain parishes rather than others. In particular, union workhouses and asylums may be located in relatively poor and remote areas where land is cheap.

Birth and death rates also need to be taken into account in analysing population change, but their effects on the spatial distribution of population may be small. A general increase in births and decline in deaths may, on average, simply raise population in each parish by the same proportion. Localised differentials in births and deaths may be accommodated by migration: e.g. rural areas with a high birth rate may exhibit substantial out-migration by young people, while genteel spa towns that attract older people will have substantial in-migration to replace residents who die. Life expectancy may also differ, e.g. being higher in rural areas than in large industrial areas.

Competition between market towns appears to have been a major factor in differential population growth. In the nineteenth century town growth provided opportunities for supplying urban dairies with milk, butchers with meat, brewers with hops and grain, market traders with fresh vegetables, and so on. Many market centres in England emerged between the late twelfth and early fourteenth centuries, and some never became more than villages or hamlets linked to manorial estates. Those that developed initially were subsequently subjected to periodic 'shake-outs' during which markets became concentrated on fewer larger centres. Early shake-outs were precipitated by population decline – e.g. the Black Death. Later shake-outs tended to be driven by advances in transport, such as canals and turnpikes, which encouraged producers to send their product to more distant larger markets than to nearby smaller ones. This study investigates whether railway investment sustained the shake-out process in the nineteenth century.

Railway building reduces freight charges and facilitates the export of local goods and services, thereby encouraging the specialisation of production at locations served by a railway. In rural areas railway access facilitates agricultural exports to feed the growing populations of industrial cities, and mineral exports to feed their factories, and also widens access to imported fertilisers (e.g. lime) and fuels (e.g. coal).

If population is reasonably mobile, workers will be attracted to locations served by railways because of the new jobs created there. Railways can also stimulate labour mobility by facilitating the removal of workers and their families to new locations, and by encouraging commuting; railways were not built specifically for commuting by ordinary workers until the 1870s, however.

In Northamptonshire and Rutland there were significant differences between towns in the dates at which railways arrived, and so if these effects are important it should be easy to capture them. The benefits of connection were, in principle, greatest in the 1840s and 1850s, when railway access was a privilege enjoyed by relatively few towns. The competitive impact of transport improvements on market shake-out could be mitigated by general population growth, however, as this could allow smaller markets to maintain a critical size.

7.3 Outline of an econometric model

The interaction of these effects can be captured by a formal model, which is set out algebraically in the appendix, where its statistical properties are also discussed. The basic hypothesis is that in each parish the equilibrium stock of population is related to the stock of land and the stock of infrastructure built upon the land at that time. Both land and infrastructure are regarded as heterogeneous.

The quantity of land is measured by the acreage of the parish, which is constant over time (see below). The quality of the land is reflected in two key geographical characteristics:

- Location:
    - access to rivers and
    - distance from important administrative centres such as the county town of Northampton and the cathedral city of Peterborough; and
- Geology, namely the type of soil.

Infrastructure encompasses:

- Transport infrastructure, comprising roads, canals and railways; and
- Institutional infrastructure, relating to administrative status and market privileges, namely whether the town was a borough, and possessed a market in 1801.

The quantity and quality of land, and the endowment of infrastructure, are summarised by a set of key parish characteristics. Some of these characteristics are constant (i.e. time-invariant) and others are variable (i.e. time-dependent).

Locational factors are mainly constant, especially those that reflect natural features. Soil fertility can change, however, as a result of changes in farming techniques (e.g. application of fertiliser and investment in drainage schemes). Rivers may be rendered navigable by cuts, and the value of non-navigable rivers may be influenced by the construction of mill streams or fish weirs. Such changes are accounted for, not by changing the location factors, but by allowing for constant location factors to have changing impacts, as explained below.

Distance to a town is measured as the crow flies; it could also be measured as distance by road, although this can change over time, and may shrink systematically if new roads are built as towns expand.

Infrastructure investment in the nineteenth century mainly involved railway building; road improvements through turnpikes were largely completed by the beginning of the century. For this reason road infrastructure is treated as constant, being given by the legacy inherited in 1801. Northamptonshire and Rutland were 'late developers' so far as canals were concerned. Some sections of canal were not completed until after 1801, but as all the canals were authorised by Parliament prior to 1801 they are treated as exogenous.

Institutional infrastructure is determined entirely by legacy effects: a town is classified as a borough or market centre depending on its position in 1801. For statistical purposes this means that all parish characteristics other than those connected with railway access are treated as exogenous: they are not influenced by population change during the period of the study. Railway investment, by contrast, is treated as endogenous, as explained below.

While the quality of land may be constant over the period of the study, the impact of quality of land on population is liable to change, as noted above. This means that constant characteristics may have variable impacts. This applies particularly to the impact of the type of soil. If incomes rise nationally, for example, consumers may substitute meat for bread, so that the price of beef rises and the price of wheat falls. Grazing becomes more profitable and grain production less profitable; as a result, population will tend to decrease in areas with sand and clay soils, suitable for arable cultivation, and increase in areas with gravel and alluvium, suitable for grazing.

The same point applies to other characteristics to some degree, e.g. the importance of inherited market status may change as the economy industrialises. The model therefore allows for the impacts of all characteristics to change over time. Thus different vintages of railway investment may have different impacts: the impact of a railway on local population may be high during the period of construction (due to an influx of labourers), only moderate immediately after opening (due to the arrival of railways staff) and more substantial again later on (due to an influx of new businesses or new residents).

Railway investment is regarded as endogenous, because it may respond to population growth. Thus while population responds to railway investment, railway investment responds to population growth. This two-way interaction is more significant for towns and cities than it is for villages. Railway lines were normally promoted to link two large centres of population by the most direct route, and were rarely diverted to serve small towns and large villages along the route. Thus whether a village was served by a railway was largely determined by the populations of the towns on either side rather than by the population of the village itself (subject to the qualifications noted below). Nevertheless two-way causation, when it occurs, creates potential problems of interpretation, as noted earlier. The problem is resolved by assuming that, while population responds immediately to railway investment, railway investment responds to population only with a lag. As explained in the appendix, this asymmetric lag structure makes it possible to identify the two effects separately.

There is a lag in the response of population to railway investment because:

- Railway construction is a lengthy process, and is preceded by an even longer process – namely raising finance and gaining Parliamentary approval. At the time this process begins, the current population of the parish will not be known; only the data from the previous census will be available. Although railway promoters will attempt to anticipate future population, they do not have perfect foresight: they cannot know future population but must estimate it from previous population or from past trends in population growth. Thus while population may adjust almost instantaneously to railway building, railway building will adjust to population increase only with a lag.
- Many railway schemes promoted during the Railway Mania 1844–6 were never built; these un-built schemes formed the basis for another set of schemes which were promoted and built during the 1860s. In general, un-built schemes often lay dormant before being revived later; e.g. un-built schemes promoted in the 1860s sometimes re-emerged as Light Railway schemes in the 1890s. This suggests that if railway building does not adjust to population growth with a single-decade lag then it adjusts with an even longer lag.

On the other hand, population responds immediately to railway investment because:

- Once it is known that a railway is to be built, people may move to the parish while land is still relatively cheap, hoping to profit from a rise in land values once the railway is built.
- Construction workers may move into the village, followed by station staff and gangers (to maintain the track after the line has been opened).

Railway investment may be disaggregated into different types of lines, because the motives for building different types of line may differ and so their impacts may differ too. Three types of railway are distinguished: trunk lines, cross-country lines and local lines.

- Trunk lines tend to connect large cities by the shortest or fastest route; most of the earliest trunk lines in England and Wales connected an industrial centre or port city to London. The large volumes of traffic carried by trunk lines suggest that their impacts on the cities they serve will be large, but if trains run non-stop then the impact on intermediate villages may be small.
- Cross-country lines developed to make direct connections between provincial centres and cut out the need for connections through London. They also served intermediate towns that had been bypassed by the trunk network. Cross-country lines can also act as feeders to trunk lines and their junctions may develop into important railway hubs.

- The construction of trunk lines and cross-country lines may in turn trigger the building of local lines. Population density is more likely to influence the route followed by a local line, and thereby determine which particular villages are served. The building of one local line may trigger the development of other local lines and thereby create a local system which is to some extent independent of the long-distance network (e.g. a suburban system). Network externalities of this kind suggest that the presence of one railway in a parish may encourage the construction of other railways serving the same parish, so that the parish then becomes a local hub.

The model distinguishes the provision of a local station from the provision of tracks. While a station normally requires a railway to serve it, it is unnecessary to build a station in every parish through which a railway passes. Building a station incurs a fixed cost of land and construction, and a recurrent cost of station staff. In addition, there is an operational cost of stopping and starting trains at intermediate stations; it lengthens journey times and can delay other traffic along the route. Local population density must be sufficiently high to create a critical mass of local traffic. Railway access can be important to the local economy independently of station access: freight sidings can be laid in to serve factories built alongside a line, and short spurs built to mines opened up nearby; indeed some mineral branch lines in Northamptonshire were built without any stations at all.

High-speed running on trunk lines, coupled with reliance on inter-city traffic, may discourage the building of intermediate stations on trunk lines, further reducing their impact on local communities. On local lines, however, low speed, coupled with reliance on local traffic, may encourage station building. It is therefore possible that local lines may be of greater benefit to local communities than trunk lines, because although trunk lines, in principle, offer immediate access to large cities, their owners may have little interest in developing local traffic. Local lines, on the hand, may be focussed on developing local traffic.

Although station access is normally provided only in parishes which are served by a railway, it is possible that a parish may be served by a station in a neighbouring parish. It is assumed that the decision to build a station is made at the same time, or somewhat later, than the decision to build the railway, so that station building, like railway building, responds to population growth with a lag. This makes it possible to distinguish statistically between the impact of stations on local population and the impact of local population on station provision.

Railway building and station building depend, in principle, not only on lagged population but on all exogenous parish characteristics, and the model allows for this; it is assumed, however, that the impacts of different characteristics on railway and station building are additive, in the same way as for population growth

Overall, therefore, the model generates five linear equations involving five dependent variables: population growth, three different types of railway building, and station building. All five variables depend linearly on constant parish characteristics and lagged parish population. Population growth depends on current and past railway building and station building, and station building depends on current and past railway building. Railway building is independent of current station building, and depends only on previous population growth. Because the equations have a recursive structure they can be estimated independently, using ordinary least squares, logit or probit regressions as appropriate.

7.4 Data sources and their limitations

This study is the second in a series of country studies examining the impact of railways on local population growth. It refines the template developed in the first study, based on Oxfordshire (Casson, 2013), and is intended as a model for subsequent studies. For this reason the sources have been selected partly on the grounds that they provide comprehensive coverage of England and Wales and are not purely specific to Northamptonshire and Rutland.

The parish is the smallest administrative unit on which good quality official population data is consistently available (information on constituent townships, hamlets and chapelries can occasionally be obtained, and some is used in the present study). The parish is an ancient administrative unit, and so a great deal is known about the history of many individual parishes. The main limitations of parish data for a study of this kind are:

- Parishes are variable in size. Some rural parishes are very large, while some urban parishes are very small. In the case of large rural parishes, the match between the extent of the parish and the pattern of settlement can be poor. The population may be clustered in one small part of the parish; there may, indeed, be several distinct settlements, none of which is centrally located. In the case of a small urban parish, the settlement may form part of an agglomeration and may span adjacent parishes. This issue is addressed by allowing all parishes in the same town to share the same characteristics, e.g. market status and railway access, independently of the particular parish in which the relevant facility, e.g. market or station, is located.
- Parishes may be irregular in shape; thus a small parish may be long and thin, so that a railway passing through one corner is remote from people residing at the opposite corner. Because river access was so useful in medieval times, parishes often run down the sides of valleys to rivers at the bottom. Some parishes even have detached portions. As a result, the type of soil, and hence the type of agriculture, may vary considerably within a parish. This means that patterns of agricultural specialisation may be difficult to detect from parish data.
- Some parishes are not only small, but have low population density, so that total population is very small. This means that population growth can be volatile: the migration of a single family can have a significant proportionate impact on overall population.

Overall, therefore, parishes can be very heterogeneous; spatial variations may be partially masked by the fact that they are internal to parishes, and therefore do not show up as inter-parish variations. A more accurate representation of spatial variation could be obtained using standard space-filling shapes such as squares or hexagons, or, even better, by using dynamic GIS in which the coordinates of parish settlements could be used, but these methods lie outside the scope of the present study.

The population data used in this study is derived from the published Census of Population, and has been standardised to adjust for changes in parish boundaries during the nineteenth century. The main method of adjustment involves consolidating the data for parishes whose boundaries changed so that boundary changes are internalised within enlarged population units. The precise methods for calculating the population and the area of a parish (or equivalent population unit) are described in Wrigley (2011); the implications of adjusting for boundary changes in a study of this type are discussed in Casson (2013).

Because county boundaries have changed, and adjacent parishes have been grouped together for analytical convenience, not all the parishes in this study belong to the modern counties of Northamptonshire and Rutland, and some parishes in these counties are not included because they have been allocated to other counties instead.

Growth rates of population are calculated for each parish in each decade. Growth is defined in proportional and not absolute terms. Growth rates are estimated in terms of differences in the logarithm of population; they are not annualised, since this is unnecessary for the statistical analysis.

As explained above, each parish has a profile with five main dimensions.

Administrative status. Parishes are classified by whether they were boroughs, and by whether they had active or defunct markets. The main source was the Gazetteer of Markets and Fairs (Letters, 2010), which synthesises information from other sources such as Beresford and Finberg (1973) and its supplement. Others useful sources for the nineteenth-century status of markets are statistical and topographical gazetteers by Bartholomew (1887), Lewis (1846) and Wilson (1870–2), excerpts from some of which appear in Vision of Britain (2013). Thirteen ancient boroughs were identified: Brackley, Daventry Finedon, Higham Ferrers, Northampton, Oundle, Peterborough, Rockingham, Rothwell, Thornby and Towcester in Northamptonshire, Oakham in Rutland and Stony Stratford just across the Buckinghamshire border. All of the boroughs except Finedon, Rockingham and Rothwell had active markets around 1800; other active markets were in Kettering, Thrapston and Wellingborough in Northamptonshire and Uppingham in Rutland. Markets may be either chartered, authorised by letter close or purely prescriptive. In addition, there were thirty-six defunct markets in Northamptonshire and five in Rutland (see Table 7.1). Many of these seem to reflect the economic and social ambitions of lords of the manor in the South Midlands during the thirteenth and fourteenth centuries. Most of the markets are identified through charters, although some are prescriptive (for further information see Partida *et al*., 2013).

Distance from major centres. Two major centres are identified: Northampton, the county town and main administrative centre, and the cathedral city of

Table 7.1 List of defunct markets

Northamptonshire				
Alderton	Aynho	Barnwell All Saints	Barnwell St. Andrew	Boughton
Brigstock	Brixworth	Bulwick	Catesby	Chipping Warden
Corby	Culworth	Fawsley	Finedon	Floore
Fotheringhay	Geddington	Grafton Regis	Harringworth	King's Cliffe
Long Buckby	Lowick	Milton	Naseby	Northborough
Rockingham	Thornby	Thorpe Mandeville	Thurning	Titchmarsh
Wakerley	Welton	West Haddon	Wilby	Wollaston
Yardley Hastings				
Rutland				
Barrowden	Belton	Empingham	Liddington	Market Overton

Peterborough (and later a major railway hub). Parishes are classified by whether the nearest part of the parish lay within five miles of each urban centre. Other salient places from which distance could be measured include London, Coventry and Leicester, but access to these is largely captured by access to relevant roads, as described below.

Type of soil. The Geology of Britain Viewer (British Geological Survey, 2012) is the main source of information for soil types. Based on their maps, nine types of soil were identified as relevant to the South Midlands, involving both subsoil and surface soil: limestone, ironstone, brash, sand, gravel, alluvium, clay, flint and mudstone. By and large, sand and limestone provide light soils while clay and mudstone provide heavy soils. Gravel and alluvium are associated with river valleys. Rock formation in the South Midlands follows complex and varied patterns, and as a consequence many parishes possess four or five different types of soil. Where ancient parishes could not be identified from the modern names used on the Viewer, Google Maps (Google, 2013) were used to locate defunct parishes preserved in the names of farms or roads. The Victoria County History (VCH) also profiles soils, but often superficially; in addition, its coverage of Northamptonshire parishes is incomplete (Serjeantson and Adkins, 1906). A more refined analysis of soil would take account of permeability (for drainage) and slope (which affects the feasibility of arable farming), and it is hoped to incorporate these factors in a later study.

Access to rivers, canals and roads. Access to water is significant for transport, mechanical power, fisheries and for meadows where animals graze. A parish is deemed to have access to water if a major river runs through the parish or (more often) forms part of its boundary. Three rivers are identified as significant: Nene, Ise and Welland. Two portions of the Nene are distinguished; the lower part from Peterborough to Northampton, suitable for large boats, and the upper part from Northampton to the border with Warwickshire. The Ise is a tributary of the Nene, joining it at Wellingborough. The Welland flows through the north of the county from Stamford in Lincolnshire towards Market Harborough in Leicestershire. In addition, the Great Ouse flows near the southern boundary of the county, but the parishes bordering it lie mainly in Buckinghamshire or Bedfordshire. Minor rivers, such as the Tove, are omitted, as they are not readily navigable and serve a relatively small number of parishes.

Northamptonshire is also served by the Grand Junction Canal from London to Birmingham, which runs through the south-west of the county between Cosgrove, near Newport Pagnell to Braunston near Daventry, where a branch is thrown off to Leicester. This branch lies almost entirely in Leicestershire and so is not included in this study. There is also a branch from Blisworth to Northampton that connects with the River Nene. The Grand Junction Canal was completed throughout in 1800, with the exception of the Blisworth tunnel, which was not opened until 1805; in the meantime, the two portions of the canal were connected by a tram-road. The Oakham Canal from Melton Mowbray to Oakham was authorised in 1793, opened in 1802 and closed in 1846, shortly after the arrival of a competing railway line. The canal linked Leicestershire to Rutland, and traversed only six parishes in Rutland. The parish of Blisworth and these six Rutland parishes are too small in number to be identified separately in this study; but instead of being excluded, they are treated as if they were served by a canal in 1801, a year before the opening of the Oakham canal, and four years before the completion of the tunnel.

Many parishes are served by the rivers because a river usually forms a boundary between parishes on the two banks, and on each bank parishes with centres of population remote from the bank throw out thin strips of land towards the river. The same does not apply to the canal, which follows the Upper Nene for some of its length but always within a parish to one side of the river or the other.

Access to roads was determined by whether the parish was served by a mail coach or stage coach service that ran on a daily basis in 1836 (Bates, 1969). This is the earliest date for which a comprehensive national source on coaching is available. Only parishes that lay directly on the route, or which included public houses and inns on the route, were included. Parishes off to the side of the route were not included. The aim of the measure was to indicate whether the parish benefited from economic activity derived directly from road traffic, rather than to indicate whether residents were conveniently located for access to the road system. An alternative approach to measuring road access would have to assess which parishes were served by turnpikes. However, turnpikes were very common by the end of the eighteenth century, and so access to a turnpike does not discriminate between parishes very effectively. In addition, a section of turnpiked road that is not connected with other turnpiked sections for onward travel may have been of limited significance for the development of through traffic. Conversely, an unturnpiked section of road within a turnpiked route might attract quite a lot of traffic.

Five routes were identified: Watling Street, running from London towards Coventry through Stony Stratford, Towcester and Daventry; a route from London to Leicester through Newport Pagnell, Northampton and Market Harborough; a route from London to Nottingham through Rushden, Kettering and

Market Harborough; another route to Nottingham through Newport Pagnell, Olney, Wellingborough, Kettering, Uppingham and Oakham; and a cross-country route from Northampton to Oxford though Towcester and Brackley. Another cross-country route, between Northampton and Peterborough, was omitted because it replicates the route of both the Lower Nene and the London & Birmingham Railway's cross-country line of 1845, and its inclusion would cause multicollinearity.

Access to railways. The railway lines serving Northamptonshire and Rutland are listed in Table 7.2. Access to railways is measured in two ways; by whether a line crossed a parish and whether a convenient station was available. Three types of line are distinguished, as explained above: A trunk line is normally double track, and is designed to carry large volumes of inter-city traffic at high speed; they are at least thirty miles in length, but may be as long as 300 miles. Cross-country lines carry moderate amounts of traffic at moderate speeds, and serve major provincial town along the way; they are normally 30–100 miles in length. Local lines carry small amounts of traffic at low speeds; they typically serve suburban or rural areas and are up to fifty miles in length. Cross-country and local lines may be either single or double track, and tend to have tighter curves and steeper gradients than trunk lines. The classification was made by consulting relevant railways histories as listed in Casson's (2009) regional bibliography.

Railways are allocated to census decades according to their opening dates as recorded by Cobb (2006). Where a given line is opened in sections, each section is dated separately and so the parish at which the sections join appears in the data set as an end-on junction between two lines.

Accurate mapping is especially important in determining whether railways cross parishes, since, unlike rivers and major roads, they do not usually run along their boundaries. Ordnance survey index maps in the Bodleian Library, Oxford, were used to determine which parishes were crossed by which line. A set of these maps for Northamptonshire is reprinted in Serjeantson and Adkins (1906). Where there are several parishes within a single town or city they are consolidated for the purposes of measuring railway access, for otherwise many urban parishes within a very short distance of a railway would be classified as having no railway at all.

Access to a station is determined, not by parish boundaries, but by whether there is a station within 1.5 miles by road from the main centre of population within the parish (the centre of population being assessed from a one-inch Ordnance Survey map). In principle, therefore, a parish can have a station but no railway, although in practice this is very rare. Using this definition of station access addresses a number of practical problems, e.g. that a railway station for a village lies on the opposite bank of a river to the village or town that it serves.

7.5 Results

A preliminary assessment of the population history can be obtained through correlation analysis. A simple application is shown in Table 7.3. The first column

Table 7.2 Railways passing through Northamptonshire and Rutland, 1830–1914

Date opened	*Type*	*Name of company and route through the counties*
1838	T	London & Birmingham (L&BR) [later London & North Western (LNWR)]: [London] – Wolverton – Roade – Blisworth – Weedon – [Rugby]
1845	C	L&BR [later LNWR]: Blisworth – Northampton – Wellingborough – Thrapston – Oundle – Wansford – Peterborough
1846	C	Midland Railway (MR): [Stamford] – Peterborough
1848	C	MR: [Melton Mowbray] – Saxby – Ashwell – Oakham – Manton – Luffenham – [Stamford]
1848	C	Great Northern Railway (GNR): Werrington – [Spalding]
1850	T	Great Western Railway (GWR): [Oxford] – Aynho – Kings Sutton – [Banbury]
1850	T	GNR: [Huntingdon] – Peterborough – Werrington
1850	L	LNWR: [Market Harborough] – Weston-by-Welland – Seaton
1850	L	Buckinghamshire Railway [later LNWR]: [Banbury] – Brackley – [Buckingham]
1851	L	LNWR: Seaton – Luffenham – [Stamford]
1852	T	GNR: Werrington – Helpston – [Essendine]
1857	T	MR: [Market Harborough] – Kettering – Wellingborough – [Bedford]
1859	L	LNWR: Northampton – [Market Harborough]
1866	C	Peterborough Wisbech and Sutton [later Midland and Great Northern Joint Railway]: Peterborough – [Wisbech]
1866	C	Northampton & Banbury Junction [later Stratford-on-Avon & Midland Junction (S&MJR)]: Blisworth – Towcester
1866	L	Kettering, Thrapstone & Huntingdon Railway [later MR]: Kettering – Thrapston – [Kimbolton]
1866	L	LNWR: Wolverton – Newport Pagnell
1867	L	LNWR: Wansford – [Stamford]
1871	L	East & West Junction Railway (E&WJR) [later S&MJR]: Towcester – Helmdon
1872	L	E&WJR [later S&MJR]: Helmdon – Cockley Brake – [Banbury]
1872	L	MR: Northampton – Ravenstone – [Bedford]
1873	C	E&WJR [laterS&MJR]: Towcester – Woodford – Fenny Compton
1877	L	MR: Kettering – Cransley
1879	T	MR: Manton – Seaton – Corby – Kettering
1879	L	Great Northern & London & North Western Joint: [Hallaton] – Weston-by-Welland
1879	L	LNWR: Seaton – Kings Cliffe – Wansford
1881	T	LNWR: Roade – Northampton – Long Buckby – [Rugby]
1882	L	MR: Ashwell – Cottesmore
1887	C	Banbury & Cheltenham Direct Railway [later GWR]: Kings Sutton – Chipping Norton
1888	L	LNWR: Weedon – Daventry
1891	C	S&MJR: Towcester – Roade – Ravenstone
1893	C	MR: Saxby – Little Bytham
1893	L	MR: Wellingborough – Higham Ferrers
1893	L	MR: Cransley – Loddington
1894	L	LNWR: Seaton – Uppingham
1895	L	LNWR: Daventry – Marton Junction – [Leamington Spa]
1898	T	Great Central Railway (GCR): [Rugby] – Woodford – Culworth – Helmdon – Brackley – [London]
1900	C	GCR: Culworth – [Banbury]
1910	T	GWR: [Bicester] – Aynho – [Banbury]

Note
T=Trunk line; C=cross-country line; L=local line. Towns outside the two counties are shown in square brackets.

shows that the rates of population growth in the same parish in successive decades tend to be negatively correlated with each other – i.e. an above-average increase in one decade tended to be followed by a below-average increase in the next decade. A natural explanation of this lies in short-term population mobility. In any decade some parishes with high growth rates will have received a transitory influx of people shortly before the second census date, following a transitory outflow just before the previous census date, while some parishes with low growth rates will have experienced the opposite effect. Such effects would be expected to disappear in the long run, and this is indeed the case. The second column of the table shows that if a double lag is applied then many of the negative correlations disappear and positive correlations predominate. Positive correlations are consistent with the view that some parishes have long-term characteristics that tend to attract people and others have long-term characteristics that repel people, with short-term mobility being superimposed on this effect. Similar results were obtained in the Oxfordshire study.

The main results are derived from regression analysis. The formal derivation of the regression equations is explained in the appendix. Table 7.4 summarises the determinants of population density over the period 1801–91 as a whole, while Table 7.5 reports the determinants of population growth decade by decade. Both population density and population growth are regressed on the five sets of explanatory variables described in section 7.3. Probability values are shown under each estimated coefficient, and significance in indicated by *(10 per cent), ** (5 per cent) and *** (1 per cent).

In Table 7.4, the first column of results relates to population in 1801, while the remaining three columns relate to population in 1891. The first of these three columns, labelled A, provides a direct comparison with the results for 1801. The next column, labelled B, includes population density in 1801 as a potential determinant of population density in 1891. This investigates whether population in 1801 can be understood as a proxy for unobserved parish-specific fixed factors that influence population growth across the entire period. If it can, as the results suggest it can, then the regression coefficients reported in this column and the next can be interpreted in terms of impacts on population growth instead of

Table 7.3 Persistence of decadal parish population growth: Pearson zero-order serial correlations between growth rates

Decade	*One-decade lag*	*Two-decade lag*
1811–21	–0.262	
1821–31	–0.013	0.035
1831–41	–0.022	–0.207
1841–51	–0.052	0.022
1851–61	0.074	0.059
1861–71	0.154	0.126
1871–81	–0.132	0.041
1881–91	0.134	0.231

population density. The final column includes all railway projects completed by 1891 as an additional determinant of population density in 1891. It provides a summary indication of the impact of railways built in the period 1831–91. Table 7.5 reports the results for each decade in a separate column.

The economic model set out above explains not only population growth but also the expansion of the railway system. For each decade there is a set of five equations, of which the population growth equation is only one. Three are concerned with investment in different types of railway, and one with station building. There is insufficient space to analyse the dynamics of railway and station building in this paper, but an example of the results that are generated is provided in Table 7.6, which reports all five regressions for 1871–81. Overall, the results for Northamptonshire suggest that local population density, even when lagged, has limited influence on railway construction through a parish. It also suggests that new railways tend to avoid parishes that are already served by railways, so that over time railways spread out across the county and do not concentrate in just a few major areas.

In Table 7.4 both distance from Northampton (NH5M) and distance from Peterborough (PB5M) are insignificant in 1801, but significantly positive in all regressions in 1891. This indicates that neither Northampton nor Peterborough were significant urban agglomerations in 1801, but that they had become agglomerations by 1891 as a result of suburban development, involving a mixture of factories and housing. In Table 7.5 proximity to Northampton is a significant factor in 1801–11, 1831–41 and 1861–91, but is significant for Peterborough only in 1821–31. Similar results were obtained for Oxfordshire in the previous study. The dominant town, Oxford, expanded dramatically, but the second town, Banbury, expanded to a more modest degree.

Comparing the effects of the three institutional factors, borough status (BORO), possession of an active market (MKT1) and location of a defunct market (MKT2), it can be seen from Table 7.4 that possession of an active market is the key determinant of population density in both 1801 and 1891, being positive and significant in all regressions. The first column shows that an active market enhanced population density in 1801, the second that it also enhanced density in 1891, the third that parishes with active markets grew faster than those without them, and the fourth that they grew faster even allowing for the impact of railways on their growth. Table 7.5 suggests that the growth of active market centres was highest relative to other parishes in 1871–81.

According to Table 7.4, possession of a defunct market also had a positive and significant impact on population density in 1801 (column 1), and this was sustained through to 1891 by legacy effects (column 2). Unlike active markets, however, defunct markets did not stimulate higher growth (columns 3 and 4). A possible reason why defunct markets exhibit high population density is that in early medieval market towns and villages land was sub-divided into small individually owned plots and large market squares were established suitable for later in-filling. This may have created a potential for later development that was

Table 7.4 Population densities at the beginning and end of the century

Acronym	*Interpretation*	*1801*	*1891(A)*	*1891(B)*	*1891(C)*
Constant		1.119***	1.131***	0.007	0.040
		(0.000)	(0.000)	(0.955)	(0.729)
Distances					
NH5M	Northampton: less than 5 miles	–0.022	0.417*	0.440***	0.397***
		(0.900)	(0.066)	(0.005)	(0.010)
PB5M	Peterborough: less than 5 miles	0.096	0.703**	0.606***	0.246*
		(0.721)	(0.032)	(0.004)	(0.074)
Institutional status					
BORO	Pre-1800 borough	–0.069	–0.168	–0.099	–0.078
		(0.798)	(0.615)	(0.492)	(0.622)
MKT1	Market active about 1800	1.618***	2.129***	0.503***	0.339**
		(0.000)	(0.000)	(0.001)	(0.037)
MKT2	Early market defunct by 1800	0.468***	0.403**	–0.068	–0.083
		(0.002)	(0.011)	(0.332)	(0.256)
Soil type					
LIME	Limestone	0.027	0.036	0.008	0.018
		(0.818)	(0.790)	(0.883)	(0.749)
IRON	Ironstone	0.245**	0.221	–0.025	–0.031
		(0.048)	(0.110)	(0.655)	(0.605)
BRASH	Brash	0.081	0.000	–0.080	–0.074
		(0.672)	(0.998)	(0.488)	(0.506)
SAND	Sand	0.025	–0.000	–0.026	–0.067
		(0.886)	(0.999)	(0.753)	(0.386)
GRAVEL	Gravel	–0.084	–0.072	0.012	0.028
		(0.561)	(0.670)	(0.871)	(0.693)
ALLM	Alluvium	0.181*	0.173	–0.010	–0.037
		(0.079)	(0.136)	(0.846)	(0.471)
CLAY	Clay	–0.098	0.033	0.132	0.111
		(0.497)	(0.846)	(0.121)	(0.165)
FLINT	Flint	0.436	0.447	0.008	0.091
		(0.072)	(0.101)	(0.934)	(0.386)
MUD	Mudstone	–0.001	0.018	0.019	0.002
		(0.992)	(0.900)	(0.750)	(0.977)
Rivers and canals					
LNENE	Lower Nene: Northampton – Peterborough	0.259*	0.590***	0.334***	0.144*
		(0.062)	(0.001)	(0.002)	(0.095)
UNENE	Upper Nene: Northampton – Dodford	–0.187	0.345	0.533	0.567
		(0.606)	(0.325)	(0.120)	(0.117)
ISE	River Ise	–0.005	0.461	0.209*	0.221*
		(0.982)	(0.508)	(0.061)	(0.053)
WELD	River Welland: Stamford – Market Harborough	–0.654	0.122	0.224	0.235
		(0.622)	(0.602)	(0.169)	(0.153)
CANAL	Grand Junction Canal and Oakham Canals	0.193	0.280	0.086	–0.025
		(0.272)	(0.145)	(0.367)	(0.827)

Roads					
WATST	Watling Street: Stony Stratford – Daventry	0.616**	0.711**	0.092	0.186
		(0.035)	(0.011)	(0.451)	(0.184)
NPNMH	Newport Pagnell – Northampton – Market Harborough	0.393*	0.755**	0.360**	0.396**
		(0.078)	(0.019)	(0.045)	(0.018)
RKMH	Rushden – Kettering – Market Harborough	0.737***	1.437***	0.697***	0.640***
		(0.002)	(0.008)	(0.002)	(0.007)
OLOAK	Olney – Oakham	0.014	0.195	0.181**	0.186**
		(0.958)	(0.488)	(0.012)	(0.020)
NTB	Northampton – Towcester – Brackley	0.699***	0.836***	0.134	–0.052
		(0.002)	(0.000)	(0.289)	(0.723)
DENS01	Population density at the start of the decade	–	–	1.005***	1.005***
				(0.000)	(0.000)
Railways and stations built 1830–90					
TRUNK30	Trunk railways opened 1830–9	–	–	–	0.090
					(0.658)
CC40	Cross-country railways opened 1840–9	–	–	–	0.190***
					(0.010)
TRUNK50	Trunk railways opened 1850–9	–	–	–	–0.012
					(0.920)
LOCAL50	Local railways opened 1850–9	–	–	–	–0.146
					(0.131)
CC60	Cross-country railways opened 1860–9	–	–	–	0.420*
					(0.054)
LOCAL60	Local railways opened 1850–9	–	–	–	0.268
					(0.164)
TRUNK70	Trunk railways opened 1870–9	–	–	–	–0.127
					(0.192)
CC70	Cross-country railways opened 1870–9	–	–	–	–0.293**
					(0.023)
LOCAL70	Local railways opened 1870–9	–	–	–	–0.035
					(0.629)
TRUNK80	Trunk railways opened 1880–9	–	–	–	0.515***
					(0.002)
CC80	Cross-country railways opened 1880–9	–	–	–	0.031
					(0.845)
LOCAL80	Local railways opened 1880–9	–	–	–	–0.252*
					(0.062)
CSTA80	Cumulative stations opened				
1830–79	–	–	–	0.077	
				(0.290)	
STA80	Stations opened 1880–9	–	–	–	–0.013
					(0.917)
R^2		0.281	0.380	0.857	0.877
Adjusted R^2		0.229	0.335	0.846	0.862
F		5.338***	8.357***	78.323***	57.003***
		(0.000)	(0.000)	(0.000)	(0.000)
N		352	352	352	352

Table 7.5 Determinants of the rate of population growth by decade

Variable	*1801–11*	*1811–21*	*1821–31*	*1831–41*	*1841–51*
Constant	0.081**	0.043	0.003	0.145***	0.017
	(0.038)	(0.424)	(0.919)	(0.008)	(0.657)
Exogenous variables					
NH5M	0.062**	0.019	0.041	0.055*	–0.009
	(0.044)	(0.430)	(0.094)	(0.068)	(0.755)
PB5M	0.058	0.002	0.159***	0.050	–0.011
	(0.140)	(0.968)	(0.000)	(0.157)	(0.809)
BORO	–0.051	0.028	0.041	0.055	–0.024
	(0.107)	(0.361)	(0.151)	(0.279)	(0.590)
MKT1	0.073*	–0.070*	0.006	0.059	–0.005
	(0.069)	(0.077)	(0.851)	(0.283)	(0.907)
MKT2	0.036*	–0.064**	–0.038***	0.056*	–0.006
	(0.096)	(0.023)	(0.008)	(0.051)	(0.764)
LIME	0.011	–0.007	0.037	–0.049	–0.010
	(0.545)	(0.720)	(0.025)	(0.027)	(0.602)
IRON	–0.029	–0.003	0.003	–0.024	–0.007
	(0.193)	(0.868)	(0.840)	(0.259)	(0.721)
BRASH	–0.027	0.013	–0.010	0.006	–0.057
	(0.332)	(0.620)	(0.696)	(0.841)	(0.112)
SAND	–0.013	–0.005	–0.034	0.033	–0.029
	(0.589)	(0.844)	(0.094)	(0.289)	(0.411)
GRAVEL	0.011	0.008	0.011	–0.036	–0.001
	(0.638)	(0.658)	(0.544)	(0.114)	(0.970)
ALLM	0.008	0.012	–0.016	–0.009	–0.009
	(0.638)	(0.409)	(0.207)	(0.544)	(0.513)
CLAY	0.033	0.011	0.007	–0.042	0.036
	(0.217)	(0.676)	(0.740)	(0.113)	(0.137)
FLINT	0.042	0.060**	–0.030	–0.104	–0.015
	(0.171)	(0.028)	(0.320)	(0.028)	(0.659)
MUD	0.004	–0.007	–0.004	–0.004	0.001
	(0.881)	(0.736)	(0.813)	(0.855)	(0.937)
LNENE	0.038	–0.004	0.049**	0.015	0.033
	(0.105)	(0.824)	(0.015)	(0.500)	(0.145)
UNENE	0.029	0.015	0.001	0.037	0.017
	(0.496)	(0.713)	(0.963)	(0.430)	(0.730)
ISE	0.017	0.004	0.042	0.035	–0.072**
	(0.640)	(0.874)	(0.109)	(0.347)	(0.012)
WELD	0.006	0.075	0.056	0.003	0.025
	(0.851)	(0.352)	(0.103)	(0.955)	(0.464)
CANAL	0.004	0.033	0.045**	–0.042	–0.040
	(0.893)	(0.136)	(0.023)	(0.218)	(0.188)
WATST	0.051	–0.022	0.026	0.029	–0.013
	(0.254)	(0.512)	(0.340)	(0.333)	(0.188)
NPNMH	–0.021	0.002	0.022	0.003	0.015
	(0.613)	(0.946)	(0.430)	(0.922)	(0.601)
RKMH	0.071**	0.004	–0.006	0.053	0.001
	(0.034)	(0.872)	(0.829)	(0.139)	(0.972)
OLOAK	0.013	–0.023	–0.018	0.094	–0.026
	(0.645)	(0.307)	(0.670)	(0.165)	(0.209)

NTB	0.015	–0.028	0.011	–0.045	0.010
	(0.690)	(0.529)	(0.678)	(0.414)	(0.749)
DENS	–0.029**	0.039***	0.018	–0.028	0.020**
	(0.036)	(0.009)	(0.200)	(0.223)	(0.033)
Railway building					
TRUNK30	–	–	–	0.036	0.019
				(0.323)	(0.503)
CC40	–	–	–	–	0.027
					(0.347)
TRUNK50	–	–	–	–	–
LOCAL50	–	–	–	–	–
CC60	–	–	–	–	–
LOCAL60	–	–	–	–	–
TRUNK70	–	–	–	–	–
CC70	–	–	–	–	–
LOCAL70	–	–	–	–	–
TRUNK80	–	–	–	–	–
CC80	–	–	–	–	–
LOCAL80	–	–	–	–	–
Station building					
STA30	–	–	–	0.482**	0.116
				(0.044)	(0.348)
STA40	–	–	–	–	0.080***
					(0.008)
STA50	–	–	–	–	–
STA60	–	–	–	–	–
STA70	–	–	–	–	–
STA80	–	–	–	–	–
R^2	0.076	0.097	0.205	0.193	0.156
Adjusted R^2	0.005	0.028	0.144	0.125	0.080
F	1.073***	1.398	3.353***	2.863***	2.048***
	(0.371)	(0.101)	(0.000)	(0.000)	(0.002)
N	352	352	352	352	352

Notes to acronyms:
See the text for explanations of exogenous variables, railway variables and station variables. DENS = Population density at the start of the decade

Table 7.5 continued

Variable	*1851–61*	*1861–71*	*1871–81*	*1881–91*
Constant	–0.079**	–0.050	–0.002	–0.176***
	(0.0293)	(0.189)	(0.976)	(0.000)
Exogenous variables				
NH5M	–0.014	0.049*	0.177*	0.112***
	(0.617)	(0.080)	(0.090)	(0.009)
PB5M	0.028	–0.034	0.025	0.011
	(0.473)	(0.353)	(0.460)	(0.812)
BORO	0.012	–0.027	–0.032	–0.116*
	(0.762)	(0.665)	(0.570)	(0.053)
MKT1	0.037	0.070	0.197***	0.071
	(0.375)	(0.336)	(0.009)	(0.249)
MKT2	–0.039**	–0.015	0.008	–0.013
	(0.022)	(0.398)	(0.748)	(0.555)
LIME	0.013	0.029	0.026	0.007
	(0.452)	(0.161)	(0.347)	(0.774)
IRON	0.023	–0.010	–0.001	0.015
	(0.236)	(0.584)	(0.957)	(0.529)
BRASH	–0.062*	–0.001	0.003	0.047
	(0.087)	(0.966)	(0.925)	(0.209)
SAND	0.008	–0.021	0.019	–0.022
	(0.809)	(0.454)	(0.626)	(0.467)
GRAVEL	0.006	0.025	–0.025	0.042
	(0.826)	(0.232)	(0.491)	(0.116)
ALLM	0.038***	0.001	–0.012	–0.005
	(0.005)	(0.953)	(0.590)	(0.801)
CLAY	0.053**	0.005	0.016	–0.006
	(0.026)	(0.844)	(0.580)	(0.830)
FLINT	0.015	0.040	0.003	0.075
	(0.666)	(0.200)	(0.957)	(0.108)
MUD	–0.002	–0.030	0.067**	0.003
	(0.919)	(0.128)	(0.018)	(0.905)
LNENE	0.031	0.028	0.010	–0.060**
	(0.228)	(0.263)	(0.762)	(0.029)
UNENE	0.036	–0.003	0.374	–0.029
	(0.512)	(0.956)	(0.231)	(0.623)
ISE	0.049	0.049	0.016	0.062
	(0.134)	(0.192)	(0.775)	(0.109)
WELD	–0.029	–0.007	0.049	0.016
	(0.257)	(0.865)	(0.129)	(0.633)
CANAL	–0.026	0.059	–0.068	0.011
	(0.315)	(0.197)	(0.222)	(0.742)
WATST	0.016	0.053	–0.057	0.102**
	(0.580)	(0.452)	(0.540)	(0.047)
NPNMH	0.187***	0.063	0.031	0.087
	(0.000)	(0.199)	(0.688)	(0.164)
RKMH	–0.004	0.068	0.198*	0.268***
	(0.918)	(0.193)	(0.067)	(0.000)
OLOAK	–0.010	–0.047	0.131**	0.008
	(0.682)	(0.262)	(0.021)	(0.794)

NTB	0.077	–0.032	–0.002	–0.076*
	(0.192)	(0.426)	(0.962)	(0.079)
DENS	0.002	0.002	–0.057**	0.039***
	(0.833)	(0.927)	(0.046)	(0.008)
Railway building				
TRUNK30	–0.014	–0.065	0.002	–0.008
	(0.686)	(0.165)	(0.982)	(0.869)
CC40	0.077***	0.022	0.038	0.070**
	(0.005)	(0.398)	(0.296)	(0.015)
TRUNK50	0.023	0.049	0.040	–0.086
	(0.492)	(0.163)	(0.475)	(0.073)
LOCAL50	–0.069**	0.050*	–0.031	–0.069
	(0.018)	(0.076)	(0.468)	(0.116)
CC60	–	0.160**	0.035	0.063
		(0.017)	(0.617)	(0.474)
LOCAL60	–	0.069*	0.054	–0.062
		(0.099)	(0.298)	(0.113)
TRUNK70	–	–	–0.018	0.039
			(0.740)	(0.301)
CC70	–	–	–0.117	–0.060
			(0.101)	(0.210)
LOCAL70	–	–	–0.026	0.009
			(0.502)	(0.793)
TRUNK80	–	–	–	–0.049
				(0.423)
CC80	–	–	–	0.027
				(0.708)
LOCAL80	–	–	–	–0.018
				(0.628)
Station building				
STA30	0.049	0.065	0.265**	0.071
	(0.391)	(0.401)	(0.039)	(0.390)
STA40	–0.046*	0.006	0.008	0.006
	(0.093)	(0.832)	(0.825)	(0.842)
STA50	0.038	0.006	–0.001	0.060
	(0.325)	(0.880)	(0.985)	(0.177)
STA60	–	–0.083	0.014	0.015
		(0.113)	(0.861)	(0.754)
STA70	–	–	0.048	–0.069*
			(0.374)	(0.059)
STA80	–	–	–	–0.014
				(0.665)
R^2	0.228	0.176	0.238	0.281
Adjusted R^2	0.151	0.0851	0.143	0.181
F	2.947***	1.934***	2.500***	2.802***
	(0.000)	(0.002)	(0.000)	(0.000)
N	352	352	352	352

Table 7.6 Estimated simultaneous equations for population growth, railway building and station building for the decade 1871–81

Variable	*Population growth*	*Trunk railways*	*Cross-country railways*	*Local railways*	*Stations*
Constant	–0.001	0.001	0.062	0.103	–0.029
	(0.993)	(0.971)	(0.166)	(0.166)	(0.428)
Exogenous variables					
NH5M	0.182*	–0.034*	0.024	0.177**	–0.061**
	(0.081)	(0.062)	(0.242)	(0.036)	(0.047)
PB5M	0.021	–0.017	–0.085*	–0.144**	0.036
	(0.554)	(0.568)	(0.053)	(0.012)	(0.276)
BORO	–0.029	–0.023	0.005	0.031	–0.033
	(0.608)	(0.744)	(0.887)	(0.756)	(0.470)
MKT1	0.188***	–0.020	0.007	0.099	–0.036
	(0.008)	(0.786)	(0.860)	(0.337)	(0.435)
MKT2	0.008	–0.013	–0.024	0.011	0.071
	(0.750)	(0.721)	(0.367)	(0.818)	(0.102)
LIME	0.028	–0.034	–0.009	0.031	0.014
	(0.312)	(0.235)	(0.653)	(0.449)	(0.541)
IRON	–0.004	0.065**	0.002	–0.081*	–0.003
	(0.886)	(0.016)	(0.947)	(0.052)	(0.910)
BRASH	–0.006	0.039*	0.095*	–0.020	0.024
	(0.884)	(0.100)	(0.060)	(0.798)	(0.595)
SAND	0.016	0.054	–0.025	–0.006	0.055
	(0.665)	(0.228)	(0.526)	(0.937)	(0.204)
GRAVEL	–0.026	–0.028	0.020	–0.038	–0.028
	(0.461)	(0.445)	(0.511)	(0.513)	(0.448)
ALLM	–0.016	0.016	–0.025	–0.088***	0.040
	(0.488)	(0.452)	(0.191)	(0.008)	(0.101)
CLAY	0.020	–0.026	–0.041	0.032	–0.020
	(0.486)	(0.130)	(0.103)	(0.575)	(0.483)
FLINT	–0.004	–0.046	0.166	–0.019	0.071
	(0.941)	(0.256)	(0.180)	(0.849)	(0.483)
MUD	0.064**	–0.001	0.023	0.085**	–0.018
	(0.028)	(0.963)	(0.372)	(0.026)	(0.567)
LNENE	0.006	–0.038	–0.057**	0.034	–0.054*
	(0.861)	(0.230)	(0.038)	(0.592)	(0.055)
UNENE	0.364	–0.026	–0.117**	–0.124*	–0.010
	(0.240)	(0.205)	(0.028)	(0.057)	(0.773)
ISE	0.016	0.126	–0.017	–0.027	–0.111**
	(0.769)	(0.271)	(0.474)	(0.503)	(0.043)
WELD	0.046	0.189	–0.014	0.164	–0.031
	(0.148)	(0.169)	(0.541)	(0.241)	(0.764)
CANAL	–0.065	–0.038*	–0.078**	–0.062	–0.025
	(0.240)	(0.063)	(0.009)	(0.255)	(0.164)
WATST	–0.067	0.007	0.215**	0.105	–0.055
	(0.467)	(0.607)	(0.029)	(0.243)	(0.376)
NPNMH	0.031	–0.014	–0.037*	–0.086	–0.018
	(0.682)	(0.555)	(0.096)	(0.461)	(0.621)

RKMH	0.189*	–0.103	–0.019	–0.117**	0.021
	(0.067)	(0.115)	(0.307)	(0.026)	(0.499)
OLOAK	0.130**	0.123	–0.002	–0.075	0.020
	(0.024)	(0.186)	(0.881)	(0.010)	(0.673)
NTB	0.014	0.050*	–0.067	–0.082	–0.007
	(0.760)	(0.068)	(0.121)	(0.619)	(0.911)
DENS71	–0.055**	0.026	0.081	0.090	0.073
	(0.049)	(0.662)	(0.322)	(0.439)	(0.192)
DENS61	–	–0.033	–0.081	–0.070	–0.058
		(0.589)	(0.367)	(0.557)	(0.348)
Railway building					
TRUNK30	0.044	–0.007	–0.026	–0.045	0.007
	(0.510)	(0.618)	(0.376)	(0.308)	(0.656)
CC40	0.039	–0.002	–0.018	–0.040	–0.010
	(0.183)	(0.962)	(0.276)	(0.237)	(0.596)
TRUNK50	0.031	–0.032	–0.047	–0.035	0.058
	(0.549)	(0.611)	(0.107)	(0.374)	(0.284)
LOCAL50	–0.039	–0.022	–0.054**	0.038	–0.060**
	(0.325)	(0.575)	(0.016)	(0.537)	(0.050)
CC60	0.031	–0.052	0.174**	0.075	–0.050
	(0.556)	(0.184)	(0.050)	(0.499)	(0.356)
LOCAL60	0.081	–0.089**	–0.020	–0.040	0.003
	(0.122)	(0.019)	(0.308)	(0.226)	(0.869)
TRUNK70	–0.019	–	–	–	0.209
	(0.723)				(0.142)
CC70	–0.112	–	–	–	0.145
	(0.118)				(0.332)
LOCAL70	–0.027	–	–	–	0.257***
	(0.473)				(0.003)
Station building					
CSTA70	0.018	0.080**	0.007	0.001	–0.054**
	(0.563)	(0.036)	(0.694)	(0.986)	(0.014)
STA70	0.051	–	–	–	–
	(0.341)				
R^2	0.232	0.171	0.168	0.121	0.266
Adjusted R^2	0.144	0.085	0.082	0.030	0.189
F	2.646***	1.991***	1.944***	1.329	3.177***
	(0.000)	(0.001)	(0.002)	(0.113)	(0.000)
N	352	352	352	352	352

missing in villages controlled by the owner of a single large estate (Hoyle, 2013).

Borough status has a negative but insignificant effect throughout. These results are broadly similar to those obtained for Oxfordshire. The main exception is that in Oxfordshire borough status was positive and significant until initial population density was controlled for, and then became negative and significant. The explanation may be that in Oxfordshire population growth was focused on just a few established centres – Oxford, Banbury and, to some extent, Witney – with other centres, including long-established boroughs, going into decline, while in Northamptonshire population growth was more widely distributed across a range of towns, such as Kettering, Wellingborough and Higham Ferrers, that included some established boroughs.

Soil types have very little influence on population density. A direct comparison of population densities in 1801 and 1891 (in the first two columns of Table 7.4) indicates that ironstone – for which Northamptonshire is particularly noted – and alluvium – characteristic of river valleys – have a significant positive impact in 1801 and an insignificant positive impact in 1891. Thus the impact of soil type on population density is greater in 1801, when Northamptonshire is still a predominantly agricultural county, than in 1891, when it is more commercial and industrial. Table 7.5 suggests that the impact of both alluvium and clay on population density was significant in 1851–61, but not at other times. Alluvium was also a significant positive factor in Oxfordshire, but otherwise soil types in Oxfordshire, as in Northamptonshire, had little influence on either density or growth.

The Lower Nene, between Peterborough and the head of navigation at Northampton, is the only river to have a consistently positive impact on population density, according to columns 1 and 2 of Table 7.4. It also has a positive impact on population growth (column 3), even when allowing for the building of the Northampton to Peterborough Railway and other railways through the valley (column 4). However, Table 7.5 shows that the impact of the Lower Nene is most significant in 1821–31, before the railways arrived, and that the effect reverses in 1881–91 once the railways are well-established. The River Ise, which joins the Nene at Wellingborough, has no impact on population density, but has a significant positive impact on population growth. This almost certainly reflects the rapid growth of population in the Kettering area due to the development of local mining and manufacturing industry. Neither the Upper Nene nor the Welland has any significant impact on population density or growth, which is not surprising since they are generally unsuitable for river traffic. More surprising, perhaps, is that canals are of no significance either. Given that the Grand Junction Canal connects London with Birmingham, with a branch serving Leicester, and affords connections to Oxford, it is surprising that its impact on population, though positive, is insignificant in 1801. One explanation may be that it only touches the south-west corner of the county, and affects too few parishes to be statistically significant. In the Oxfordshire study rivers were much more important for initial population density than they were for subsequent growth.

Roads are very important in Northamptonshire, despite having little importance in Oxfordshire. At the beginning of the nineteenth century a high volume of traffic was carried over the trunk roads radiating north-west from London that passed through the county, and along some of the cross-country routes that intersected them. According to the first two columns of Table 7.4, four of the five main roads identified in this study had a significant impact on the population densities of the parishes through which they passed, and the fifth – the road from Olney to Oakham (part of a London–Nottingham stage coach route) – had a positive effect on population growth. According to Table 7.5, the impacts of the Rushden–Kettering–Market Harborough road and the Olney–Oakham road became most significant in the late nineteenth century. The two roads entering Market Harborough from the south had positive impacts on both population density and growth. These results suggest that the Northamptonshire road system was a more important influence on the distribution of population than were rivers and canals (with the exception of the Lower Nene). The continuing relevance of the road system to nineteenth-century growth, given the decline of stage coach traffic, may be due to the emergence of local traffic connected with the spread of rural industry. Northampton developed a substantial industrial district connected with shoes and leather, while Kettering and Wellingborough became the hubs of another district manufacturing metal products. Market Harborough developed as a railway hub and manufacturing centre, and this may have led to an increase in road-based feeder traffic.

The 1801 population density that is included in the 1891(B) and 1891(C) regressions is strongly significant and has a coefficient very close to unity. This is consistent with the view that initial population density is a proxy for a group of hidden fixed factors that influence equilibrium parish population density in the long run. The fact that the coefficient is slightly greater than one suggests that these factors increased in importance marginally over the century. This interpretation is also consistent with the results obtained in the Oxfordshire study.

Because its coefficient is close to one, subtracting initial population density from the right-hand side effectively eliminates population density from the regression, while subtracting it from the left-hand side transforms the dependent variable into proportional population growth 1801–91. This shows that population growth over the entire period 1801–91 is largely independent of population density in 1801. Lagged population density is significant in certain decades, however, according to Table 7.5; its impact on growth is significantly positive in 1811–21, 1841–51 and 1881–91, and significantly negative in 1801–11 and 1871–81.

The impact of railways on population density in 1891 is surprisingly modest. This impact is measured in Table 7.4, column 1891(C), using initial population density as a control. The impacts of railway lines and railway stations are estimated separately. Five sets of railways schemes emerge as significant, and two of these have negative effects on parish population. Cross-country lines opened in the 1840s and 1860s have significant positive impacts, as do trunk lines opened in the 1880s. However, cross-country lines opened in the 1870s and local lines opened in the 1880s have negative effects. The opening of stations has no significant effect on parish population whatsoever.

To interpret these results it is useful to consider the specific schemes involved, as listed in Table 7.2. Consider first the schemes with positive impacts. The cross-country routes of the 1840s were the main railways that opened up the county. Four separate schemes were involved, covering different parts of the county; as a result numerous parishes were affected, and so it is relatively easy for an impact to register statistical significance. Parishes served by the railway may have gained commercial advantage from the fact that parishes some distance from the line had no access to an alternative railway. The cross-country lines opened in the 1860s served only small portions of Northamptonshire before entering other counties, but the areas of Northamptonshire they served were increasingly commercial and industrial: Peterborough, Kettering, Thrapston, Blisworth and Towcester. The solitary trunk line scheme of the 1880s was a Roade to Rugby loop line that finally put Northampton on the trunk line railway map, after it had been bypassed by the London & Birmingham Railway. Although the line was not built to serve the villages through which it passed, many of the villages, like Long Buckby, were close to Northampton and were developing new industries that could benefit from the railway.

By contrast the schemes with negative impacts commenced later, in the 1870s, and were more rural in character. The only cross-country line of the 1870s was the Stratford-on-Avon and Midland Junction Railway, which extended the Blisworth to Towcester line westward to Stratford. It was owned and operated by a small independent company and never attracted the type of traffic its promoters envisaged. The two local lines of the 1880s were dead-end branches, one a mineral line and the other a passenger line that was later extended, but it is unclear precisely why their impact was so small. The insignificance of the impacts of the Midland Railway trunk line schemes of the 1850s and 1870s is also puzzling.

Table 7.5 suggests that where railway building had a positive impact, the impact was greatest in the decade during which the railway was constructed and in the decade immediately following. The cross-country routes of the 1840s were the only railways to have a sustained impact over later periods.

These results confirm some of the findings for Oxfordshire railways – namely that the impact of railways was relatively modest, and could even be negative. However, there are important differences. In Oxfordshire local lines appear to have had the greatest positive impact, and trunk lines the least impact (the impact of cross-country lines was hard to assess because there were few of them). The negative impact of trunk lines was explained in the Oxfordshire study by 'shake-out' in markets towns along the route, but this process does not seem to have occurred to the same extent in Northamptonshire.

There seems to have been much less market shake-out in Northamptonshire than in Oxfordshire. This is reflected in the results for both markets and railways; many established markets remained buoyant, and railways had relatively little impact, whether positive or negative, on their growth. Small market towns in Northamptonshire prospered more than the small market towns in Oxfordshire, probably because Northamptonshire became more industrialised. Unlike Oxfordshire, which remained predominantly agricultural, Northamptonshire was relatively close to Coventry and the industrialising West Midlands, and possessed an expanding leather and shoe industry which benefited from the growing market for improved footwear and luxury goods. It was in a geocentric position, and therefore well-placed for distribution by road as well as by railway and canal.

Although Northamptonshire was criss-crossed by railways – with lines radiating from London intersecting cross-country lines – railway building seems to have been driven mainly by competition between large trunk-line companies – the London & North Western and the Midland, and to some degree the Great Northern too. These large companies built lines in order to enhance the competitive position of their own network. Railway hubs emerged in response to regional rather than local needs, and the railways left it to the roads to convey local traffic to the nearest hub. The fact that station building had little impact on parish population growth reinforces the view that local traffic was not a priority for the railway companies.

The impact of railways may have been felt more in terms of occupational change than population levels. The coming of the railways may have been associated more with population turnover than with aggregate population change, e.g. small-scale industry moving into business premises in towns and villages with declining markets, or business premises in declining sectors being converted to purely residential use. Roade, for example, was a four-way railway junction on the London to Birmingham main line. Apart from the establishment of coal merchants, it was little affected by the railway until the 1860s. When the growth of factory employment in Northampton adversely affected local self-employed artisans, the town developed into a dormitory for workers at the Wolverton railway carriage works nearby. Local occupations changed significantly, but total population did not (Riden, 2002, p. 357).

7.6 Railway and station building

Table 7.7 analyses how railway and station building responded to parish characteristics over the period 1801–91. There were no openings over the period 1801–31 and no closures over the entire period. In column 1 the dependent variable measures the number of railways serving a parish in 1891. This is disaggregated by type of railway in columns 2–4. Columns 5 and 6 relate to station access in 1891, which is a binary variable, independent of the number of stations, as explained above. Column 5 analyses station building independently of local railways, while column 6 examines the connection between railways and their stations. Some of the logit and probit regressions did not converge to meaningful results, and so all the reported results are based on OLS.

Table 7.7 Estimated equations for railway building and station building, 1801–91

Variable	*All railways*	*Trunk railways*	*Cross-country railways*	*Local railways*	*Stations*	*Stations*
Constant	0.071	–0.156**	0.153	0.074	0.055	0.040
	(0.700)	(0.043)	(0.144)	(0.509)	(0.570)	(0.642)
NH5M	0.180	0.026	–0.031	0.185	–0.058	–0.110
	(0.409)	(0.973)	(0.746)	(0.154)	(0.533)	(0.249)
PB5M	1.222***	0.341***	1.131***	–0.251***	0.442***	0.183
	(0.001)	(0.007)	(0.000)	(0.000)	(0.001)	(0.264)
BORO	0.257	–0.044	0.254	0.048	0.189	0.132
	(0.487)	(0.804)	(0.161)	(0.806)	(0.317)	(0.455)
MKT1	0.708	0.066	0.284*	0.357	0.121	–0.057
	(0.071)	(0.721)	(0.097)	(0.109)	(0.547)	(0.748)
MKT2	–0.165	–0.087	–0.001	–0.077	0.049	0.093
	(0.107)	(0.175)	(0.981)	(0.277)	(0.519)	(0.166)
LIME	–0.060	0.008	–0.174***	0.106*	–0.022	–0.016
	(0.600)	(0.887)	(0.006)	(0.090)	(0.675)	(0.739)
IRON	0.133	0.108*	0.017	0.009	0.050	0.018
	(0.190)	(0.051)	(0.786)	(0.891)	(0.409)	(0.748)
BRASH	0.124	–0.006	0.155	–0.026	0.115	0.090
	(0.550)	(0.934)	(0.284)	(0.813)	(0.259)	(0.334)
SAND	0.275*	0.115	0.143	0.017	0.111	0.047
	(0.095)	(0.157)	(0.134)	(0.870)	(0.214)	(0.552)
GRAVEL	–0.145	0.004	–0.100	–0.049	–0.084	–0.049
	(0.290)	(0.954)	(0.261)	(0.556)	(0.255)	(0.437)
ALLM	–0.065	0.042	–0.048	–0.059	–0.016	0.001
	(0.454)	(0.347)	(0.312)	(0.253)	(0.739)	(0.985)
CLAY	0.153	0.067	–0.001	0.088	–0.019	–0.060
	(0.268)	(0.185)	(0.990)	(0.301)	(0.777)	(0.342)
FLINT	0.450**	0.084	0.176	0.190	0.236	0.124
	(0.024)	(0.454)	(0.243)	(0.245)	(0.147)	(0.385)
MUD	0.242**	0.022	0.088	0.132**	0.036	–0.026
	(0.024)	(0.675)	(0.183)	(0.049)	(0.580)	(0.658)
LNENE	0.606***	–0.020	0.672***	–0.046	0.055	–0.072
	(0.001)	(0.759)	(0.000)	(0.532)	(0.489)	(0.428)
UNENE	–0.396*	0.075	–0.146	–0.316**	–0.189**	–0.088
	(0.087)	(0.564)	(0.161)	(0.018)	(0.014)	(0.300)
ISE	0.396*	0.556***	–0.189***	0.030	–0.114	–0.220**
	(0.059)	(0.003)	(0.004)	(0.799)	(0.267)	(0.042)
WELD	0.662**	0.260	0.072	0.329	0.289	0.116
	(0.017)	(0.120)	(0.654)	(0.130)	(0.109)	(0.483)
CANAL	0.223	0.296**	–0.003	–0.076	–0.085	–0.138**
	(0.205)	(0.016)	(0.967)	(0.388)	(0.246)	(0.020)
WATST	0.421	0.009	0.257	0.156	0.040	–0.061
	(0.107)	(0.931)	(0.140)	(0.162)	(0.693)	(0.507)
NPNMH	0.119	–0.025	–0.049	0.193	–0.060	–0.098
	(0.619)	(0.813)	(0.505)	(0.238)	(0.618)	(0.425)
RKMH	–0.280	–0.123	–0.064	–0.093	0.002	0.073
	(0.330)	(0.481)	(0.554)	(0.556)	(0.989)	(0.649)
OLOAK	0.214	0.189	0.150	–0.125	0.087	0.042
	(0.192)	(0.120)	(0.108)	(0.161)	(0.396)	(0.641)
NTB	0.605*	0.188	0.295	0.122	0.156	0.011
	(0.088)	(0.220)	(0.142)	(0.556)	(0.406)	(0.947)
DENS01	0.038	0.034	–0.026	0.030	0.054**	0.043*
	(0.454)	(0.128)	(0.471)	(0.284)	(0.039)	(0.088)
TRUNKALL	–	–	–	–		0.248***
						(0.000)
CCALL	–	–	–	–		0.217***
						(0.000)
LOCALALL	–	–	–	–		0.280***
						(0.000)
R^2	0.306	0.214	0.467	0.136	0.167	0.356
Adjusted R^2	0.253	0.154	0.426	0.070	0.104	0.301
F	5.744***	3.552***	11.415***	2.057***	2.622***	6.387***
	(0.000)	(0.000)	(0.000)	(0.003)	(0.000)	(0.000)
N	352	352	352	352	352	352

These results clearly indicate that Peterborough was a more important railway hub than Northampton. They also indicate that Peterborough was a trunk line hub rather than a local hub – as it remains today. The role of Northampton is understated, however, because some of the important junctions feeding traffic into the town were at Blisworth and Roade rather than Northampton itself.

Railways tended to be built over sand (in lower lying areas) and flint and mudstone (in higher areas). As in Oxfordshire, railways tended to follow the river valleys – notably the cross-country route along the Lower Nene. A Midland Railway trunk line traversed the ironstone district along the River Ise, while the London & Birmingham line followed the course of the Grand Junction Canal. Some road routes also appear significant, but this may be misleading as the railways did not generally follow the lines of particular roads, but rather intersected other railways at key points along them.

Local population density had no significant impact on railway provision. After controlling for the factors above, only station building is influenced by population density in 1801. This reinforces the view that local railway provision was largely driven by economic forces operating on the towns and cities at either end of a line, and that intermediate parishes were generally served if and only if they lay along the most direct and convenient route between them. Once the route was decided, however, stations would be located near to major centres of population. The results in column 6 confirm the obvious point that railway provision encouraged station access, but show that this effect is surprisingly weak. The strongest impact appears to have come from the building of local lines, which is reasonable given that they were built mainly to serve local communities. Measured impacts may be weak partly because stations were not always located in the parish where the centre of population was located, but in a neighbouring parish. This kept noise and nuisance from the railway to a minimum, allowing the railway to follow a direct route and reduced the costs to the railway of buying land.

These results indicate that there was relatively little feedback from increases in population density onto railway building, as well as little impact from railway building on population density, as shown in section 7.5.

Railway and station building can also be analysed decade by decade. An example was given in Table 7.6 relating to 1871–81. The decadal results show that new railways tended to avoid parishes already served by a railway. The tendency for railways to agglomerate at hubs was more than outweighed by the tendency for railways to seek out local monopolies by avoiding areas already served by other railways.

7.7 Conclusions and implications for future research

This chapter has presented a model of equilibrium parish population in which equilibrium population is influenced by railway building and railway building is influenced by lagged population growth. The impact of railway building and population density on station building is also considered, together with the impact of station building on population growth. The results indicate that in Northamptonshire and Rutland the impact of railways on local population growth was small.

Overall, parish population growth was driven mainly by proximity to Northampton, location on the Lower Nene, location on a road of strategic importance and the legacy of an active market. The initial population density provided the base level from which population levels evolved, but did not itself influence the subsequent pace of growth to any significant extent. Fast-growing parishes did not grow faster than others in every period, but on average they experienced more decades of fast growth than their slower-growing neighbours. Within any decade, growth appears to have been driven mainly by changes in the impact of legacy factors such as active markets inherited from the eighteenth century. According to the model, changes in impacts appear to have been driven by changes in the economic environment (e.g.in relative commodity prices) that were common to all parishes in the county.

Railways followed routes dictated by engineering, geography and the urban business interests. In Northamptonshire this stimulated the building of both trunk lines and cross-country lines. Northampton was bypassed by the busy West Coast main line but Peterborough became a major hub on the somewhat quieter East Coast main line. Railways did not respond to local population density except through the siting of stations.

Northamptonshire developed a range of relatively small-scale industries, some of which were linked to local resources such as ironstone, but most of which seems to have exploited Northamptonshire's geographical centrality within the wider national economy. Road transport remained strategically important after canals and railways had been built.

The model used in this chapter is perfectly general. It is flexible, and can be tailored to the specific circumstances of different counties. It is also capable of further refinement. Using dynamic GIS models of road, railway and canal infrastructure it would be possible to provide more meaningful measures of access to transport at the level of local settlement and parish. The structure of local transport systems could also be examined in greater depth, e.g. analysing the differences between trunk and local lines using a deviation index and a measure of number of stations per route mile.

It is also possible to examine in more detail the potential causal mechanisms linking railways to population growth. For example, high population density may encourage more frequent train services, greater use of stopping trains and the introduction of workmen's fares. This may in turn attract new residents to a parish, leading to subsequent population growth. These effects could be accommodated by introducing timetables and fares as additional endogenous variables and postulating a lag structure that identifies their effects.

The data sources used in this study are essentially national sources, so that the study can be replicated on any English county. The model has been piloted on Oxfordshire and, having been extended and refined, is ready to be applied more generally.

Acknowledgements

We are grateful to John Poulter and Janet Casson for comments and criticisms on an earlier draft, and to the staff of the Bodleian Library map room for advice on early maps.

Appendix

Let parishes be indexed $i=1,\ldots, N$. Decennial census dates 1801–91 are indexed $t=0, 1,\ldots, 10$. With ten census dates, there are nine decades, each of which is analysed separately. Let z_{it} be the natural logarithm of the population of parish i at the beginning of decade t. The instantaneous proportional rate of population growth over decade t is $dz_{it}=z_{it+1}-z_{it}$. The logarithm of the acreage of the ith parish is a constant, w_i. The logarithm of population density of parish i at the beginning of decade t is $y_{it}=z_{it}-w_i$. Let $dy_{it}=y_{it+1}-y_{it}$ be the change in the logarithm of the population density over decade t; then $dy_{it}=dz_{it}$.

Let y_{it} be the equilibrium population density in parish i at the beginning of decade t. Let x_{ki} be the kth time-invariant characteristic of parish i ($k=1,\ldots, K$). Railway access is the only time-dependent parish characteristic. Three types of railway infrastructure are distinguished: *trunk line* ($h=1$), *cross-country route* ($h=2$) and *local line* ($h=3$). Let the integer variable dr_{hit} measure the number of new railway lines of type h opened through a parish in decade t; then railway access at the beginning of period t is measured by the cumulative stock of railway investment $r_{hit}=\sum_j dr_{hij}$, ($j=0,1,\ldots, t-1$). Railway infrastructure is highly durable, and so railway investment has significant legacy effects, which vary with the time elapsed. Thus it is the investments dr_{hij} rather than the level of access r_{hit} that determines equilibrium population density in period t.

Parish growth will be further stimulated if a local station is built. There is no intrinsic advantage to multiple stations, since when several lines serve the same parish a single station is usually more convenient than separate ones. Let the binary variable $ds_{it}=1$ if the parish acquires its first station during decade t, and zero otherwise. Let $s_{it}=1$ if the parish has access to a railway station at the beginning of decade t, and be zero otherwise. In the absence of station closures, $s_{it}=0$ only if $ds_{it}=0$ for all $j=0, 1,\ldots, t-1$, and conversely $ds_{it}=0$ if $s_{it}=1$.

In addition, there is an unobservable composite parish characteristic, z_i, which is constant over time but has time-dependent impacts on parish population, and a transitory shock in period t, u_{1it}.

Assuming that the impacts of all parish characteristics are linear, equilibrium population at the beginning of decade t is

$$y_{it}=\alpha_{10t}+\sum_k\alpha_{1kt}x_{ik}+\sum_h\sum_j\beta_{1hjt}dr_{hij}+\sum_j \gamma_{1jt}ds_i+\lambda_t z_i+u_{1it}$$
$$(h=1, 2, 3; j=0, 1,\ldots, t-1; k=1,\ldots, K) \qquad (1)$$

where α_{10t} is a constant, α_{1kt} measures the impact of the kth time-invariant factor, β_{1hjt} measures the impact of railway investment of type h in decade j, γ_{1tj}

measures the impact of access to a station opened in decade j, and λ_t is the impact of the unobservable characteristic. Note that all the parameters are decade-specific.

Each decade population adjusts fully from its initial level at the start of the decade, y_{it}, towards the equilibrium level for the end of the decade; because adjustment is complete the equilibrium level corresponds to the observed level, y_{it+1}.

Let $v_t = (\lambda_{t+1}/\lambda_t) - 1$ be the proportional increase in impact of the unobservable characteristic on equilibrium population over decade t. Comparing the equilibrium equations for the beginning and end of decade t makes it possible to eliminate the value of the unobservable characteristic and obtain the population growth equation:

$$dy_{it} = a_{10t} + \sum_k a_{1kt} x_{ik} + \sum_h b_{1htt} dr_{hit} + \sum_h \sum_j b_{1hjt} dr_{hij} + c_{1tt} ds_{it} + \sum_j c_{1tj} ds_{ij} + v_t y_{it} + (u_{1it+1} - (1 + v_t) u_{1it})$$
$$(h = 1, 2, 3; j = 0, 1, \ldots, t-1; k = 1, \ldots, K) \quad (2)$$

where $a_{10t} = \alpha_{10t} - (1 + v_t)\alpha_{10t-1}$, $a_{1kt} = \alpha_{1kt} - (1 + v_t)\alpha_{1kt-1}$, $b_{1htt} = \beta_{1htt}$, $b_{1hjt} = \beta_{1hjt+1} - (1 + v_t)\beta_{1hjt}$, $c_{1tt} = \gamma_{1tt}$, $c_{1tj} = \gamma_{1jt+1} - (1 + v_t)\gamma_{1jt}$. Equation (2) shows that the proportional rate of growth of population during each decade is determined by the change (from the previous decade) in the impacts of the time-invariant parish characteristics, changes in the legacy impacts of railway building and station opening, and the impact of current railway building and current station opening.

Railway building is modelled as follows. The target level of railway access at the beginning of decade t, r_{hit}, reflects the number of lines of type h that railway promoters consider appropriate to the parish. This number depends on the same time-invariant factors that influence population, x_{ik}. It also depends on the population density that prevailed at the beginning of the previous decade, y_{t-1}, which reflects the most recent official published population data for the parish. While the stock of previous lines will influence new railway building, the vintage of the lines is less likely to be important to railway investment than it is to population growth. Target railway access may also be influenced by the presence of an existing station, as a station may provide facilities for a junction with an existing line. This suggests that the target railway access of type h for parish i at the beginning of decade t is

$$r_{hit} = \alpha_{2h0t} + \sum_k \alpha_{2hkt} x_{ik} + \sum_m \beta_{2hmt} r_{mit-1} + \gamma_{2ht} s_{it-1} + \delta_{2ht} y_{it-1} + u_{2hit}$$
$$(h, m = 1,2,3; k = 1, \ldots K) \quad (3)$$

where α_{2h0t} is a constant, α_{2hkt} measures the impact of the kth time-invariant parish characteristic, β_{2hmt} measures the impact of the local stock of previous railway investment of type m on the target stock of type h, γ_{2ht} measures the impact of a previously constructed station, δ_{2ht} the impact of previous population density and u_{2hit} is an unobservable parish-specific transitory shock affecting railway investment of type h. It follows that

$$dr_{hit}=a_{2h0t}+\sum_k a_{2hkt}x_{ik}+\sum_m \beta_{2hmt+1}dr_{mit-1}+\sum_m b_{2hmt}r_{imt-1}+\gamma_{2ht+1}ds_{it-1}+c_{2ht}s_{it-1}+\delta_{2ht+1}y_{it}-\delta_{2ht}y_{it-1}+u_{2hit+1}-u_{2hit} \quad (h, m=1, 2, 3) \tag{4}$$

where $a_{2h0t}=\alpha_{2h0t+1}-\alpha_{2h0t}$, $a_{2hkt}=\alpha_{2hkt+1}-\alpha_{2hkt}$, $b_{2hmt}=\beta_{2hmt+1}-\beta_{2hmt+1}$, $c_{2ht}=\gamma_{2ht+1}-\gamma_{2ht}$.

Equation (4) relates current railway construction of a given type to the same time-invariant parish characteristics that influence population density, and to the inherited stock of lines of all different types at the beginning of the decade, the inherited stock of stations, the population density at the start of the decade and the population density at the start of the previous decade. The stock of lines of an existing type has an ambiguous impact on current construction because, while existing lines may attract new lines, they may also satiate the demand for lines of that type. Railway construction is more closely related to the rate of increase of population density than its level, as this will maintain proportionality between population density and railway provision as population changes.

The target level of station access at the start of a decade depends on the time-invariant characteristics of the parish, the amount and composition of railway infrastructure in place at the start of the decade and the population density at the start of the previous decade. Station access depends on current railway access rather than lagged railway access because stations are often built at the same time as the railways that serve them. Thus:

$$s_{it}=\alpha_{30t}+\sum_k \alpha_{3kt}x_{ik}+\sum_h \beta_{3ht}r_{hit}+\delta_{3t}y_{it-1}+u_{3it} \quad (h=1, 2, 3;\ k=1,\ldots, K) \tag{5}$$

whence

$$ds_{it}=a_{30t}+\sum_k a_{3kt}x_{ik}+\sum_h \beta_{3ht+1}dr_{hit}+\sum_h b_{3ht}r_{hit}+c_{3t}s_{it}+\delta_{3t+1}y_{it}-\delta_{3t}y_{it-1}+u_{3it+1}-u_{3it} \quad (h=1, 2, 3;\ k=1,\ldots, K) \tag{6}$$

where $a_{30t}=\alpha_{30t+1}-\alpha_{30t}$, $a_{3kt}=\alpha_{3kt+1}-\alpha_{3kt}$ and $b_{3ht}=\beta_{3ht+1}-\beta_{3ht}$. Equation (6) relates current station building to time-invariant parish characteristics, current railway building, the inherited stocks of the different types of line, inherited station access, population density at the start of the decade and population density at the start of the previous decade. Because of the way that it is defined, inherited station access always has a negative impact on station building.

There are five equations that need to be estimated. The dependent variables are current population growth, dy_{it}, (equation (2)), current investments in trunk lines, dr_{1it}, cross-country lines, dr_{2it}, and local lines, dr_{3it} (equation (4)), and station building, ds_{it} (equation (6)). Each dependent variable is influenced by a set of time-invariant parish characteristics, and by population density, railway access and station access at the start of the decade concerned.

The five equations have a recursive structure: while population growth depends on both current station building and current railway building, neither current station building nor current railway building depends on current population growth. Furthermore, while current station building depends on current railway building, current railway building does not depend on current station building. The direction of causation is therefore unambiguous in every case.

The set of five equations (2), (4), (6) may be arranged in the general form:

$$dr_{1it}=z_{21it}+e_{21it} \tag{7.1}$$

$$dr_{2it}=z_{22it}+e_{22it} \tag{7.2}$$

$$dr_{3it}=z_{23it}+e_{23it} \tag{7.3}$$

$$-b_{31t}dr_{1it}-b_{32t}dr_{2it}-b_{33t}dr_{3it}+ds_{it}=z_{3it}+e_{3it} \tag{7.4}$$

$$-b_{11t}dr_{1it}-b_{12t}dr_{2it}-b_{13t}dr_{3it}-b_{14t}ds_{it}+dy_{it}=z_{1it}+e_{1it} \tag{7.5}$$

where the terms in z reflect the impact of fixed characteristics and lagged values of time-dependent factors, while the terms in e represent uncorrelated transient random shocks; they are uncorrelated with each other and with the exogenous variables.

Since causation is unambiguous, there is no simultaneous equation bias, and each of these equations can be estimated independently of the others. Given the logarithmic transformation of population there is no reason to believe that heteroskedasticity will be a serious problem. If the time-invariant parish characteristics are correctly specified then there should be no residual spatial autocorrelation.

References

Bartholomew, John (ed.) (1887) *Gazetteer of the British Isles: Statistical and Topographical*, Edinburgh: A. and C. Black.

Bates, Alan (compiler) (1969) *Directory of Stage Coach Services, 1836*, Newton Abbot: David & Charles.

Beresford, Maurice and H.P.R. Finberg (1973) *English Medieval Boroughs: A Handlist*, Newton Abbot: David & Charles.

British Geological Survey (2012) *Geology of Britain Viewer*, www.bgs.ac.uk/discoveringGeology/geologyOfBritain/viewer, accessed 06/2012–01/2013.

Casson, Mark (2009) *The World's First Railway System*, Oxford: Oxford University Press.

Casson, Mark (2013) The determinants of local population growth: A study of Oxfordshire in the nineteenth century, *Explorations in Economic History*, 50(1), 28–45.

Cobb, Michael H. (2006) *The Railways of Great Britain: An Historical Atlas*, 2nd edn, Shepperton: Ian Allan.

Google (2013) Google Maps, www.google.com, accessed 06/2012–01/2013.

Gregory, Ian N. and Jorde Henneberg (2010) The railways, urbanisation and local demography in England and Wales (1825–1911), *Social Science History*, 34(2), 199–228.

Hoyle, Richard (2013) Private communication.

Kennedy, John (2013) Measuring the impact of the Beeching Axe, *Journal of the Railway and Canal Historical Society*, 216, 21–32.

Letters, Samantha (2010) *Gazetteer of Markets and Fairs in England and Wales to 1516, Centre for Metropolitan History*, University of London, www.history.ac.uk/cmh/gaz/gaz2 (last updated 15/07/2010), accessed 06/2012–01/2013.

Lewis, Samuel (1846) *A Topographical Dictionary of England*, London: S. Lewis.

Partida, Tracey, David Hall and Glenn Foard (2013) *An Atlas of Northamptonshire: The Medieval and Early Modern Landscape*, Oxford: Oxbow Books.

Patmore, J.A. (1966) The contraction of the network of railway passenger services in England and Wales, 1836–1962, *Transactions of the Institute of British Geographers*, 38, 105–118.

Riden, Philip (ed.) (2002) *A History of the County of Northampton, Volume V: Cleley Hundred*, London: University of London Institute of Historical Research.

Serjeantson, R.M. and R.D. Adkins (eds) (1906) *Victoria History of the County of Northampton, Vol. II*, London: James Street.

Vision of Britain (2013) www.visionofbritain.org.uk, accessed 06/2012–01/2013.

Wilson, John Marius (1870–2) *Imperial Gazetteer of England and Wales*, Edinburgh: A. Fullerton.

Wrigley, E.A. (2011) *The Early English Censuses 1801–1851*, Oxford: Oxford University Press.

8 Women's landownership in England in the nineteenth century

Janet Casson

8.1 Introduction

This chapter analyses land ownership by women during the nineteenth century using information on 23,966 plots of land. Four regions of England were selected for study, encompassing urban and rural areas involved in agriculture, industry and commerce. Studies of land ownership on this scale have never been attempted before because there was no registration of land in nineteenth-century England, and therefore no official source of information on the subject. This study exploits a new source of information on land ownership, namely the books of reference produced by railway and canal promoters, together with their accompanying maps. This wide-ranging source has never before been used systematically for historical research into land ownership.

This study reappraises conventional wisdom on women's ownership of land. It is widely believed that the subordinate status of women in the nineteenth century discouraged, or even prevented, their ownership of land. Married women in particular, it is supposed, could not own land in their own right, so that whatever land was owned by women was owned exclusively by widows and spinsters. This view ignores the extensive use of trusts, often established through marriage settlements, to protect property inherited by wives from parents and other family members. Furthermore, the legal status of women changed significantly towards the end of the century as a consequence of the Married Women's Property Acts of 1870 and 1882. A more nuanced picture of land ownership by women is therefore required.

There are two main ways of researching women's ownership of land. One is to identify women and investigate whether they owned land, and the other is to investigate land and determine how much was owned by women. It is difficult to construct a random sample using either approach, but it is much easier to construct a representative sample using the latter approach. Identifying women from registers of wills, for example, is biased towards the inclusion of wealthy women, while the information provided is weighted towards the assets that testators held shortly before their death. There is a risk that this approach may just provide a snap-shot of the land accumulated by wealthy women to provide for their old age. By contrast, this study identifies women from the land that they owned. The approach is comprehensive, within the limits set out below, and encompasses women from all social backgrounds and of all ages.

Books of reference contain information about every plot of land adjacent to a proposed canal or railway scheme, including the uses of land and buildings, together with the names of the owners, lessees and occupiers. The information for a railway scheme relates to plots up to 100 yards (91.4 m) on either side of the midpoint of a rural railway line and ten yards (9.14 m) on either side of an urban line. If any part of a plot of land came within that coverage then it was included, but no information is given about the size of any individual plot. Some plots consisted of a field, others contained several agricultural or industrial buildings and a few consisted of smaller units such as a privy or a hen house. The plot information is arranged by parish in the order that the parishes appear along the route. When a railway runs along a parish boundary information is provided for plots on both side of the boundary.

Key advantages of the source material are:

- Geographical coverage is representative because there were many railway schemes – far more schemes were promoted than were actually built. Although a majority of schemes served industrial areas and river valleys, rural areas were well-served too. Through an appropriate selection of schemes it is possible to provide a relatively comprehensive coverage of land for the specific areas selected for study.
- Coverage spans the entire century. Railways were extensively promoted from 1830, following on immediately from the construction of canals, and continued to be promoted until the end of the century. It is therefore possible to track changes in women's ownership over time.
- Many social groups are included: railway development was so widespread that it affected land owned by people from all social and economic groups, from the wealthiest landowners to the owners of small rural cottages.
- The information is reliable, because the documents had legal status and were scrutinised by both Parliament and the public.
- The information links plots, not only to their owners, but also to occupiers and lessees. This makes it possible to investigate whether women owners favoured women lessees or occupiers. By contrast, most information about nineteenth-century land holdings just links individual plots with occupiers, making such investigation difficult (Lindert, 1987).
- Multiple claims to ownership are noted, often providing information about trustees involved, including those for wills, marriage settlements, lunatics and charities. This provides valuable insights into the social, legal and institutional context of land ownership.
- Information on each parish can be linked to information on adjacent parishes through which a railway passes. This may reveal whether a woman who owned land in a given parish also owned land in adjacent parishes. Using a nation-wide source therefore provides a wider perspective on

ownership than can be obtained from purely parish-based sources such as tithe maps (Wall, 1984, p. 447).

- The source also provides full details on other types of owner, such as men and institutions. It is often assumed that the most important differences in patterns of land ownership were between men and women, but it is possible that differences between men and women on the one hand and institutions on the other may be important too.

8.2 Literature review

There are very few historical studies of women's landownership. Most previous studies have focused either on the wealthiest personal landowners, who are mainly male, or have embedded the study of land ownership within a wider study of women's property.

Nineteenth-century research into land ownership, based on information supplied by the landowners, was regarded by many contemporaries as inaccurate. Glyde's (*c.*1855, pp. 327–9) study of the ownership of 57,899 acres in Suffolk found that only 2,365 acres were owned by women, only two of whom owned more than 200 acres. The Earl of Derby criticised statistics on the ownership of land based on the 1861 Census, considering the information to be misleading because the designated landowners had been self selected and half of them were women (Sanderson and Roscoe, 1894, p. 140). In an attempt to clarify the situation the Local Government Board carried out the 1873 Return of Owners of Land on behalf of Parliament which was subsequently reassessed, modified and updated by John Bateman (Great Britain, Local Government Board, 1875; Bateman, 1883).

Rubinstein (2006, pp. 251–6) has used information from the 1873 Return, together with probate records and income tax data, to examine the concentration of wealth. He focuses on land ownership by wealthy men, but includes a list of women owners in an appendix. Collins and Havinden (2005) have examined the ownership of large estates in Berkshire and Oxfordshire, 1500–1914, supplementing information from the 1873 Return with information from Victoria County Histories and the Dictionary of National Biography. Shea and Trew (2007) have used the railway books of reference to examine whether directors and shareholders in newly promoted railways owned land on the routes through which their railway passed. Few women appear in any of these studies, and the way the studies are designed means that only wealthy women appear.

Another group of studies uses information relating to wills, and involves local or regional case studies. They provide direct evidence of middle-income women inheriting and purchasing freehold real property for housing or business purposes and purchasing real estate to provide rental income. All the studies relate to the late eighteenth or first half of the nineteenth century. Berg (1993) studied women in Birmingham and Sheffield, Davidoff and Hall (2002, p. 276) studied Birmingham and Essex; Owens (2000, p. 85) studied Stockport, and Lane (2000, pp. 176–7, 187–94) studied two Leicestershire towns. Owens (2006, pp. 23–4) also examined Bank of England will registers. In addition, Seeliger (1996) used manorial and tithe records over three centuries 1650–1900 to find evidence of land ownership in Hampshire by women from one of the lowest socio-economic groups.

The most focused study of women's land ownership is by Combs (2006). Using Death Duty, Succession Duty and Estate Duty registers, she analyses 60 lower middle-class female shopkeepers in Leeds, Liverpool and London. She tests whether the greater protection for financial assets accorded to women by the Married Women's Property Act of 1870 led these women to increase their financial holdings and reduce the amount of real property, including land, in their asset portfolios. Comparing the estates bequeathed by 30 women who married before the Act with those of 30 women who married after the Act, she finds that on average the expected reduction in real property holdings did indeed occur.

Utilising local trade directories, local newspapers, court records, diaries and private correspondence, Barker (2006, pp. 2–3, 9–10, 135–7, 167–73) showed how during the late eighteenth and early nineteenth centuries lower middle-class businesswomen in Manchester, Sheffield and Leeds operated within the constraints of the common law but independently of male family members, including their husbands. Drawing on insurance policies from the archives of the Sun Fire Office, Kay (2003) showed that female lodging house keepers combined business success with social respectability, while a parallel study of trade cards revealed women as owners and managers of a range of retail businesses (Kay, 2006, p. 152). Phillips (née Pullin) also used fire insurance records to examine London businesswomen, while Doe has examined women's role as managing owners of ships in British ports (Phillips, 2006, pp. 17–18, 26, 160–3, 257–8; Pullin, 2001, pp. 172–3; Doe, 2009, pp. 66–7, 127–48). These studies identify catering, retailing and the provision of lodgings as occupations particularly favoured by businesswomen.

A synthesis of this literature suggests a number of issues that warrant further investigation.

- It has often been claimed that women's ownership of land was small, and confined largely to widows and spinsters, although some local case studies question this view.
- Low levels of ownership, as alleged, are often attributed to common law limitations on the ability of women to enforce their property rights. If this is correct then changes in legislation in 1870 and 1882 should have led to changes in women's ownership of land. On the other hand, family customs, combined with legal trusts and the increasing application of equity law, may have meant that the common law position had less influence on property ownership than has been alleged.
- It is possible that women might not have wished to own much land, even if they had been free to do so, because they were averse to risk (Cromie and Haynes, 1988). Income from land can be uncertain and prone to fluctuations,

unlike income from fixed-interest securities such as bonds. Given a choice, therefore, women might prefer investments in government bonds rather than investments in land. Female risk aversion is, however, difficult to reconcile with the well-documented role of women speculators during the South Sea Bubble of 1720 and the Railway Mania of 1844–5 (Laurence, 2006; Carlos and Neal, 2004; Hudson, 2001, pp. 8, 14–19, 103).

- There may have been a long-term trend in women's ownership of land. Changing social attitudes may have encouraged women to seek greater economic independence, leading to an upward trend in ownership. On the other hand, the growth of capital markets, the development of joint-stock banking and the spread of provincial stock exchanges during the nineteenth century may have encouraged financial investment instead.
- Women's occupational choices may have influenced their ownership of land. A link between occupation and ownership would emerge where women wished to own the premises from which they worked rather than to rent their premises from others. In this case women's ownership of land may have been highest for plots that included residential or retail premises, or afforded opportunities for activities such as poultry keeping.
- Women's residential preferences may also have influenced the location of land they owned. While some women may have owned residential property for rental income, others may have chosen owner-occupation. Ownership of residential property may have been highest in areas in which women preferred to live, or in areas where the kind of people that women preferred as occupiers were likely to live. Literary sources, such as Gaskell's *Cranford*, suggest that women may have had preferences for market towns or genteel suburban areas (Gaskell, 1853, pp. 175–6).

8.3 The measurement of key variables using the source material

Railway schemes covered the entire country, and by 1900 almost every town and village in England was within five miles of a railway station. Four areas of the country were selected for detailed study – two in the North and two in the South: Oxfordshire and surrounding counties are predominately agricultural; County Durham is a mixture of agriculture and mining; West Yorkshire is mainly industrial, while the London area (Middlesex and surrounding counties) is mainly administrative and commercial.

The analysis is highly disaggregated. It focuses on individual plots of land, as identified by professional railways surveyors. A typical rural plot would be a field and a typical urban plot a house and garden. Land was sub-divided to the point where each plot could reasonably be owned by a different person. This avoids the problems associated with higher levels of aggregation, at which units of land (e.g. farms or estates) may have sub-divided ownership and multiple uses. The books of reference list plots sequentially, beginning at one end of the line and finishing at the other. Lines were selected in order to meet a target of a minimum of 400 plots within each region in each decade. Subject to meeting this target, lines were selected to provide a representative geographical coverage of the region over the century as a whole.

There is no evidence that railway promoters discriminated for or against women-owned plots. Railways needed to be built as straight and level as possible, so that once the origin and terminus had been fixed the route could not easily be varied. From the 1840s onwards not even aristocrats could prevent the building of a railway through their land. There is no evidence that women owners were regarded as a 'pushover' where railway construction was concerned: indeed, two of the most obstinate objectors to the pioneering Liverpool and Manchester Railway of 1830 were a pair of women housing speculators (House of Commons, 1825, pp. 330–2, 357, 364–5).

Once a scheme had been selected for the study, all available plots were used until the target number of plots had been met. Targets were exceeded in many cases in order to include all plots on a particular line. The database was constructed in SPSS using 23,966 rows and 64 columns. The uses of each plot were recorded in full. Additionally, for all plots where a woman was involved in ownership or leasing, the names of all owners, lessees and occupiers were recorded. The same information was also recorded for a random 10 per cent sample of all plots in which women owners were not involved. This sample was then used as a control group for women owners. The names of individuals were recorded on a separate qualitative database as were the names of all the parishes, townships, etc.

Overall, five main types of variable were constructed for purposes of statistical analysis: they concern

- Ownership
- Plot use
- Parish characteristics
- Time-dependence
- Interaction (e.g. linking plot uses and parish characteristics).

All relevant information was quantified in binary form; thus ownership characteristics, plot uses and time-dependent factors were all expressed as dummy variables. Some of the qualitative information was also used for case study analysis, but the case studies are not discussed in this chapter.

Women's ownership. The principal ownership variable is a binary dummy variable that takes the value of one if a woman was involved at all in the ownership of a plot and is zero otherwise. Women owners were further sub-divided according to their status and to the number and type of their co-owners (see below).

Plot use. Plot use was categorised into 17 main uses: housing, retail, agricultural buildings, all kinds of agricultural land, wood, domestic land, waste land and/or common, water, roads, utilities, mining, quarrying, railway, canal, public building, industrial and miscellaneous. A further seven subsidiary categories were used where the source material provided a sufficient number of observations (normally ten per region) to make them viable for statistical purposes; housing was sub-divided into house, cottage, tenement or court, and agricultural land into arable, meadow and pasture. Some plots had multiple uses.

Parish characteristics. The inclusion of parish characteristics in the study makes it possible to investigate whether women revealed a preference for investing in particular kinds of locality. The five parish characteristics are all continuous variables:

- Population density: this is linked to urbanisation (see below); where density is very high it may be linked to poverty and social deprivation;
- Population per house: a possible indicator of over-crowding; it may, however, reflect the sub-division of large houses into separate dwellings;
- Proportion of the population aged below 20 years: the only indicator of age distribution that is available;
- Proportion of the population born outside the county; a measure of immigration into the county;
- Proportion of females in the population: by comparing this ratio with the proportion of land in the parish owned by women it is possible to assess whether high women's ownership of land is linked to a high proportion of female residents in a parish.

The information to calculate the parish variables was obtained from the Enumeration Abstract in the Abstract of the Answers and Returns, of 1843 (Parliamentary Papers, 1843 (496)). This Abstract was based on the 1841 Census which immediately preceded the Railway Mania when many of the schemes used in the study were first proposed, and it contains demographic information not included in subsequent census reports. The use of parish information from a specific year avoids confusion between effects caused by spatial variation and effects caused by temporal variation. Linking parish data to information from the books of reference requires care in cases where one of the sources sub-divides parishes into townships, chapelries, etc., and the other does not.

An urban area is defined as a parish or group of adjacent parishes that have a high population density, a total population exceeding 15,000 and a wide range of cultural amenities. Urban areas correspond to towns and cities with extensive retail, legal, banking and recreational facilities. Urbanisation is captured by two dummy variables. Urban centrality takes a value of one if some part of a parish is within one mile of the centre and is zero otherwise, while urban peripherality is defined analogously for distances of more than one and less than five miles. In London, for example, where St Paul's Cathedral was designated as the centre of the city, East End parishes with very high population but limited cultural amenities were not identified as separate urban areas. The 'Phillimore Atlas' was used to determine which London parishes were central or peripheral (Humphery-Smith, 2003).

In many nineteenth-century parishes land was owned by a small number of people and this could affect the ability of outsiders to purchase plots. Although there are indications in the books of reference that plot ownership in some parishes was restricted to a few people it was impossible to determine if these were 'close' parishes as defined by Mills and others (Lane, 2000, p. 188; for a full discussion see Mills, 1980, and Banks, 1988). Lords of the manor sometimes owned a high proportion of plots in a single parish. The presence of a female lord of the manor could therefore generate high levels of women's ownership in certain parishes. Such women might also exercise considerable influence over the occupation of other land in the parish. For many parishes, however, it proved difficult to determine whether there was a female lord of the manor at the time the book of reference was compiled. Manorial ownership was in decline throughout the nineteenth century and, perhaps as a result, records are incomplete. The analysis described in this chapter was replicated using a smaller sample of parishes for which manorial information is available. The results confirmed that female manorial ownership was an important factor in certain parishes, but the wider implications were inconclusive because of the smaller number of parishes involved.[1]

Time-dependent variables. The date of the railway scheme was included to assess long term time trends. A Railway Mania dummy variable was designed to identify if there was an upward or downward spike in the ownership of plots by women during the speculative period 1843–6. Two other dummy variables, linked to the Married Women's Property Acts of 1870 and 1882, were included to allow for step changes in the propensity for women to own plots arising from the changes in their legal position. Unlike the Railway Mania dummy, the aim was not to identify spikes at the times of the Acts, but rather continuing legacies of the Acts. Since each Act has its own dummy variable, the study can analyse the period 1870–82 between the two Acts as well as compare and contrast the periods before and after them.

It should be emphasised that a change in women's ownership at the time that the values of a dummy variable changes cannot be unambiguously attributed to any one specific event that occurred at that date. It is always possible that some other event at about the same time was responsible for the change instead. To identify the appropriate cause of change it is normally necessary to have additional information from women's diaries or letters, and in particular their correspondence with bankers and solicitors. This is a special case of the general problem of 'omitted variables' in statistical inference, but it is particularly relevant to the use of time-dependent dummy variables.

Cumulative railway mileage is a continuous time-dependent variable that tracks the growth of the railway system. It is based on route mileage and includes both single-track and double-track lines. This variable is useful in accounting for the increasing proportion of plots that are used by other railway companies whose lines are intersected by a proposed scheme.[2]

Interaction variables. An interaction variable was introduced, generated by multiplying together two dummy variables: housing use and urbanisation. The variable is useful in testing the hypothesis that women own land in urban areas because they wish to live there as well (living either in their own property or some other property nearby).

The variables are summarised in Table 8.1.

Table 8.1 Alphabetical list of acronyms for plot uses, parish and ownership characteristics and other variables

Acronym	*Meaning and examples from plots*
AGBL	agricultural buildings: barn, cart shed, hen cote, piggery, slaughter house, stables
AGLD	agricultural land general: arable, common field, field, grass, market garden, meadow, nursery ground, open field, pasture, potato ground
COTT	cottage specified
CRWD	cumulative railway development, i.e. the miles of railway line laid by the year of the scheme from which plot taken
DATE	date of the railway, canal or turnpike road scheme from which the plot is taken
DOML	domestic land: allotment, garden, greenhouse, orchard, paddock, park, physic garden, greenhouse, pleasure garden, orchard, shed, shrubbery
FMWA	First Married Women's Property Act. Plot identified as pre-1871 or 1871 and later
HSEG	general housing: cottage, farm house, homestead, house, lodge, manor house, mansion, tenement, toll house
HOUS	house specified: farm house, homestead, house, lodge, manor house, mansion, toll house
INDS	industrial: airshafts, brewery, brickworks, carpenter's, cattle market, chimney stack, cider house, coal yard, counting house, drying kiln, forge, joiners, harbour, lime kiln, malt house, mill–corn/water, pottery, powder magazine, printers, rope walk, sawpit, stone yard, warehouse, weighing house, wharf, wheelwright, workshop
MIQU	combined mining (coal or metal), quarrying and gravel-pit variable
MISC	miscellaneous independent units: archway, building land, city walls (York), croft, enclosure, garth/yard, ground – cricket/drying/football/recreation/running, hovel, Hyde Park, hydraulic ram, island/eyot, Park, Knightsbridge barracks, kitchen, land unspecified, privies, proposed railway site, shuttle, theatre (being constructed), wash house/kitchen
PBTW	percentage of population in parish/township below 20 years of age – young people (including both children and young adults)
PFEM	percentage of females within parish/township
PIMM	percentage of immigrants, people not born in the county, in the parish/township
POPD	population density of parish or township, i.e. average number of people per acre
PPHS	population per house in parish or township, i.e. average number of people per house
RAIL	railway: tramway, wagon way
RETA	retail: bake house, beer house, brew house, photographic studio, public house, riding school, shop, studio, surgery, tea gardens
RMAN	Railway Mania scheme 1843–6
ROAD	roads: bridge, landing stage, mews, passage, paths – bridle & foot, roads – occupational, turnpike, private
SMWA	Second Married Women's Property Act. Plot identified as pre-1882 or 1883 or later
UCEN	urban central, plot within one mile of the centre of a town of 15,000 inhabitants, modified for London parishes
UDIF	urban differentiation, plot has a general housing use and is either urban central or peripheral
UPER	urban peripheral, plot over one mile but less than five miles from the centre of a of town 15,000 inhabitants, modified for London parishes
UTIL	utilities: gas, gasometer, retort house, sewers, telegraph posts & wires, water
WAST	waste: waste specified, common, common balk, cow common, fell, heath, moorland, upland
WATR	water: culvert, ditch, drain, foreshore, fountain, marsh, mill dam/pond, pond, pump, reservoir, sea, septic tank, stream, river, water closet, watercress beds, weir, well
WOOD	wood: coppice, copse, osier/willow bed, plantation, underwood, wood

8.4 Modelling the propensity for women to own specific types of land

Given the research questions set out in section 8.2, and the availability of data as set out in section 8.3, the natural way to proceed is to relate women's ownership of land to a range of plot characteristics. The characteristics of any plot encompass its use, the demographic characteristics of the parish in which it is located, the values of the time-dependent variables at the date it was recorded and the corresponding value of the interaction variable.

Some of the research questions can be answered fairly simply using descriptive statistics. A good indication of whether women's ownership of land was substantial can be obtained from the overall percentage of plots owned by women. Similarly, this use of trends can be addressed fairly easily by comparing ownership in the earlier decades with the ownership in later decades.

To obtain a more nuanced view, however, it is appropriate to conduct a regression analysis (Wooldridge, 2006, p. 862). The dependent variable is a binary variable, indicating whether a woman was involved in the ownership of a given plot; the independent or explanatory variables are the plot characteristics. Using regression analysis it is possible to allow for changes in women's ownership driven by land becoming more residential, commercial and industrial over time, and parishes became more urban (or suburban in many cases).

If women have a high degree of personal agency, they can choose whatever type of land they wish to own. They may have distinctive preferences for certain plot uses and certain locations, and these preferences may change as cultural attitudes or their legal position alter over time. In practice, however, women face competition for land from both men and institutions (e.g. crown, church, colleges, charities). Women need sufficient wealth to out-bid men and institutions.

Within the land market there will be transaction costs, but over time there will be a tendency towards equilibrium. In a market equilibrium plots will be owned by those who value them most highly. In equilibrium women will tend to own the types of land that, relative to men and institutions of similar wealth, they most prefer. This is true even where inherited land is concerned, provided that property is not entailed. A woman who does not particularly value land that she has inherited can sell out and invest the proceeds in financial assets instead.

The regression model outlined above assumes that the explanatory variables are exogenous: while these variables affect women's ownership, women's ownership does not affect them. Given that there is a fixed stock of land, and that women had little or no control over legislation in the nineteenth century, it is evident that the time-dependent variables are indeed exogenous. Furthermore, apart from female manorial ownership, it seems unlikely that female ownership would significantly impinge on parish demography. The main concern must be plot uses, where it is conceivable that the uses of land might change if women had a significant presence in the land market. No evidence was found that women changed the uses of plots they purchased or inherited, or that plot uses were changed in order to attract women buyers.

Because the dependent variable is binary, it is appropriate to use probit or logit regression, and logit was selected for the purposes of this study. The conventional reason for rejecting ordinary least squares (OLS) regression is that it may predict that the value of the dependent variable is less than zero or exceeds one. A much stronger reason for using logit regression in the present case, however, is that the latent variable used in the logit model corresponds to an equilibrium price premium for land, as explained below.

A separate regression was estimated for each region. The hypothesis that the regional regressions were based on similar populations was statistically rejected. This result is plausible, because it suggests that each of the regions had a distinctive land market with a distinctive price structure. Because of the large number of observations in each region, statistically significant results were obtained even though the overall proportion of the variation in women's ownership explained is modest.

Let

- x_{hij} be the value of the ith characteristic relating to the jth plot in the hth region (time-dependent variables have the same value for every plot in a given railway scheme);
- y_{hj} be a binary variable indicating whether the jth plot of land in region h has a woman owner; and
- z_{hj} be the women's price premium on the jth plot of land in region h; it measures the excess of the maximum price that a woman in prepared to pay for the plot in question over the maximum that a non-woman (man or institution) is prepared to pay for it.

The women's price premium varies according to the plot characteristics. Suppose that plot characteristics impact additively; then

$$z_{hj} = a_h + \textstyle\sum_i b_{hi} x_{hij} + u_{hj} \quad (i = 1, \ldots, N) \tag{1}$$

where a_h is a regional constant setting a base-line level of women's ownership; b_{hi} measures the impact the ith plot characteristic on the price premium in the hth region; and u_{hj} is a random variable influencing the price premium of the jth plot in the hth region. This variable includes all the unobservable factors that affect the premium.

All the u_{hj} are independently distributed and follow the same logistic distribution with cumulative distribution function $F(u)$.

A woman will own the plot if and only if they out-bid their non-woman rivals:

$$y_{hj} = 1 \text{ if } z_{hj} > 0 \tag{2}$$

$$= 0 \text{ otherwise}$$

It may be deduced that the probability of a woman owning the jth plot in the hth region is $F(z_{hj})$. This makes it possible to estimate the parameters a_h, b_{hi} by maximum likelihood.

By comparison, under OLS it is assumed that women's ownership of a plot is determined directly by plot characteristics without the intervention of price competition:

$$y_{hj} = z_{hj} \tag{3}$$

The difference between logit and OLS is that logit uses a threshold approach in which the price mechanism creates a tipping point where non-women owners are replaced by women owners, whereas OLS assumes a continuous linear relationship between plot characteristics and women's ownership. The logit approach has a stronger basis in the economic theory, as explained above. Unlike OLS coefficients, however, logit coefficients cannot be directly interpreted as measuring marginal impacts.

Another limitation of logit is a practical one. While OLS estimates can be derived through linear operations on the data set, logit involves non-linear computations which may fail to converge – which sometimes occurs in this study because of the large number of dummy variables used. While logit is the first choice estimation method, OLS is a useful backup; comparing OLS and logit results when logit estimation converges suggests that OLS affords a good approximation to logit in the present case. Although the values of the estimated coefficients are not directly comparable, their signs and significance are.

The goodness of fit of a logit regression is sometimes assessed by the percentage of cases in which a correct prediction of the value of the dependent variable (i.e. women's ownership) is achieved. However, the predictions generated by this method are heavily influenced by the critical level of the probability that is used to separate the binary outcomes in the predictive exercise. If the actual probability of a positive outcome (i.e. ownership by a woman) is, say, normally less than 20 per cent, and the critical level used for prediction is 50 per cent, then it is possible that no cases of women's ownership will be correctly predicted; on the other hand, the overall goodness of fit may appear to be good because almost every null outcome will be correctly predicted. A more realistic assessment can be achieved by matching the critical probability more closely to the actual probability; in the case above this would increase the probability of correctly predicting a positive outcome, but also increase the probability of wrongly predicting a negative outcome, almost certainly making the overall performance apparently worse. For this reason, predictive success is not used to assess goodness of fit in this study (Tunali, 1986; Wooldridge, 2006, pp. 589, 867).

8.5 The results

Descriptive statistics

Descriptive statistics reveal that on average women were involved in the ownership of 12.4 per cent of plots. However, there was considerable regional variation with women involved in the ownership of 14.7 per cent of plots in Oxfordshire, around 12 per cent in Yorkshire and Durham and only 8.9 per cent in London.[3]

Cross-tabulations indicate that women's ownership varied systematically according to plot use, and that the pattern of variation differed somewhat between regions (see Table 8.2). In order to clarify the extent to which women in a given region specialised in owning certain types of plot an index of specialisation was constructed. This index is the ratio of the propensity of women to own a plot in a given use (i.e. the proportion of all plots with that use that were owned by women) to the average propensity of women to own any plot (i.e. the proportion of all plots owned by women whatever their use). Separate indices were calculated for each region and are shown in Table 8.3. An index of 1.0 indicates no specialisation, above 1.0 indicates specialisation and below 1.0 indicates avoidance of the plot use.

In Oxfordshire, women favoured houses, cottages, domestic land, retail and industrial plots. Of the 17 Oxfordshire retail plots 13 were in Cheltenham in 1845 and included unspecified shops and a bake house. Two women were owner-occupiers engaged in trade: Maria White who owned a house, shop and brewery and Mary Hill who owned a shop, house and passage.

Ownership of agricultural buildings and retail plots was popular with Yorkshire women, and London women favoured cottages, retail plots and utilities. London women also favoured agricultural buildings because of their need to stable horses, and some women also owned quarries. Durham women preferred houses, domestic land and railways (the railways were connected with collieries that the women owned).

To examine the relation between women's ownership and parish characteristics, Pearson zero-order correlations were calculated (see Table 8.4). Parishes with a high population density attracted women owners from every region except Durham. A large number of residents per house was a significant deterrent to ownership in Yorkshire and Durham. Oxfordshire and Yorkshire women avoided parishes with high numbers of young people but were attracted to parishes with high levels of immigration and those with high percentages of females. Oxfordshire women were attracted to urban central parishes and Yorkshire women were attracted to urban peripheral parishes. However, such peripheral parishes were avoided by Durham and London women.

Table 8.5 examines the status of women owners. In all four regions at least 39 per cent of women owned land either on their own or with other women. In other respects there were significant regional differences, however. In the north (Yorkshire and Durham) over 20 per cent of women owners co-owned with specified

Table 8.2 To what extent do certain types of plot usage attract women more than men or institutions as owners? Cross-tabulation of women's involvement in plot ownership and plot use across all four regions

Region	*Oxfordshire*		*Yorkshire*		*Durham*		*London*	
Plot uses	*Plots (%)*	*Plots in given use with women owners (%)*	*Plots (%)*	*Plots in given use with women owners (%)*	*Plots (%)*	*Plots in given use with women owners (%)*	*Plots (%)*	*Plots in given use with women owners (%)*
HSEG	11.5	20.4	23.0	12.6	5.8	14.4	41.9	9.6
HOUS	7.5	19.0	17.6	11.7	4.6	18.0	38.8	8.8
COTT	3.9	22.7	4.7	13.5	1.1	1.9	3.0	19.2
RETA	0.8	29.3	1.5	15.2	0.5	0.0	5.1	11.5
AGBL	4.3	16.4	4.5	18.1	6.2	11.3	6.0	14.8
AGLD	58.2	12.4	50.2	13.3	62.5	11.1	28.3	8.5
WOOD	3.0	10.8	1.5	9.5	4.9	12.1	1.1	6.7
DOML	19.1	23.8	11.4	11.7	11.7	16.3	25.1	7.0
WAST	1.1	19.0	0.7	11.1	1.3	15.6	0.7	0.0
WATR	11.8	12.5	8.2	14.3	9.7	13.9	6.5	8.3
ROAD	21.7	11.1	12.8	11.4	14.3	11.1	12.3	5.1
UTIL	0.5	3.0	0.3	4.3	0.7	6.3	1.8	11.8
RAIL	0.6	2.4	0.8	5.1	1.8	14.0	0.9	5.1
INDS	1.6	19.5	2.3	11.6	1.1	13.2	3.7	6.4
MIQU	0.6	14.3	0.5	3.0	0.5	7.7	0.1	25.0
All plots		14.7		12.4		12.0		8.9

Note
Oxfordshire includes parts of the counties of Berkshire, Buckinghamshire, Gloucestershire, Northamptonshire, Warwickshire and Worcestershire. Yorkshire includes both the East and West Ridings. Durham includes parts of the North Riding of Yorkshire and Northumberland. London includes parts of Essex, Hertfordshire, Kent, Middlesex and Surrey.

Table 8.3 Coefficients of specialisation by women in the ownership of plots according to plot use

Explanatory variables: plot uses	*Region*			
	Oxfordshire	*Yorkshire*	*Durham*	*London*
	Specialisation indices			
HSEG	1.4	1.0	1.2	1.1
HOUS	1.3	0.9	1.5	1.0
COTT	1.5	1.1	0.2	2.2
RETA	2.0	1.2	0.0	1.3
AGBL	1.1	1.5	0.9	1.7
AGLD	0.8	1.1	0.9	0.9
WOOD	0.7	0.8	1.0	0.8
DOML	1.6	0.9	1.4	0.8
WAST	1.3	0.9	1.3	0.0
WATR	0.9	1.1	1.2	0.9
ROAD	0.8	0.9	0.9	0.6
UTIL	0.2	0.3	0.5	1.3
RAIL	0.2	0.4	1.2	0.6
INDS	1.3	0.9	1.1	0.7
MIQU	1.0	0.2	0.6	2.8

Table 8.4 Pearson zero-order correlations between women’s ownership of plots and parish characteristics

Parish characteristics	*Regions*			
	Oxfordshire	*Yorkshire*	*Durham*	*London*
POPD	.088	.030	–.041	.078
	(.000)	(.008)	(.005)	(.000)
PPHS	.012	–.037	–.050	–.008
	(.286)	(.001)	(.000)	(.594)
PBTW	–.053	–.049	–.002	.022
	(.000)	(.000)	(.876)	(.154)
PIMM	.126	.032	.004	–.017
	(.000)	(.005)	(.763)	(.253)
PFEM	.113	.059	.031	–.028
	(.000)	(.000)	(.032)	(.066)
UCEN	.071	.014	.018	.016
	(.000)	(.226)	(.212)	(.295)
UPER	.000	.030	–.071	–.039
	(.978)	(.010)	(.000)	(.011)

husbands, and more than 23 per cent co-owned with a man of the same name, who could have been a husband, father, brother or more distant relative. By contrast, in the south (Oxfordshire and London) sole-ownership or ownership with other women was more common, and co-ownership with husbands or men of the

Table 8.5 Comparative regional analysis of the ownership of plots by subgroups of women owners expressed as a percentage of all plots owned by women

Parish characteristics	*Regions*			
	Oxfordshire	*Yorkshire*	*Durham*	*London*
Co-ownership structure				
Sole owner or women co-owners	57.8	39.4	39.4	55.1
Men of the same surname	11.5	31.5	23.3	16.8
Husband	8.2	21.5	24.3	1.8
Husband with others	7.8	16.8	10.8	1.8
Trust	15.1	10.7	13.0	20.5
Title				
'Mrs'	8.6	7.4	6.8	24.9
'Widow'	2.1	1.2	0.0	0.3

same name much smaller. This may reflect a regional cultural divide. There are also indications of a divide between London and the provinces, as co-ownership with males was very low in London, even though the use of the title 'Mrs' was relatively common.

It is possible that the high percentage of the 'Mrs' group in London arose because such females were widowed and moved to the capital to avail themselves of the increased cultural and social opportunities, and specified the title as a mark of their respectability. In Oxfordshire the use of the title 'Mrs' is more ambivalent, but nevertheless suggests the independence of the women in that region too.

London had the greatest percentage of 'trust' plots, all of which were post 1845. Twelve plots were subject to marriage trusts, but the vast majority were trusts set up to administer an estate after death. It is possible that the large number of trusts found in the London sample reflected the fact that, as the metropolis, it afforded easy access to legal advice. According to Stebbings (2002, pp. 4–5), trusts, which were unregulated and mostly private, were widely used, even by working class people, in the nineteenth century and it was assumed at the time that about a tenth of all property was subject to a trust. According to Morris (1994, pp. 176–7; 2005, p. 113), most trusts were via a will with the property left in trust for a widow or daughter and there were three trustees, usually male. Morris has examples of Leeds men setting up such trusts in the 1820s and 1830s for widows and daughters, the aim being to protect a business so that a minor could inherit later or to prevent money from falling into the hands of future husbands. The complexities of trusts are fully explored by Okin (1983) and Staves (1990, pp. 42–55), who highlight the situation of a husband being a trustee to his wife's trust.

Regressions

The regressions for each region were estimated separately and the results reported in Table 8.6 confirm that there was considerable regional heterogeneity.

Table 8.6 Comparative regional analysis of the propensity of women to be involved in the ownership of plots, analysed according to plot use, parish characteristics and time

Region	*Oxfordshire*	*Yorkshire*	*Durham*	*London*
Type of regression	*Logit*	*Logit*	*OLS*	*OLS*
Percentage of plots with women involved in ownership	14.7	12.4	12.0	8.9
Explanatory variables:				
HSEG	–.178 (.097)	.241 (.023)		.183 (.000)
RETA	.542 (.068)		–.131 (.044)	.038 (.058)
AGBL		.495 (.001)		.056 (.003)
AGLD		.514 (.000)		
DOML	.605 (.000)		.046 (.003)	
WAST	.700 (.017)			–.105 (.040)
WATR	–.323 (.006)			–.031 (.094)
ROAD	–.357 (.000)		–.027 (.045)	–.040 (.004)
RAIL	–2.125 (.037)			
INDS				–.041 (.076)
POPD			–.008 (.001)	.001 (.000)
PPHS	–.200 (.012)	–.397 (.000)		–.022 (.000)
PBTW	.029 (.019)			
PIMM	.039 (.000)	.043 (.000)		
UCEN		–.179 (.000)		
UPER			–.096 (.000)	–.027 (.036)
UDIF				–.201 (.000)
DATE	.012 (.000)	.014 (.000)	.004 (.000)	.001 (.003)
RMAN		.339 (.000)		.031 (.054)
FMWA	–.263 (.037)	–1.921 (.000)	–.225 (.000)	.049 (.002)
SMWA	–.645 (.000)	.817 (.007)		–.117 (.000)
Constant	–24.356 (.000)	–26.561 (.000)	–8.055 (.000)	–1.470 (.008)
No. observations	7299	7539	4861	4267
R^2			.049	.045
Adj. R^2			.048	.042
F			35.637 (0.000)	13.501 (0.000)
χ^2	253.987 (0.000)	167.503 (0.000)		
CSR^2	.034	.022		
NR^2	.060	.042		

Note
Significance levels
(p-values) are in brackets underneath estimated coefficients.

The analysis focuses on the signs and significance of the coefficients and not on their magnitudes. Logit regressions only converged in two regions, so OLS regressions are reported for the other cases. Running both types of regressions in the first two regions showed that, although the coefficients are not directly comparable, the OLS results approximate the logit results so far as sign and significance are concerned. To facilitate interpretation, insignificant variables were eliminated backwards using a threshold of 10 per cent significance. The large number of observations means that the overall regressions are significant even though the proportion of the variation in women's ownership that is explained (as measured by the R^2 or pseudo R^2) is relatively small.

The location of plots was found to be important not only at the regional level but also at the local parish level, especially with regard to ownership. The overall conclusion was that women's ownership of plots was concentrated in areas of relative gentility, with good local amenities but no problems with overcrowding or large family groups. The regional patterns suggest that women adapted their ownership to local economic conditions rather than everywhere favouring plots of land with the same uses.

The relationship between housing uses and the types of parishes in which suitable housing is located make it more difficult to disentangle the specific role of housing. Women favoured the ownership of housing in Yorkshire and London, though not in Oxfordshire, where the coefficient was negative and only just significant at 10 per cent. However, Oxfordshire women did have a strong preference for certain parish characteristics which suggest that quality of housing was important to them, in particular avoiding parishes with a high number of people per house, a trait which was shared by women owners in several regions. High levels of immigration into a parish also encouraged ownership by women in Oxfordshire and Yorkshire.

During the Railway Mania of 1843–6 there was a spike in plot ownership by women from Yorkshire and London, suggesting that women in those regions may have been less risk averse than suggested in the literature and were possibly speculating in land which could increase in value if purchased by railway companies (Green and Owens, 2003; Maltby and Rutterford, 2006).

After 1870, the date of the first Married Women's Property Act, women's plot ownership in Oxfordshire, Yorkshire and Durham declined. London women owners however were encouraged into plot ownership after the first Act, a finding that differs from that of Combs, who included London women in her study. The results show that the passing of the 1882 Act had an ambiguous effect on women's plot ownership. In Oxfordshire and London the impact was significantly negative, but in Yorkshire it was significantly positive and in Durham it was of no significance. Both Acts were therefore important and on balance their effect was to reduce women's ownership of plots of land.

The time trend was the most consistent of all the variables, being significantly positive for women owners across all regions. It indicates a multiregional, and potentially nation-wide, long-term trend towards greater female involvement in the ownership of property, and provides the context within which the other changes need to be understood. In particular the step changes in response to the two Married Women's Property Acts need to be considered in the context of this overall steadily rising trend.

8.6 Conclusions and implications for future work

Overall, the research highlights the importance of women owners of land and real property in the nineteenth century. It revises the view that legal and social constraints confined women's ownership of land to wealthy widows and spinsters and shows that ownership was far more widespread than has been supposed.

There are a number of ways in which the study should be developed. The interpretation offered should be investigated for other geographical regions since a major advantage of the source is that it enables the study to be replicated.

Case studies of some of the women landowners should be carried out. This would extend information about women owner-occupiers and absentee owners. A similar investigation could be carried out to see if women involved in the ownership of land were related to members of provisional committees promoting the various railway schemes used in this study. Both studies would involve extending the research to include census records and the records of the railway companies. It might even be possible to check a sample of ownership details against the papers of a solicitor who acted in the land purchase for a railway company but this would only be possible if such business records had been deposited with a county record office or the National Archives. Such case studies would also make it possible to investigate if any of the women found in this study were among those from Leeds studied by Barker (2006) or Morris (2005).

The information about trusts could be subdivided into the different kinds – specified marriage, general husband and wife, post death, committee for lunatic, etc. – to provide more information. However, the numbers of plots with trusts is comparatively small overall and this would probably only provide statistically viable information if the sample was considerably extended. Extending the study to other regions and/or a greater number of plots per region would also enable the manorial study to be extended.

Acknowledgements

I am grateful to my supervisor Jane Humphries for guidance and criticism, and to Chris Day and Mark Smith for their encouragement. The research could not have been completed without valuable advice from various county record offices. An earlier version of this chapter was presented at the Economic History Society Annual Conference at Cambridge, 2011, and at various seminars and workshops in Oxford.

Notes

1 The books of reference often gave manorial ownership. If not, the relevant Victoria County Histories were consulted along with local directories and histories, but sometimes it could not be established; see Friar (2002, pp. 258–9).

2 Sources included Mitchell and Deane (1962), p. 225; Parliamentary Papers, 1852, (21), p. 393; Parliamentary Papers, 1854, p. vi; Parliamentary Papers, 1900, pp. iv–v. The earliest railways in this study were isolated local lines so the Parliamentary statistics are adequate, even though the first railway listed is 1825 and the early horse-drawn tramways were only added to the statistics when they were upgraded to normal railway status.

3 The examples given in this section are all drawn from Casson (2013).

References

Banks, Sarah (1988) Nineteenth-century scandal or twentieth-century model? A new look at 'Open' and 'Close' parishes, *Economic History Review*, New Series, 41(1), 51–73.

Barker, Hannah (2006) *The Business of Women: Female Enterprise and Urban Development in Northern England 1760–1830*, Oxford: Oxford University Press.

Bateman, John (1883) *The Great Landowners of Great Britain and Ireland* (repr. 1971, ed. David Spring), Leicester: Leicester University Press.

Berg, Maxine (1993) Women's property and the Industrial Revolution, *Journal of Interdisciplinary History*, 24(2), 233–50.

Carlos, Ann M. and Larry Neal (2004) Women investors in early capital markets, 1720–1725, *Financial History Review*, 11(2), 197–224.

Casson, Janet Penelope (2013) *Women and Property: A study of women as owners, lessors, and lessees of plots of land in England during the nineteenth century as revealed by the land surveys carried out by the railway, canal and turnpike companies*, Oxford University, DPhil submission.

Collins, Ted and Michael Havinden (2005) Long term trends in landownership, 1500–1914: Berkshire and Oxfordshire, *Oxonensia*, 70, 27–49.

Combs, Mary Beth (2006) Cui Bono? The 1870 British Married Women's Property Act, bargaining power, and the distribution of resource within marriage, *Feminist Economics*, 12(1–2), 51–83.

Cromie, S. and Haynes, J. (1988) Towards a typology of female entrepreneurs, *Sociological Review*, 36(1), 99–107.

Davidoff, Leonore and Catherine Hall (2002) *Family Fortunes: Men and Women of the English Middle Class 1780–1850*, rev. edn, Abingdon: Routledge.

Doe, Helen (2009) *Enterprising Women and Shipping in the Nineteenth Century*, Woodbridge: Boydell Press.

Friar, Stephen (2002) *The Local History Companion*, Stroud: Sutton.

Gaskell, Elizabeth (1853) *Cranford* (repr. 2007), London: Bloomsbury.

Glyde, John, Jn. (*c.*1855) *Suffolk in Nineteenth Century: Physical, Social, Moral, Religion and Industrial*, London: Simpkin Marshall.

Great Britain, Local Government Board (1875) *England and Wales (Exclusive of the Metropolis): Return of Owners of Land, 1873*, 2 Vols, London: Eyre and Spottiswoode for HMSO.

Green, David R. and Alastair Owens (2003) Gentlewomanly capitalism? Spinsters, widows, and wealth holding in England and Wales, *c.*1800–1860, *Economic History Review*, New series, 56(3), 510–36.

House of Commons (1825) *Proceedings of the Committee of the House of Commons on the Liverpool and Manchester Railroad Bill*, Session 1825, London: Thomas Davidson.

Hudson, Sarah (2001) *Attitudes to Investment Risk amongst West-Midland Canal and Railway Company Investors, 1760–1850*, Coventry: University of Warwick, PhD thesis.

Humphery-Smith, Cecil R. (ed.) (2003) *The Phillimore Atlas and Index of Parish Registers*, 3rd edn, Chichester: Phillimore.

Kay, Alison C. (2003) A little enterprise of her own: Lodging-house keeping and the accommodation business in nineteenth-century London, *London Journal*, 28(2), 41–53.

Kay, Alison C. (2006) Retailing respectability and the independent woman in nineteenth-century London, in R. Beachy, B. Craig and A. Owens (eds) *Women Business and Finance in Nineteenth-century Europe*, Oxford: Berg, pp. 152–66.

Lane, Penelope (2000) Women, property and inheritance: Wealth creation and income generation in small English towns, 1750–1835, in Jon Stobart and Alastair Owens (eds) *Urban Fortunes: Property and Inheritance in the Town 1700–1900*, Aldershot: Ashgate, pp. 172–94.

Laurence, Anne (2006) Women investors, 'That nasty South Sea affair' and the rage to speculate in early eighteenth-century England, *Accounting, Business & Financial History*, 16(2), 245–64.

Lindert, Peter, H. (1987) Who owned Victorian England?: The debate over landed wealth and inequality, *Agricultural History*, 61(4), 25–51.

Maltby, Josephine and Janette Rutterford (2006) 'She possessed her own fortune': Women investors from the late nineteenth century, *Business History*, 48(2), 220–53.

Mills, Dennis R. (1980) *Lord and Peasant in Nineteenth Century Britain*, London: Croom Helm.

Mitchell, B. R., and Phyllis Deane (1962) *Abstract of British Historical Statistics*, Cambridge: Cambridge University Press.

Morris, R. J. (1994) Men, women, and property: The reform of the Married Women's Property Act 1870, in F. M. L. Thompson (ed.), *Landowners, Capitalists, and Entrepreneurs: Essays for Sir John Habakkuk*, Oxford: Clarendon Press, pp. 171–191.

Morris, R. J. (2005) *Men, Women and Property in England, 1780–1870: A Social and Economic History amongst the Leeds Middle Classes*, Cambridge: Cambridge University Press.

Okin, Susan Moller (1983) Patriarchy and married women's property in England: Questions on some current views, *Eighteenth-Century Studies*, 17(2), 121–38.

Owens, Alastair (2000) Property, will making and estate disposal in an industrial town, 1800–1857, in Jon Stobart and Alastair Owens (eds) *Urban Fortunes: Property and Inheritance in the Town 1700–1900*, Aldershot: Ashgate, pp. 79–107.

Owens, Alastair (2006) Making some provision for the contingencies to which their sex is particularly liable: Women and investment in early nineteenth-century England, in Robert Beachy, Beatrice Craig and Alastair Owens (eds) *Women, Business and Finance in Nineteenth-century Europe*, Oxford: Berg, pp. 20–35.

Parliamentary Papers, 1843 (496), Abstract of the Answers and Returns made pursuant to Acts 3 & 4 Vic. c.99, and 4 Vic. c.7, intituled respectively 'An Act for taking an Account of the Population of Great Britain', and 'An Act to amend the Acts of the last Session for taking An Account of the Population'. Enumeration Abstract, MDCCCXLI, Part 1, England and Wales.

Parliamentary Papers, 1852 (21), Accounts and Papers, Railways, 1852, XLVIII, Comparative Summary of traffic upon railways of England and Wales, Scotland and Ireland.

Parliamentary Papers, 1854, Reports from Commissioners (20), Railways; Woods and Forests: Local Acts (Preliminary Inquiries), Session 31 January–12 August 1854, Vol. XXXVII.

Parliamentary Papers, 1900, LXXVI, Accounts and Papers: Railway: Returns for England and Wales, Scotland, and Ireland for the year 1899.

Phillips, Nicola (2006) *Women in Business 1700–1850*, Woodbridge: Boydell Press.

Pullin, Nicola (2001) *'Business is Just Life': The Practice, Prescription and Legal Position of Women in Business, 1700–1850*, Royal Holloway, University of London, Ph.D. thesis.

Rubinstein, W. D. (2006) *Men of Property: The Very Wealthy in Britain since the Industrial Revolution*, rev. 2nd edn, London: The Social Affairs Unit.

Sanderson, Sir T. H. and E. S. Roscoe (eds) (1894) *Speeches and Addresses of Edward Henry, XVth Earl of Derby K.G.*, London: Longmans, Green.

Seeliger, Sylvia (1996) Hampshire women as landowners: Common law mediated by manorial custom, *Rural History*, 7(1), 1–14.

Shea, Gary, and Alex Trew (2007) Dynamic financial conditions and economic growth in the UK, *Economic History Society, Annual Conference, Abstracts*, Exeter, 239–40.

Staves, Susan (1990) *Married Women's Separate Property in England 1660–1833*, Cambridge, MA: Harvard University Press.

Stebbings, Chantal (2002) *The Private Trustee in Victorian England*, Cambridge: Cambridge University Press.

Tunali, Insan (1986) A general structure for models of double-selection and an application to a joint migration/earnings process with remigration, *Research in Labor Economics*, 8(B) (ed. Ronald G. Ehrenberg), pp. 235–83.

Wall, Richard (1984) Real property, marriage and children: The evidence from four pre-industrial communities, in Richard M. Smith (ed.) *Land, Kinship and Life-cycle*, Cambridge: Cambridge University Press, pp. 443–80.

Wooldridge, Jeffrey M. (2006) *Introductory Econometrics: A Modern Approach*, 3rd edn, Mason, OH: Thomson South Western.

9 The diffusion of steam technology in England

Ploughing engines, 1859–1930

Jane McCutchan

9.1 Introduction

The perception that the English countryside is a product of 'hoe agriculture' is enduring; a concept which the organisers of the opening scene of the 2012 London Olympic Games called 'Green and Pleasant' (Boyle, 2012). On entry to the Olympic Stadium in East London, the audience saw a scene that represented a traditional and idyllic view, complete with meadows, fields and rivers, and featuring farmers tilling the soil, while real farmyard animals grazed – including 12 horses, three cows, two goats, ten chickens, ten ducks, nine geese, 70 sheep and three sheep dogs. The scene changed, to represent the Industrial Revolution, but the pageant offered few clues to suggest how farming achieved this transformation, from the idyllic panorama of 'hoe agriculture', to feeding the masses toiling in the factories – other than by rolling up the turf.

By 1840, farmers were emulating the factories and using coal, which was not produced on the farm, as a source of energy. When steam engines replaced horses, the energy derived from coal was harnessed for field cultivation. However, supplying horses for other uses, such as road haulage, was part of a farm's business, and horse supply was perceived as sustainable because farmers could breed their own replacements, unlike steam engines, which depreciated in value and eventually wore out.

The argument in favour of using steam power was first and foremost one of cost, but there were considerable technical challenges which had to be overcome before Howard (1867) could claim in his firm's trade catalogue that traction engines adapted for steam ploughing, called ploughing engines, performed a greater amount of work in less time, more efficiently and at a lower price – together with tasks that had otherwise not been practical, for example, breaking heavier ground, digging deeper than previously possible and tackling jobs unsuitable for using horses. Table 9.1 compares the costs and benefits for horses and ploughing engines; horses, for example, are agile and can be used for other purposes, while ploughing engines are large, cumbersome and heavy, and driving them along narrow country lanes and turning into field gateways presents a challenge.

This chapter explores the impact of agricultural steam mechanisation in the UK, 1859–1930. The paucity of data on agricultural steam mechanisation in the

Table 9.1 Comparative benefits and costs of ploughing engines and horses

	Ploughing engines	*Horses*
Power	High power for breaking heavy ground and digging deeper (heavy soils such as clay are made friable and porous) Ability to pull heavy machinery (horses would require large teams and cumbersome harnesses)	Effective on light soils for ordinary field cultivation
Fuel	Coal is expensive (although cheaper if delivered by rail). Available all year round	Require fodder such as oats and hay, which is not freshly available in winter
Accommodation	Shed	Stable
Utilisation	Requires lighting up time before starting, and removal of fire and clinker deposits at end of day	Ready to work when awake; horse manure from stable can be spread on the land
	Benefits from continuous operation once fired up, but this can make lunch breaks, etc. difficult; may require a relief crew	Periodic stops for rest
Manoeuvrability	May be too large for gates, tight turning in a farmyard, etc.	Agile
Versatility	Ploughing engines can be used in stationary mode for other purposes; different accessories can be fitted for such purposes	Naturally versatile (e.g. for sport, recreation, pulling carts and carriages)
Purchase price	Very high	Modest: horses reproduce themselves
Maintenance	Requires regular oiling and preventative maintenance on worn parts	Liable to illness
Labour	Skilled labour for driving and maintenance	Experienced unskilled labour
Economics	Covers a large area of land quickly and therefore requires a large estate to be utilised effectively. The alternative is use by an independent contractor with several customers	Covers land slowly and can be kept fully occupied on a small farm

nineteenth century means that it has not attracted the extensive and systematic treatment that factory mechanisation and other uses of steam power, such as railways, have received. It is relatively neglected in mainstream economic history literature, but has received attention from agricultural historians and historians of technology. New estimates for the regional variations in the timing, pace and extent of early usage of steam engines in Britain during the eighteenth century have been suggested by Nuvolari *et al*. (2011) using an updated version of the list of engines originally compiled by Kanefsky and Robey (1980). Power availability and agricultural productivity in England and Wales between 1840 and 1939 are discussed by Collins (1996); the evolution and economic impact of steam mechanisation are described in Brown (2008) and Dewey (2008). Steam power was widely used on the farm, as well as for land reclamation and drainage schemes, and various types of stationary, portable and traction engines were developed, as summarised in Table 9.2.

The key data on steam ploughing engines used in this study are derived from the unpublished business records of John Fowler & Co., which are held at Museum of English Rural Life (MERL), University of Reading, together with records of other agricultural engineers. The research was made possible by Robert Oliver of the Steam Plough Club (SPC), who painstakingly, over a period of seven years, teased individual ploughing engine records from the Fowler archive, and made them available for academic research.

9.2 Getting up steam: the evolution of the ploughing engine

In 1866, Royal Agricultural Society of England inspectors made a tour to investigate steam ploughing with portable engines, and their reports give a picture of the transition between the use of horses, 'portables' and self-moving ploughing engines on the farm (Read, 1867). Horses pulled portable steam engines into position, to work the 'Roundabout ploughing system', suitable for small, irregular-shaped fields. Records suggest that trees and hedges were eventually torn out and field size increased in order to replace the 'portables' with self-moving ploughing engines, which were much more powerful.

Haining and Tyler (1970: 287–91) describe how the Roundabout system worked (see Figure 9.1):

> Engine and portable windlass were lined up and chocked so that the rope could run around the field and back, thus completing the circuit. The plough was incorporated in the circuit at a corner of the field opposite the engine, and the rope was pulled the width of the field via the windlass. Several labourers were required to move the corner guide pulleys after each traverse, the width of the furrow being ploughed, towards the engine, which then took up the surplus rope and commenced the next traverse.

When Fowler patented his 'Double-engine system', farmers who adopted the system claimed that the acreage under cultivation and the variety and quantity of crops that could be grown on the new land increased dramatically (Read 1867). This suggests it was the mighty single and double cylinder steam ploughing engines, described in Floud and McCloskey (1981) as 'failures', which were the hissing, whispering, whistling giants of the agricultural steam revolution. They

Table 9.2 Types of agricultural steam engine

Type	*Technical specification*	*Functionality*	*Accessories*
STATIONARY	Horse power: low – medium		
Beam engines	Large engines, often 'built in' to a structure, i.e. 'house built'	Pumping	Scoop wheel
Reciprocating		Milling	Grinding stones
Horizontal engines	Small engines with the horizontal cylinders and crank shaft supports attached to a rectangular base plate. Boiler and coal store often in separate buildings	Barn use	Threshing machine, crusher, kibbler, root cutter, etc.
Vertical engines	Almost always single cylinder, crank shaft supported overhead; can fit into a confined space	Barn use	As above
Early portable	Vertical engine on wheels	On the farm where needed	Saw bench
PORTABLE	Horizontal locomotive-type tubular boiler and firebox, single or compound cylinder, 6–8 feet high chimney, wheels but had to be horse-drawn horse power: medium, 6–8–10 hp	Barn, stack yard, field work	Threshing machine; saw bench Ploughing gear for Roundabout system Plough, windlass, driven either by belt or solid coupling from the engine, anchors, rope, rope-porter, pulleys
TRACTION	Any self-moving engine		
Saturated steam	Horse power 8, 10, 12, 14, 16	Heavy work	
Superheated steam	Horse power 20, 25, 30	Very heavy work	Smokebox-superheater
Ploughing engine	Traction engine adapted for steam ploughing	Field cultivations	
Single cylinder ploughing engine	Crankshaft converts the reciprocating linear motion of the piston(s) into rotation, by the action of the piston(s) and connecting rod(s) on the offset web of the crank. Chain, slanting shaft or gear driven models	Mole drainage Ploughing, cultivating, pulling out stumps and hedges	Single or double winding drums Ploughing gear Anti/balance plough Mole plough Cultivator Living van Water cart
Double cylinder ploughing engine	Low and high pressure cylinders side-by-side	As above	Single winding drums, engines work in pairs. As above

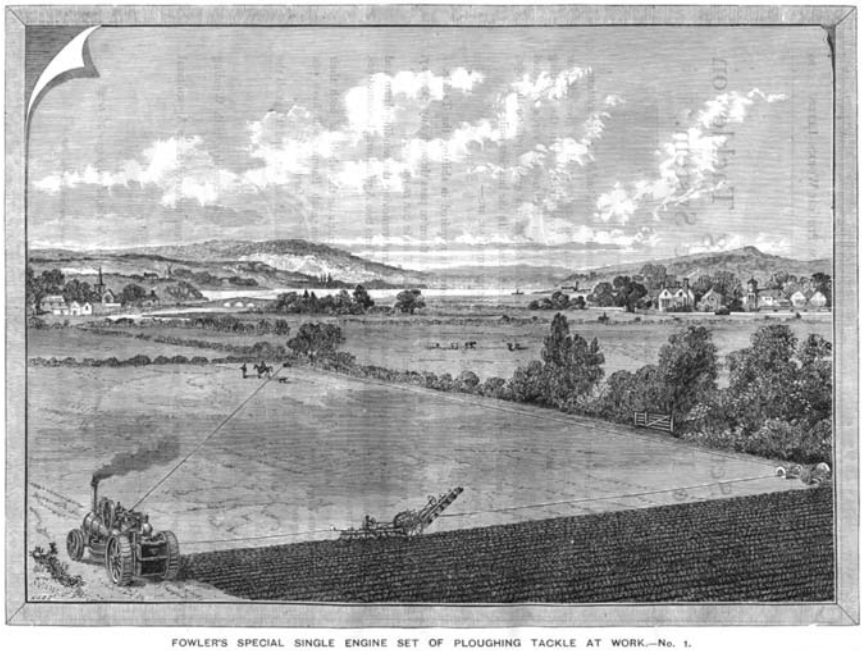

Figure 9.1 Fowler's 'Roundabout' system of steam ploughing.

Figure 9.2 Fowler's 'Double Engine Set' of ploughing tackle.

ripped out trees and hedgerows, brought new land into cultivation and increased its productivity. The two winding engines, working on opposite headlands, alternately pulled the plough across the field; the engine not in work paying out the rope, while moving into position for the return journey. Any kind of ploughing 'gear' could be used, but a Fowler Balance Plough was selected for the engraving in the trade catalogue (Figure 9.2).

9.3 Sources: the ploughing engine database

Identifying and accessing sources

Data on the production of single and double engine steam ploughing sets was collected in ACCESS from the extensive collections of primary and secondary material held at MERL and elsewhere. These comprised published literature, contemporary journals and company ledgers, which describe sequentially the manufacture date and type of each individual ploughing engine and record the name and address of the buyer. The key deposits of source material are listed in the key to Table 9.3.

Company order books are diverse and entries run into thousands, and it is difficult to find ploughing engine sales imbedded in the text; Taskers, for example, developed a reputation for making bridges as well as traction engines, steam wagons and road haulage equipment and they were the only company not to develop an export trade for their ploughing engines.

The number of ploughing engines produced by the minor manufacturers is small and they may be considered 'bespoke', but they contributed to engine development and are of considerable interest; for example Burrell's experiments with 'direct traction' ploughing engines were not successful, but advanced the cause of cable ploughing.

McClaren's business records were destroyed in a fire, and so the provenance details compiled by steam enthusiast Alan Duke, in the Duke archive at MERL, were consulted. Duke's lists also contain information from vehicle licensing records. Company ledgers and Duke's transcriptions are hand-written, so illegibility poses a problem. These secondary sources provided a key to finding ploughing engine records in the companies' total output and a 10 per cent random sample was cross-checked with the original documents to confirm accuracy.

Structure of the database

To better understand the characteristics of ploughing engine equipment, for example if features were built-in, or if accessories could be added, and if these were machine-specific or interchangeable, ploughing engine specifications were collected under the following headings:

- The firm that produced the engine
- Date of production
- Engine number

Table 9.3 Manufacturers of ploughing engines in England: archive sources

Company	*County*	*MERL*	*Lincoln*	*Suffolk*	*Hants*	*VLA*	*Alan Duke*	*SPC*	*Original records survive*
John Allen	Oxon	x						x	
Aveling & Porter	Kent		x				x	x	x
Charles Burrell & Sons	Norfolk						x	x	
John Fowler	West Yorkshire	x						x	x
Richard Garrett & Sons	Suffolk			x			x	x	x
Richard Hornsby & Sons	Lincs			x			x	x	x
McLaren	West Yorkshire					x	x	x	
Robey	Lincs							x	
Taskers	Hants				x			x	x
Wallis & Stevens	Hants	x					x	x	x
R.J. Wilder	Berks							x	

Key
MERL: Museum of English Rural Life, University of Reading
Lincoln: Lincolnshire County Record Office, Lincoln
Suffolk: Suffolk County Record Office, Ipswich
Hants: Hampshire county Council Archive, Winchester
VLA: Vehicle Licensing Authority
Alan Duke: Alan Duke papers, Museum of English Rural Life, University of Reading
SPC: Steam Plough Club of the UK Archives

- Cylinder type
- Horse power
- Key variants (modifications)
- Date and provenance of second-hand sales from 1859–1930 (where known)
- In the case of engines sold for export, the destination and country
- Details of new boilers which were supplied and fitted
- Price was not included because of lack of information

The agricultural engineering companies included in the study represent firms making cable ploughing engines in England between 1859 and 1930. The recording process took three months and included visits to different archives. Permission was obtained from Hampshire County Council Archaeology Department to examine the Tasker order books, stored in an industrial warehouse off a major roundabout in Winchester.

The data was transferred to an EXCEL spread-sheet for statistical analysis. The county names were standardised for ease of classification. Engines that went straight to export were allocated to a separate group. Missing counties were filled in, where the county name was manifestly known, e.g. *Northumberland* Steam Plough Company or where only the town was known e.g. Bedford. Several new 'pseudo' counties were created for classification purposes;

- *Blank***,** where the county could not be ascertained e.g. 'War Dept.' or the delivery address was simply missing;
- *Fowler*, where the engine was returned to the factory and resold or used for other purposes, e.g. demonstration or experimentation;
- *Show***,** where the engine went to an agricultural show before being sold;
- *Explode***,** where there was a catastrophic boiler explosion or other total loss thus taking the engine out of the study from a certain date; and
- *Export*, for engines which were exported.

Missing data

Some dates of production were missing and these were interpolated by arranging the engines by maker and engine serial number so that the dates of adjacent engines in the series could be ascertained. The missing dates could be recreated with a high degree of certainty; for example, the manufacture date for Fowler ploughing engine number 924K in the early Fowler/Kitson series was missing, but Fowler 923K and Fowler 925K were both built in 1861 and so the missing date was imputed as 1861. The dates of 58 engines were imputed in this way, i.e. 2.3 per cent of the all engines produced.

Results

Revised estimates of ploughing engines manufactured by producers in England between 1859 and 1930 are shown in Table 9.4. This table compares

Table 9.4 Revised estimates of ploughing engine output by manufacturers in England

Company	*Previously published estimate(s)*	*New estimates*		
	Total sales	*Total sales*	*Domestic sales*	*Export sales*
John Allen/Oxford SPC	7	7	7	0
Aveling& Porter	291	291	161	130
Charles Burrell & Sons	198	198	65	133
John Fowler	Various	6013	2146	3867
Richard Garrett & Sons	32	32	2	30
Richard Hornsby & Sons	36	36	33	3
McLaren	258	258	42	216
Robey	1	1	0	1
Taskers	6	13	13	0
Wallis & Stevens	1	1	0	1
TOTAL	830+Fowler	6849	2469	4380

Source: Database.

the new estimates from the database with estimates derived solely from the previous literature. The database provides the first definitive numbers for Fowler ploughing engines, and also revises upwards the number of Tasker engines.

When the new data are analysed by date of production it is evident that production can be divided into three phases. This is illustrated by the graph for the domestic market and exports in Figure 9.3.

- 1859–79: This is the main period of production for the domestic market;
- 1880–1915: The main period of overseas trade, until exports fell below domestic sales during the First World War;
- 1916–30: A boost in production arising from UK government contracts to increase self-sufficiency in national food production through intensive cultivation was quickly followed by the demise of the industry in 1928, when petrol-driven tractors superseded steam power.

Total production for the domestic and export market in each period is shown in Table 9.5. The switch from the domestic market to the foreign market and then back to the domestic market is analysed in more detail in Table 9.6, using data for Fowler engines. It presents the average annual prodcution for each market over each period, and also the proportion of production exported in each period.

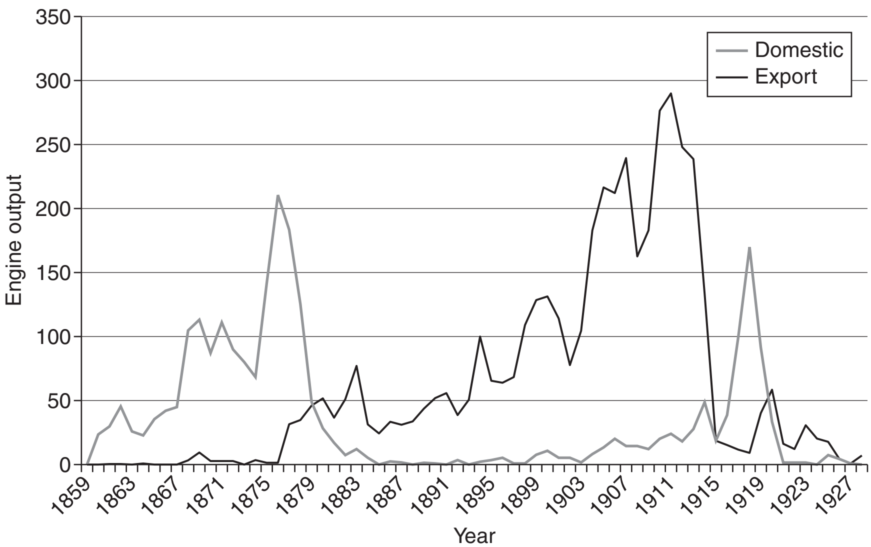

Figure 9.3 Ploughing engine output 1859–1930: domestic sales and exports (source: database).

Table 9.5 Total number of ploughing engines produced in the UK, 1859–1930

	1859–79	*1880–1915*	*1916–30*	*Total*
Domestic	1638	376	455	2469
Exports	150	3981	249	4380
Total production	1788	4357	704	6849

Source: database.

Table 9.6 The changing importance of the domestic and export sales for Fowler engines

Period	*Average annual sales*			*Export percentage*
	Domestic	*Export*	*Total*	
1859–79	78.00	7.14	85.14	8.39
1880–1915	10.44	110.58	121.03	91.37
1916–30	30.33	16.60	46.93	35.37
1859–1930	34.29	60.83	95.13	63.62

9.4 Time series analysis of production for the domestic market

The remainder of this chapter focuses on the use of steam ploughing engines in the UK. Export demand was very different from domestic demand; it was strongly influenced by the expansion of export-led farming in settler countries, including grain exports from Canada and Australia, and by cotton exports from Egypt; consequently it had a very different time profile, as noted above.

Table 9.1 highlighted competition between steam power and animal power. In the nineteenth century it was the heavy horse, such as the Suffolk Punch, that was the established competitor when steam ploughing commenced in 1859, rather than the oxen that had been used earlier. Steam engines were expensive one-off purchases, and so financing them could prove difficult. Qualitative evidence suggests that many of the early engines were sold to wealthy owners of large estates who wished to show off their engines as novelties; some seem to have been dedicated agricultural improvers, and others retired city bankers and merchants who had become hobby farmers. As adoption increased, however, agricultural contractors became a more important source of demand; some of the larger contractors owned fleets of engines which were hired out across the region in which they were based. It seems that the contractors normally used their own employees to operate and maintain the engines. Cost-based competition with horses therefore probably intensified as the use of steam engines diffused.

As Table 9.2 indicates, ploughing engines were only one type of steam power used on the farm; portable engines and traction engines were also used, e.g. for threshing, sawing and general haulage, but they are excluded from this study.

Unlike ploughing engines, horses were relatively cheap to acquire as they could be bred on the farm. One steam engine could replace several horses, however, as its horse power rating indicated. Horses had high running costs, as they required feeding and stabling and also regular periods of rest. The operating costs of a horse can be proxied by the price of oats, which was the most expensive and necessary form of animal feed, while the operating cost of the steam plough can be proxied by the price of coal. Steam engine operation requires a different type of skill to horse management, but there are no wage data that capture the skill differential involved in a satisfactory way.

The price and output in any market depend on a range of exogenous factors that affect demand and supply in that market. The analysis is confined to the output of steam ploughing engines because price data are incomplete, as noted above. Although the demands for steam engines and for horses are interdependent, the focus in this chapter is on the output of engines alone, and the technique of analysis is therefore based on a single equation regression that is estimated by ordinary least squares (OLS).

Examination of Figure 9.3 shows that the main challenge in explaining the time profile of engine production is to account for the peaks in 1878 and 1919, and the intervening low level of production in between – a time when production was mainly devoted to exports instead. The peak in 1878 can be explained by a standard diffusion model, in which the adoption of new engines in any year depends positively on the existing stock of engines, which increases over time, and the potential demand for engines that remains unsatisfied, which depends negatively on the existing stock of engines and positively on any factor that stimulates the potential demand for them. This diffusion model can be proxied by a quadratic time trend, in which output is related positively to time and negatively to the square of time, giving an inverted U-shape to the output profile which peaks at the point where the positive and negative forces on output balance out.

The figure also shows that the rise and decline on either side of the 1878 peak is generally smooth, apart from a brief plateau about 1870. The simplest way to account for this is in terms of the persistence of shocks; namely that the effects of any shock to demand or supply persists for several periods. This suggests that the regression should include lagged terms in the explanatory variables. Unobservable shocks may also exhibit persistence, and this suggests that lagged values of the dependent variable (i.e. past levels of output) should also be included in the regression (see Chapters 1–3).

If the diffusion process is accounted for by fitting a quadratic trend, then the exogenous variables need to be chosen to reflect variations in potential demand. This is the demand for a target stock of engines related to the exogenous factors, and not the demand for output in any period. Output in any period, net of diffusion, is explained, on this view, by changes in the target stock, plus any demand for replacement engines. The demand for replacement engines is difficult to analyse, and is assumed to be negligible for the purposes of this analysis; the focus is on explaining variations in the target stock.

The explanatory variables assumed to influence target stock are:

- *Prices of agricultural outputs that benefit from steam cultivation*. These are arable crops that require field cultivation: wheat, barley, oats and potatoes. Potatoes were an important cash crop under progressive six or seven course crop rotations (potatoes, wheat, mangolds, wheat, seeds, wheat; or potatoes, wheat, oats or barley, green rye, peas or tares with a final crop of either barley, seeds or wheat). Prices are taken from Clark (2004).
- *Prices of agricultural outputs that do not benefit from steam power*. The natural choice of substitute is wool, since sheep-rearing was a prominent use of pasture, and pasture was the major use of land taken out of arable production. The price of wool is also taken from Clark (2004).
- *Prices of inputs used intensively in steam cultivation and not in horse cultivation*; the price of coal is taken from Church (1986).
- *Prices of inputs used intensively in horse cultivation and not in field cultivation*: the natural choice is the price of oats, as described above. Note that the price of oats has a dual role in the analysis, namely as an output supplied to food processors and ordinary households, and as an input purchased by owners of heavy horses. These roles should, in theory, reinforce each other, as a high price of oats should both stimulate arable production and encourage such production to be powered by steam rather than horses.
- *The cost of capital used to finance the purchase of steam engines*. Since steam engines were highly durable assets, the appropriate cost of capital, as indicated by standard neoclassical economic theory, is the long-term interest rate. The long-term interest rate was low and stable for much of the second half of the nineteenth century, however, and therefore its variation is minimal. Short-term interest rates, such as bank rate, were more volatile, however. High short-term interest rates may have encouraged farmers to postpone their purchases until interest rates fell, while low interest rates may have encouraged them to bring purchases forward instead; the short-term interest rate may therefore be more important than the long-term interest rate in influencing the timing of purchases, and hence the time profile of output.
- *Other factors influencing the demand*. Rising prosperity may affect the demand for arable products; it may increase demand if people eat more, but possibly reduce it if they eat better, e.g. by switching from bread to meat. Rising population will tend to increase demand. In principle, foreign demand for UK agricultural products should also be considered, but in the period being studied exports of agricultural products were relatively low and imports predominated instead. The impact on arable agriculture of the UK prosperity and rising UK population may be captured by variations in aggregate UK income, as reflected in UK GDP. Both bank rate and GDP were measured using Bank of England historic data series (2013).

All the prices described above are treated as exogenous variables. This could be questioned on the grounds that competitive prices are in general endogenous, as explained in Chapter 2. However, the prices used here are not the prices for steam engines themselves, but for other commodities. While the prices of some of these commodities might themselves be influenced by the adoption of steam in agriculture, it would be the past adoption of steam rather than the current adoption that was relevant. Furthermore, steam ploughing engines never achieved a high degree of dominance in the market for powering ploughs, and so any influence of past production on current prices is likely to have been modest.

The derivation of the regression specification is explained in detail in the appendix. The regression results are reported in Table 9.7. Because the regression equation contains lags of up to three years, a missing observation in any year causes four degrees of freedom to be lost. There are two years for which accurate production data are missing (because the dates of production of certain engines are unrecorded), but the missing output can be interpolated using the method described above. The left-hand column of numbers reports the results without the missing observations and the right-hand column with them. The results are fairly robust to the variation, except with respect to the price of oats, as explained below.

In the columns each cell reports the estimated coefficient, and the significance level, expressed as a probability value, in brackets. Following the example set in Chapter 3, the lags associated with each variable are expressed in terms of the level lagged one period, together with changes in level lagged up to two periods (or three periods in the case of the output). Current output is therefore regressed on a constant term, the lagged level and lagged changes in output, the lagged levels and lagged changes in relevant prices (namely for coal, barley, wheat, oats, potatoes and wool), current and lagged GDP and the current bank rate. Lags in the bank rate were excluded because it was postulated the bank rate merely fine-tuned the timing of purchases that would have taken place anyway.

The results show that:

- There is significant inertia in output, but the inertia relates not to the influence of past levels of output on the current level of output but to the influence of past changes in output on the current level of output.
- Both the linear and squared components of the quadratic trend are significant and have the predicted signs. This confirms that the output of ploughing engines for the domestic market is to some extent demand-driven by a diffusion process.
- The coal price is insignificant. It is often asserted that the demand for steam power was driven by the declining price of coal, but if this was true of any period, it was true of the eighteenth century rather than the late nineteenth century when the price of coal was no longer falling. Ploughing engines benefit from good quality steam coal, and there was a huge demand for such coal both from railways and shipping industries. Although new mines were opened in the UK in the late nineteenth century, these were mostly deep mines that were costly to operate and were justified commercially only by the high price of steam coal. There were also periodic spikes in the price of coal caused by industrial disputes in the mines; however, there appears to be

Table 9.7 Time series regression for annual domestic ploughing engine sales, 1859–1914

Explanatory variables	*With missing observations*	*With interpolated observations*
Constant	–478.129*** (0.000)	–517.710*** (0.000)
Sales lagged 1 year	–0.065 (0.512)	0.001 (0.993)
Change in sales lagged 1 year	0.741*** (0.000)	0.710*** (0.000)
Change in sales lagged 2 years	0.457*** (0.000)	0.324*** (0.006)
Change in sales lagged 3 years	0.201** (0.041)	0.159 (0.138)
Time	6.793*** (0.003)	6.174** (0.011)
Quadratic time	–0.142*** (0.000)	–0.105*** (0.002)
Coal price lagged 1 year	–0.255 (0.513)	–0.498 (0.174)
Change in coal price	0.237 (0.387)	0.052 (0.851)
Change in coal price lagged 1 year	–0.297 (0.386)	–0.522 (0.220)
Barley price lagged 1 year	90.466*** (0.001)	97.831*** (0.000)
Change in barley price	–9.461 (0.437)	5.847 (0.610)
Change in barley price lagged 1 year	–73.661*** (0.000)	–63.323*** (0.000)
Wheat price lagged 1 year	–35.533*** (0.009)	–26.637** (0.047)
Change in wheat price	–10.529* (0.069)	–5.475 (0.356)
Change in wheat price lagged 1 year	37.770*** (0.000)	25.760*** (0.006)
Oats price lagged 1 year	74.615 (0.199)	57.403 (0.373)
Change in oats price	102.373 (0.780)	74.798* (0.099)
Change in oats price lagged 1 year	–10.998** (0.014)	–58.745 (0.186)
Potato price lagged 1 year	7.285** (0.014)	1.638 (0.588)
Change in potato price	2.883** (0.028)	2.619** (0.038)
Change in potato price lagged 1 year	–1.836 (0.189)	0.582 (0.638)
Wool price lagged 1 year	12.059 (0.000)***	10.439 (0.000)***
Change in wool price	7.365*** (0.000)	5.839*** (0.010)
Change in wool price lagged 1 year	–6.669*** (0.004)	–6.333*** (0.005)
GDP lagged 1 year	0.048 (0.365)	0.071 (0.197)
Change in GDP	0.025 (0.532)	0.063 (0.222)
Change in GDP lagged 1 year	0.112*** (0.007)	0.095** (0.025)
Bank rate	–4.742** (0.029)	–0.064** (0.025)
R^2	0.988	0.978
Adjusted R^2	0.967	0.951
F-statistic	45.533 (0.000)	35.709 (0.000)
Durbin–Watson	2.363	2.691
Number of observations	44	51

Note
Significance level based on White robust standard errors: *** significant at 1 per cent; ** significant at 5 per cent; * significant at 10 per cent.

no association between these disputes and peaks and troughs in steam engine output. For a variety of reasons, therefore, it seems that fluctuations in the price of coal cannot explain fluctuations in steam engine output.

- The lagged level of the barley price has a significant positive effect on steam engine output, as expected, but the lagged level of wheat price has an unexpected negative sign. The change in the wheat price lagged one year has a significant positive effect, however. Overall, the results are consistent with the view that barley and wheat prices both influence the demand for steam engines, but in rather different ways.
- There is no clear pattern in the impact of the price of oats. If oats proxied the cost of horse power as well as demand for arable products then a strong positive effect would be expected. The absence of a significant positive effect suggests that substitution between steam engines and horses may not have been as sensitive to relative fuel costs as standard neoclassical economic theory would suggest. It lends indirect support to the view that ploughing engines in the UK were to some extent 'playthings' for wealthy landowners, whose decisions to purchase were motivated by a desire to enhance their social status in the county community rather than to improve the profitability of their estates. This does not mean that the same was true of the export market, however.
- Changes in GDP are insignificant, but bank rate has the predicted negative effect, although the size of the effect is sensitive to the omission/interpolation of observations.

The overall fit of the regression is impressively high, although much of the fit is accounted for by the lagged output variables. The Durbin–Watson statistic is acceptable, but indicates that some positive serial correlation remains in the residuals even when quite extensive lags are used. This suggests that the shocks impinging on the market for steam engines are highly persistent.

9.5 The spatial diffusion of steam engines across the UK

Additional insights into the demand for ploughing engines can be obtained through cross-section analysis. This is because it is possible to identify the owners to whom the ploughing engines were supplied, and thereby to identify the locations in which they were first used. It is also possible to track many of the engines throughout their working lives, but this analysis is not presented here.

It is possible to construct a county profile of the stock of steam ploughing engines in 1879, based on the cumulated output of engines supplied to purchases in each county. Mapping this distribution shows a concentration of the stock in the South-east of the country, and along much of the East Coast as far north as County Durham. Engines can also be found in the Midlands as far north as Staffordshire and Nottinghamshire, but there are very few in Wales, Scotland and the South-west (beyond Wiltshire and Somerset). A range of county characteristics for 1871 was compiled from the 1871 Census returns and from tables in the relevant volumes of the *Agricultural History of England and Wales*. A summary

Table 9.8 County profiles: summary information

UK agricutural region classified by Caird	*County*	*County size (acres)*	*Ploughing engine domestic sales 1859–79*	*Miles to London*	*Miles to Manchester*	*Rural pop. (in 000s) 1871*	*Horses kept for agri-culture (in 000s) 1875*
1	Cambridge	525,152	38	62	167	114	19
1	Suffolk	947,681	35	78	215	214	32
1	Essex	1,060,549	122	38	202	171	32
1	Herts	391,141	34	26	193	81	11
1	Beds	295,582	25	56	152	92	9
1	Hunts	229,544	19	66	146	39	7
1	Norfolk	1,354,301	25	114	188	265	41
2	Lincoln	1,775,457	123	141	86	268	48
2	East Yorks	838,970	35	214	100	93	23
2	Kent	1,039,419	121	41	240	177	25
3	Surrey	478,492	30	33	197	44	11
3	Sussex	936,911	39	61	261	178	21
3	Hampshire	1,070,216	59	66	215	144	22
3	Berkshire	451,210	61	44	156	110	12
3	Nottingham	52,5076	48	127	73	72	14
3	Leicester	514,164	8	100	103	78	12
4	Rutland	95,805	0	100	110	20	2
4	N'hants	630,358	44	66	136	154	14
4	Warwick	563,946	51	95	100	106	15
4	Oxford	472,717	30	59	156	109	13
4	Bucks	466,932	35	43	168	81	12
4	Middlesex	180,136	6	10	196	NA	4
5	Shropshire	826,055	19	160	70	154	19
5	Worcester	472,165	44	132	106	80	14
5	Hereford	391,141	2	133	123	85	13
5	Monmouth	368,399	NA	127	143	50	6

5	Gloucester	805,102	39	101	142	137	18
5	Wiltshire	865,092	53	113	190	151	18
5	Somerset	1,047,220	27	166	208	230	22
5	Dorset	632,025	37	126	234	116	13
5	Devon	1,657,180	3	196	236	249	35
5	Cornwall	873,600	3	284	319	160	22
6	N'umberland	1,249,299	75	316	178	94	14
6	Durham	622,476	45	269	126	42	11
6	North Yorks	1359600	19	234	69	137	26
6	West Yorks	1,685,409	33	186	43	164	33
7	Cumberland	1,001,273	6	307	116	99	14
7	Westmorland	485,432	2	234	62	38	4
7	Lancashire	1,219,221	12	219	58	105	22
7	Cheshire	707,078	0	192	48	103	13
7	Derbyshire	658,803	12	129	59	55	13
7	Staffs	728,468	43	141	40	107	15

Sources: The list of counties of England by area is from the Census of Population, 1871, as cited in *Encyclopaedia Britannica* (1890), Vol. 8, p. 220; the distance from the county town (as in 1860) to London is from the Google distance calculator; Corn cultivation: the percentage of cultivated area under corn crops is from Collins (2000, pp. 136–7); Wheat cultivation data is from the *Chamber of Agriculture Journal and Farmers' Chronicle*, 27 March 1871, cited in Collins (2000, pp. 132–3); Barley cultivation: John (1989, pp. 1048–50, data from Mark Lane Express); Oats cultivation: John (1989, pp. 1048–50; data from Mark Lane Express); Potato cultivation: *Annual Agricultural Statistics and Returns* (MAFF 1866–1914). Number of horses per county kept for agricultural purposes is from Afton and Turner (2000, pp. 1800–4). For Caird's regional classification see Craigie (1883) cited in John (1989, pp. 1048–50).

of the data on county location, size, number of ploughing engines, rural population and stock of horses is presented in Table 9.8. Information was also obtained of the areas under cultivation for specific crops. Comprehensive data were available for only 39 English counties, comprising all English counties except Buckinghamshire, Middlesex and Monmouthshire.

A cross-section linear regression was estimated to analyse the impact of county agricultural characteristics on the density of ploughing engines, measured as the stock of engines per million acres. The basic hypothesis was that the adoption of ploughing engines was influenced by the nature of the crops grown (as dictated by the nature of the soil), and by the distance of the county from major markets. Two major markets were identified; one in the South-east centred on London and another in the industrial North-west, centred on Manchester. If it had been possible to disaggregate to below the county level then more centres could have been included in the analysis. Two measures of distance were used for each centre; the distance of the county town from the market centre, and a dummy variable indicating whether the county town was within 50 miles of the market centre. This allowed for the possibility that lands where the adoption of steam cultivation was marginal were reasonably close to a market centre. Rural population was included, both as a measure of local labour supply and as a measure of local demand for food. A high rural population, it was hypothesised, could have a mixed impact: it could stimulate large-scale agriculture and boost steam cultivation, or it could supply cheap manual labour for use in conjunction with horse power. The population of horses used in agriculture was also included, and it too was expected to have mixed effects. A high density of horses might reflect conditions unsuitable for steam cultivation, but on the other hand it could reflect a buoyant agricultural sector affording scope for steam cultivation. Neither of these variables is ideal as an explanatory variable, but the range of data available for profiling county agriculture is rather limited. The remaining variables all concern the proportion of land under tillage for various crops.

The regression results are presented in three versions, depending on whether tillage is measured using only a composite dummy variable (column A), only disaggregated variables (Column B), or a combination of the two (column C).

The results are presented in Table 9.9. They are unambiguous so far as distance to market and type of tillage is concerned. A high level of tillage in general, but above all high tillage of wheat, is a paramount influence. A distance less than 50 miles from London is also crucial. Both effects are positive and the coefficients are relatively large. Nothing else seems to matter so far as statistical significance is concerned. It should be noted, however, that no strong effects are predicted for rural population or horse population in any case. The R^2 statistics show that tillage of wheat and proximity to London together explains more than half the variation across counties in 1879.

The strong positive results for wheat in these cross-section regression contrast with the more ambiguous results for wheat in the time series analysis. The two results are not inconsistent, however. The time series results relate to the price of wheat and the cross-section results to the amount of wheat cultivation. In the

Table 9.9 Cross-section regression of ploughing engine sales 1859–79 by county characteristics

Explanatory variables	*A: Overall tillage only*	*B: Types of tillage only*	*C: Overall tillage and types of tillage*
Constant	26.214	–66.744	–72.551
	(0.406)	(0.328)	(0.243)
Proportion of agricultural land under tillage exceeds 50 per cent in 1871	32.733**		21.182
	(0.037)		(0.179)
Proportion of tilled land growing wheat in 1871		2.556**	1.984*
		(0.021)	(0.061)
Proportion of tilled land growing barley in 1871		0.695	0.748
		(0.349)	(0.293)
Proportion of tilled land growing oats in 1871		–0.147	–0.234
		(0.860)	(0.773)
Proportion of tilled land growing root crops exceeds 20 per cent in 1871	4.138	21.874	14.355
	(0.777)	(0.117)	(0.401)
Proportion of acreage growing potatoes in 1866	–0.436	–1.188	0.691
	(0.954)	(0.915)	(0.953)
Horses per acre	–0.052	0.241	0.294
	(0.903)	(0.607)	(0.525)
Rural population per acre	1.291	2.260	2.577
	(0.586)	(0.451)	(0.452)
Miles from London	–0.057	–0.099	–0.055
	(0.586)	(0.313)	(0.610)
Within 50 miles of London	40.753**	51.269***	47.078**
	(0.020)	(0.006)	(0.022)
Miles from Manchester	–0.008	0.007	0.019
	(0.930)	(0.943)	(0.848)
Within 50 miles of Manchester	7.448	10.496	16.071
	(0.776)	(0.705)	(0.581)
R^2	0.538	0.594	0.619
Adjusted R^2	0.394	0.428	0.443
F-statistic	3.751	3.587	3.516
	(0.003)	(0.003)	(0.004)
Number of observations	39	39	39

Note
Significance level based on White robust standard errors: *** significant at 1 per cent; ** significant at 5 per cent; * significant at 10 per cent. Includes all English counties other than Buckinghamshire, Middlesex and Monmouthshire, for which some data were missing.

short run there may be a weak connection between the two due to lags in adjusting agricultural practices and regional styles of farming to changes in market prices. The spatial pattern may therefore indicate a long-run spatial equilibrium, while the time profile of production may reflect the dynamics of diffusion, status-seeking by early-adopters and the influence of short-run financial conditions on the timing of purchases.

9.6 Conclusions and implications for further research

This chapter has presented the first comprehensive and definitive database of steam ploughing engines produced in the UK. The industry has sometimes been dismissed as a failure, but this judgement seems premature. Steam ploughing engines were an important export in late-Victorian and Edwardian England, and were a key technology in opening up the prairies and the bush in settler economies for large-scale export-oriented agriculture. Imports into Britain from the settler economies ultimately undermined the economic basis of much intensive arable farming in the UK, but not before English farmers had explored the potential of the steam ploughing engine. Between 1859 and 1879 production for the domestic market showed sustained growth, interrupted only slightly around 1870.

Although the main determinants of demand for ploughing engines are fairly clear, more work remains to be done to clarify certain issues. It is possible, for example, that the influence of proximity to London on the adoption of steam power could be due to heavy clay soil found in the London area, which made steam power particularly effective on land that was very tiring for horses. The permeability of soil is also an issue which requires further investigation. More attention should also be given to the use of steam engines in the cultivation of root crops, some of which were used to feed the population of London and other major urban centres.

Access to the railway network could also be a factor favouring steam power near to London. The main lines converging on London from all directions, coupled with the early development of the suburban railway network, meant that many farms near London were very close to the railway network. It is known that many steam traction engines were delivered by rail, and that in some cases they were unloaded between stations at a point near to the farm rather than at a station. They despatched from the works completely erected, except for the flywheel and chimney, and the wheels might be dismounted and re-fixed on arrival at their destination. They were craned onto ‘well carriages’ at the factory and chained; for example, Lord Zetland’s Fowler Ploughing Engine No. 99 (939K) was sent to Richmond, in the North Riding of Yorkshire in 1862 on the York & Newcastle line. The fact that the trunk railways network in the South-west was largely broad gauge until 1892, and that all major engine producers were served only by standard gauge lines, could be a factor that inhibited diffusion to the South-west. Furthermore, transport infrastructure facilitated not only the delivery of engines, but also the delivery of the coal to power them; and it also provided access to market for the agricultural products they produced.

Marketing strategies also need to be considered. Steam plough producers seem to have been quite familiar with the concept of diffusion, and the importance of demonstration and word-of-mouth recommendations by influential people. This may explain the early adoption by local opinion-leaders; it is possible that they were targeted by the manufacturers and offered favourable terms. Large engineering companies like Fowler’s promoted sales of their ploughing engines through demonstrations and trade stands at agricultural shows. Evidence for their presence is provided by the prizes that they won – which they subsequently listed in their trade catalogues. Fowler appeared to target markets near his Leeds factory, in the Midlands and North-west, although he went as far north as Newcastle on the East coast and to Ayr on the West coast of Scotland.

All of these issues relating to the pattern of diffusion are currently under investigation in a follow-up study.

Appendix: specification of a time series model for ploughing engine output

This is a demand-driven rather than supply-driven model. Demand is determined by partial adjustment of the inherited stock of ploughing engines to a target stock of ploughing engines. The target stock depends on exogenous factors such as input and output prices.

Stock-flow relationships imply that output is equal to gross investment, where gross investment is the sum of net investment and depreciation. Net investment is the change in the capital stock from the end of the previous year to the end of the current year, while depreciation is the amount of the capital stock at the end of the previous year that does not survive to the end of the current year. Depreciation is normally assumed to be proportional to the amount of inherited capital stock.

Let y_t be the stock of PEs at the end of year t (a period is one year) ($t=1,\ldots, T$). Let q_t be the output of PEs during year t. Let d be the proportional rate of depreciation of the stock of PEs; then

$$q_t=(y_t-y_{t-1})+dy_{t-1}=y_t+(1-d)\,y_{t-1}$$

Assume for simplicity that $d=0$; then we have:

$$q_t=y_t-y_{t-1} \tag{1}$$

Equation (1) implies that if we can explain the time path of y_t then we can also explain the time path of q_t.

Let x_{jt} be the value of the jth exogenous variable in year $t(j=1,\ldots, M)$. The target capital stock for the end of year t is y^*_t, where

$$y^*_t=a+\textstyle\sum_j b_j x_{jt} \tag{2}$$

Demand for output is based on the discrepancy between the target out and the inherited output. Due to uncertainty and/or costs of adjustment, farmers plan to eliminate only a proportion $k(0<k\leq 1)$ of the discrepancy in any given period; k is known as the partial adjustment factor. Adjustment also involves a random shock, v_t:

$$\begin{aligned} y_t&=y_{t-1}+k(y^*_t-y_{t-1})+v_t \\ &=ky^*_t+(1-k)y_{t-1}+v_t \end{aligned} \tag{3}$$

A preliminary analysis of the data indicates high serial correlation in output, which suggests that the impacts of random shocks may persist for several periods. It is therefore appropriate to assume that v_t follows a first order autoregressive process:

$$v_t = hv_{t-1} + u_t \tag{4}$$

where h $(0 \leq h < 1)$ is the persistence factor. The u_t are serially uncorrelated by assumption, but the v_t are not. It is assumed, as usual, that u_t are uncorrelated with any of the exogenous variables (with or without lags).

Substituting (2) into (3) and substituting the result into (4) shows that

$$y_t = (1-h)ka + (1-k+h)y_{t-1} - h(1-k)\, y_{t-2} + k\textstyle\sum_j b_j x_{jt} - hk\sum_j b_j x_{jt-1} + u_t \tag{5}$$

Solving the auxiliary equation

$$\lambda^2 - (1-k+h)\lambda - h(1-k) = 0$$

shows that the system is stable when the assumed parameter restrictions $0 \leq h < 1$, $0 < k \leq 1$ hold. Applying equation (1) shows that

$$q_t = (1-h)ka + (h-k)y_{t-1} - h(1-k)\, y_{t-2} + k\textstyle\sum_j b_j x_{jt} - hk\sum_j b_j x_{jt-1} + u_t \tag{6}$$

Equation (6) is a special case of the general equation

$$q_t = \alpha + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \textstyle\sum_j \gamma_{1j} x_{jt} + \sum_j \gamma_{2j} x_{jt-1} + u_t \tag{7.1}$$

where

$$\alpha = (1-h)ka \tag{7.2}$$

$$\beta_1 = h - k \tag{7.3}$$

$$\beta_2 = -h(1-k) \tag{7.4}$$

$$\gamma_{1j} = kb_j \tag{7.5}$$

$$\gamma_{2j} = -hkb_j \tag{7.6}$$

Equation (7.1) may be estimated by OLS by regressing output on lagged stock (lagged both one year and two years), and on current and one-year lagged values of the exogenous variables. Linear and quadratic time trends can be added to help to capture the effects of omitted variables; lags of these variables do not need to be included in the regression.

References

Afton, B. and M.E. Turner (2000) The statistical base of agricultural performance in England and Wales, 1850–1914, in E.J.T. Collins (ed.), *Agrarian History of England and Wales*, *Vol. 7: 1850–1914*, Cambridge: Cambridge University Press, pp. 1757–2140.

Bank of England (2013) *Statistical Interactive Database*, www.bankofengland.co.uk/boeapps/iadb/BankStats (January 2013).

Boyle, Danny (2012) *Green and Pleasant*, London: London Olympic Games Press Office.

Brown, Jonathan (2008) *Steam on the Farm, A History of Agricultural Steam Engines 1800 to 1950*, Ramsbury: Crosswood Press.

Church, Roy (1986) *The History of the British Coal Industry, Vol. 3*: *1830–1913: Victorian Pre-eminence*, Oxford: Clarendon Press.

Clark, Gregory (2004) The price history of English agriculture, 1209–1914, *Research in Economic History*, 22, 41–123.

Collins, E.J.T. (1996) Power availability and agricultural productivity in England and Wales 1840–1939, *Discussion Paper Series*, Reading: University of Reading.

Collins, E.J.T. (2000) Rural and agricultural change, in E.J.T. Collins (ed.), *Agrarian History of England and Wales*, *Vol. 7: 1850–1914*, Cambridge: Cambridge University Press, pp. 72–233.

Craigie, P.G. (1883) On statistics of agricultural production, *Journal of the Royal Statistical Society*, 46, 40–2.

Dewey, Peter (2008) *Iron Harvests of the Field: The Making of Farm Machinery in Britain since 1800*, Lancaster: Carnegie.

Floud, R. and D. McCloskey (eds) (1981) *The Economic History of Britain since 1700, Vol. II, 1860 to the 1870s*, Cambridge: Cambridge University Press, p. 182.

Haining, J. and C. Tyler (1970) *Ploughing by Steam: A History of Steam Cultivation over the years*, Hemel Hempstead: Model & Allied Publications.

Howard, J. and F. (1867) Catalogues for Steam Cultivation.

John, A.H. (1989) Statistical Appendix, in G.E. Mingay (ed.), *Agrarian History of England and Wales*, *Vol. VI: 1750–1850*, Cambridge: Cambridge University Press, pp. 972–1155.

Kanefsky, J.W. and J. Robey (1980) Steam engines in 18th century Britain: a quantitative assessment, *Technology and Culture*, 21 (2), 161–86.

MAFF (1866–1914) *Annual Agricultural Statistics and Returns*, London: HMSO.

Nuvolari, A., B. Verspagen and N. von Tunzleman (2011) The early diffusion of the steam engine in Britain, *LEM Working Paper Series*, Pisa: Santa Anna School of Advanced Studies.

Read, H. (1867) Reports of the Committees appointed to investigate the present state of steam cultivation, *Journal of the Royal Agricultural Society of England*, Volume III, 2nd Series.

10 Cupidity and crime

Consumption as revealed by insights from the Old Bailey records of thefts in the eighteenth and nineteenth centuries

Sara Horrell, Jane Humphries and Ken Sneath

10.1 Introduction

Consumption has taken on ever-increasing importance in explaining the economic growth that both preceded and accompanied industrialisation. Indeed, Jan de Vries' (2008) 'industrious revolution' assigns the power to motivate increased market-orientation and labour supply to the emergence of new, particularly oriental, goods and the acquisitive desires they prompted. People worked harder and longer and more frequently for wages to earn more money in order to be able to buy these novel goods. The new orthodoxy has industriousness as precursor to the Industrial Revolution itself (see e.g. Allen, 2009).

Certainly exotic goods became available. From the late-seventeenth century there were imports of tea, sugar, spices, porcelain and silks from the East. The era of the 'New Draperies' witnessed the manufacture of lighter, more appealing clothing. Homes were transformed with window curtains, new kinds of furniture, such as the switch from chests to drawers, and comfortable feather mattresses. Clocks and mirrors aided time-keeping, appearance and light diffusion. Eating became a new experience as people swapped pewter and wooden vessels for earthenware, glassware and metal cutlery.

There is no doubt that there was a 'comfort revolution' in the early modern period, an improvement frequently attested to through the composition and value of probate inventories. But whether an industrious revolution with its drivers in consumption really foreshadowed the industrial revolution is less certain. Questions remain unanswered.

- Did all classes and all geographical regions share in this expanded consumption? In particular, did those who putatively increased their work effort, labourers, craftsmen, women and tertiary sector workers share in this consumer boom? Furthermore, were they motivated as in the model by avaricious impulses?
- If, instead, the growth in consumption arose as a result of price declines brought about by improvements in transport, technology and technique, then an alternative underlying mechanism would be implied.

- Does the timing of these consumption shifts match periods of economic expansion in the early modern and industrial eras, and so ground industrialisation in the Smithian growth of the long eighteenth century?

Probate evidence provides few answers to these questions. Although the requirement to list goods on death remained until 1792, the numbers of surviving inventories thinned out rapidly after 1720 (Cox and Cox, 2000). Moreover, even for periods when inventories are plentiful, they largely pertain to the 'middling sorts', tradesmen, artisans, small farmers and upwards, rather than the labourers and work people whose industriousness is central to any attempt to link expanding consumption to an increasing supply of labour. More generally then, the evidence from inventories shows increased ownership among the elite in urban areas but a much slower uptake in rural areas and further down the social strata (French, 2007; King, 1997; Overton *et al*., 2004; Shammas, 1990; Sneath, 2009; Weatherill, 1998). And, if the reach of probate leaves out exactly those households whose behaviour is central to the industriousness hypothesis, the evidence itself might also be in doubt.

Inventories required a valuation to be given to the goods recorded but the accuracy of these is uncertain, and the quality, age and other detail of the good typically remain unspecified. A value was sometimes ascribed to a cluster of items in a room or store cupboard. However, these valuations show the increased quantity of consumer goods owned to be attained with a static overall cost, implicating falling prices rather than cupidity in expanding demand (Shammas, 1990). At any rate, probate evidence cannot resolve the key questions about the role of consumption in economic growth.

Alternative data sources are hard to come by. Although from 1792 we have information on the food brought by ordinary households, starting with the work of Davies (1795) and Eden (1797), we have nothing on the larger durable purchases and rarely anything on clothing. Few commentators were sufficiently interested in the standard of comfort of the mass of the people to provide descriptions of how they lived until the conditions in factory towns in the mid-nineteenth century prompted authors such as Engels (1845) to report on this aspect of people's lives. This information gap bedevils attempts to demonstrate links between new wants, increased industriousness and economic growth.

However, with ingenuity and care, we can bring another source to bear. Specifically we use crime data for London to reveal desirable items of consumption over time. The digitisation of the Old Bailey records (Old Bailey, 2012) and improved adeptness at converting largely qualitative accounts into databases which can be subjected to econometric analysis using standard statistical software has facilitated the use of this source for the oblique investigation of trends in goods available.

Over the period we consider, 1750–1821, the Old Bailey covered all serious crimes committed in the City of London and the county of Middlesex, essentially the area north of the Thames, which accounted for 60 per cent of London's urban population in 1680 and more subsequently (Shoemaker, 1991, ch. 1). At this time, London was the fashion centre whose extensive emporia showcased new goods (Boulton, 2000, p. 325; Schwarz, 2000, p. 648). The provinces lagged behind.[1] If a consumer revolution was happening in England and Wales, this is where the action must have been, quite by accident, etched into the capital's crime statistics. While stolen goods do not equate to consumption, they do represent the shifting nexus between demand and supply. As *Social Trends* puts it summarizing current crime statistics:

> The nature of crime may change over time. Some crime, such as burglary, may stay the same in that it involves the breaking and entering of households and theft of goods, but the type of articles stolen in a burglary will change, reflecting amongst other things fashions, technological developments, and the desirability and availability of various household goods.
>
> (ONS, 2002, p. 153)

Then as now, the nature of stolen property reflects the stock of goods exploited by opportunistic or planned larceny and reveals the preferences of thieves; both imply trends in consumption. The records do not provide an accurate account of ownership in the wider population, but they do indicate when items emerged as readily available and document some of their key characteristics.

The first task is to allay the doubts of readers on this point and establish the suitability of the Old Bailey records as an incidental commentary on consumption. Utilising extant sources to investigate issues unconnected to the purpose for which the data were originally collected and retained is fraught with difficulty and presents an obvious possibility of inherent bias. Indeed the susceptibility of crime statistics to unrepresentativeness and mutability over time is widely known. To use these data further we need to reassure ourselves that they are fit for purpose. Much of this chapter assesses this suitability but we also report some results from our project by looking at what the crime data can reveal about consumption of a specific item about which a considerable amount is already known: clothing. These results help establish the validity of our method.

10.2 Sources: a critical appraisal

Three issues are key to whether or not the Old Bailey records are fit for purpose as indicators of trends in consumption.

- Were the cases documented in the Papers and Proceedings a reliable record of prosecutions or were they a biased sample?
- Were prosecutions a reliable record of thefts or were they a biased sample?
- Were stolen goods broadly representative of availability and attractiveness or were they selected according to specific characteristics, such as portability and saleability?

The first issue concerns the papers as a record of prosecutions, and here, since they provide one of the most detailed and extensive sources available to historians and have been widely used in the investigation of crime and legal practice, their strengths and weaknesses are well known. John Langbein (2003), an authority on the source, lists three potential problems:

- The focus on London, which gives an urban slant to any dependent historical inquiry;
- The evolution from sensation-mongering chap books in the Elizabethan period to more complete and quasi-official accounts in the crime calendar of the later eighteenth century; and
- Their variable scope and reliability, as the compressed accounts of the 1710s gave way to ever more detailed reports.

These attributes constitute a serious handicap to historians trying to use the papers to bring crime itself into historical perspective or to establish trends in courtroom practice (Beattie, 1986, 2001; Devereux, 1996). However, our interest is not in the substance of the trials or their revelations about legal procedure but in the proceedings' unintentional inclusion of extraneous information, which with some ingenuity can be used to unlock puzzles in economic history. How do the known drawbacks of the source bear on our project?

- In our context, the skew to London is a good thing. If a consumer revolution was happening anywhere it was here in the fashion centre of the country, whose extensive emporia showcased new goods unavailable in the provinces. New shopping streets in Piccadilly, St James and the Haymarket displayed a whole range of new and fashionable British goods (Berg, 2005). Little sign of new goods in London would be devastating to the hypothesis of a consumer revolution.
- The metamorphosis of the source is also unproblematic since the focus of our study is the eighteenth century, by which time the papers had become 'quasi-official (Langbein, 2003, p. 183). Indeed, from 1729 the Session Papers are acknowledged to be a complete, albeit sometimes brief, record of the trials held at the Old Bailey (Beattie, 2001, p. 24).
- The variable scope and reliability of the series turns out to be not such a problem either. In the period studied, the proceedings continued to cater to a predominantly middle-class readership, selecting cases accordingly. Financial pressures meant that, for example, in 1774–7 a large fraction of cases, particularly acquittals, were reported in cursory 'squib' accounts, in order to hold down size and hence publishing costs (Shoemaker 2008). Langbein warns against 'quantitative analysis from such incomplete data' (2003, p. 185). However, his interest is in juridical developments, where compressed reporting could be seriously misleading. Our concern is with mundane features of each crime, where even the squib reports are sufficient; they specify the crime, the swag, its valuation and erstwhile ownership. We

proceed on the basis that the reports of the prosecuted thefts for this period are essentially complete.

This leads to a second issue that concerns the extent and nature of the biases involved in whether a theft actually came to be prosecuted. The nature of the legal process in the eighteenth and early nineteenth centuries has led historians of crime to conclude that the underlying level of criminal activity remains unknown and unknowable (Beattie, 1986, p. 199; Hay, 1982, p. 117). Various stages had to occur before a crime appeared in the records as an indictment. A putative offence had to occur; the alleged perpetrator had to be caught and brought before a Justice of the Peace, a responsibility which fell to the injured party; and the magistrate had to determine the nature of the crime and the appropriate response. Only then, and only if the crime was deemed sufficiently serious, was an indictment drawn up for the case to be heard by High Court circuit judges at an Assizes (Beattie, 1986; Shoemaker, 1991). The, often considerable, financial and opportunity costs of bringing the case to court fell to the victim. Such disincentives to bringing a prosecution might suggest that very few thefts, and then only the most serious ones, became indictable offences. Indeed, in his study of the Middlesex and Westminster quarter sessions for 1660–1725, Shoemaker (1991) shows that, although most thefts technically should have been treated as felonies and, if reported should therefore have resulted in indictments, very often it seemed that victims were satisfied with the return of their property and, maybe, some compensation. In lieu of prosecution, the issue of a recognisance, binding the alleged perpetrator to appear in court on forfeit of a bond, or summary justice in terms of a fine, a whipping or a brief stay in a House of Correction were dispensed. None of these would remain in the historical record as an indictable offence.

The convoluted legal process of the time suggests that we should be wary, if not downright sceptical, about the possibility of auditing stocks of consumer goods through indictment records. For example, the types of goods observed in registers of stolen property may be biased towards those owned by people more able and willing to pursue a prosecution. Fortunately, such problems are mitigated by legal administration in the capital. Although the City and Middlesex magistrates held Sessions of the Peace eight times a year and dealt with misdemeanours and nuisance, as elsewhere, they dealt with few charges of theft. Even cases of petty larceny (a misdemeanour rather than a felony involving thefts of less than one shilling in value) were rarely prosecuted at the sessions. Nearly all cases of theft were referred to judges at the Old Bailey, with the consequence that ‘all data concerning the prosecution and trial of property crime in the City can be derived from the records of the Old Bailey’ (Beattie, 2001, p. 17).

However, changing propensities to prosecute remain an issue for this observed record. Four factors affected the rate of prosecution: the likelihood of apprehension; the costs of prosecution; the chances of a guilty verdict, which itself depended on prevailing attitudes to specific crimes; and, the public acceptability of the punishment. These factors all changed over the long eighteenth century.

The chances of apprehending suspected criminals and retrieving stolen property improved over time. Contributing factors included:

- Better communications, with newspapers and an increasingly effective postal system publicising descriptions of stolen property and suspects and advertising rewards for retrieval or capture;
- The early eighteenth-century use of (often dubious) thief-takers;
- Improvements in street lighting and better surveillance by night watchmen, funded out of the rates from 1700 onward;
- The establishment of the Bow Street officers, the forerunners of the police, recruited by Henry and John Fielding from 1748; and
- The formation of prosecution associations to apprehend suspects and share costs of prosecution in the second half of the eighteenth century.

These all increased the probability that a suspect would be detained, while moves to regularise the availability of magistrates in the later 1700s helped to improve the likelihood of indictment (Beattie, 1986, pp. 35–73).

Changes in the apportionment of costs also encouraged prosecutions. An Act of 1752 made provision for State reimbursement of some costs in the case of convicted felonies. In the 1760s, it has been estimated that around one-third of prosecutors' witnesses at the Surrey assizes were granted costs with labourers more frequently reimbursed than those from other groups (Beattie, 1986, pp. 43–4; King, 1984, pp. 32–3). In 1778 aid was extended so that costs could be paid to those in 'poor circumstances' regardless of verdict, so that nearly two-thirds of eligible cases in Surrey were awarded costs in 1792–4 (Beattie, 1986, pp. 44–5). Bennet's Act in 1818 extended such legal aid. These changes likely caused an increase in the rate of prosecution from any given level of crime. Additionally, anxieties about 'crime waves' which occurred regularly in the capital may have encouraged both prosecutions and guilty verdicts. Against these trends was a shift in what was perceived as appropriate punishment.

After the Restoration, there were essentially two types of crime: felonies and misdemeanours. Each was associated with a specific form of punishment: capital (the death penalty) and public whipping, respectively. Crimes were also classified according to whether or not they were eligible for benefit of clergy. Benefit of clergy allowed leniency for capital offences if the accused were clergy, later to be more broadly interpreted as literate. Concern that access to benefit of clergy was weakening the deterrent effect of the law prompted the reclassification of many property crimes as non-clergyable in the early eighteenth century. Robbery, burglary and housebreaking, which often involved physical violence as well as the loss of property, were always thought sufficiently serious to merit capital punishment and so were denied benefit of clergy. But the 'lesser' crimes of theft from a house, shop, warehouse, ship or manufactory, theft of livestock, shoplifting and pick-pocketing also became non-clergyable, tightening the noose around the necks of perpetrators who were found guilty. Petty larceny remained a non-capital offence. It seems many prosecutors of theft felt the death sentence too severe a penalty and either failed to prosecute or downgraded the accusation to a clergyable larceny until alternative forms of punishment became available (Beattie, 1986, pp. 140–98). However, burglary and housebreaking offences continued to be prosecuted at consistent levels into the nineteenth century (Beattie, 1986, pp. 165–7).

Might these changes over time impact systematically on the types of good stolen? Although apprehending the perpetrators of more serious crime may have become easier, the use of night watchmen and street lighting may have enabled the detection of more petty theft: any overall effect on the type of good stolen remains uncertain. The awarding of costs may have encouraged people further down the social scale to prosecute but labourers were regularly awarded costs as early as the 1740s and 50s (King, 1984, pp. 32–3). Moreover, crime historians have argued that from the Restoration onwards the view that justice was available to all permeated down the social scale. By the mid-eighteenth century, ordinary people prosecuted maybe half of all offences (Shoemaker, 1991, introduction). In 1743–53 14 per cent of the prosecutions at quarter sessions in Surrey were brought by labourers, 26 per cent by artisans and 6 per cent by women (Beattie, 1986, p. 193). Thus, there is uncertainty whether help with costs changed the social composition of litigants (Beattie, 1986, pp. 196–7) and so, importantly for our project, affected the observed composition of stolen goods. Growing distaste for hanging as an all-purpose punishment may have deterred prosecution in general, but the public seemed firmly behind the designation of burglary and housebreaking as capital offences. Contemporary commentary suggested that these crimes were often committed by 'professional' thieves for whom there was little public sympathy.[2] Indeed increased policing and surveillance did not deter burglary whose rewards became ever greater (Beattie, 1986, pp. 161–7). Overall then, while prosecutions may not reflect underlying levels of crime, trends in indictment for housebreaking and burglary were less affected by the changes over the eighteenth century than were other forms of theft and it would seem that the social composition of those who brought indictments for these felonies did not vary greatly or systematically.

However, even if goods recorded as stolen represent a consistent sample over time, were the specific commodities themselves a selected sample? Were stolen goods selected for their transportability and ease of onward sale? The exchange of goods in early modern London involved multifarious arrangements. At one end of the scale were the burgeoning retail outlets which sold new goods and at the other peddlers and hawkers who often dealt in second-hand (Lemire, 2006). In between, there were street sellers, markets, dealers, pawnshops, auctions held in people's homes or in inns, and straightforward barter. Some goods provided a medium of exchange as cash and coinage could be scarce, and many workers were, at least in part, paid in-kind, through perquisites or by truck. Gifts or inheritances, also, may have needed conversion into other commodities. Although authorities attempted regulation, informal and unlicensed trading persisted. For instance, at the end of the eighteenth century there was concern about the "'vast numbers of unlicensed Hawkers and dealers in old Cloaths assembling every afternoon' in some London districts" (Lemire, 2006, p. 247). The alehouse provided a popular market place, maybe particularly for stolen goods and there was one on almost every street corner. In 1796, Patrick Colquhoun estimated there were around 6,000 licensed public houses in London (Lemire, 2006, p. 249). Dishonest shopkeepers knowingly bought and resold stolen goods like tea and tobacco while purloined clothing easily found its way to the pawnshop. There was ample opportunity to slip dishonestly acquired goods of all sorts into this tangled web of formal and informal channels of exchange and permanent and transitory outlets.

Who would buy these goods? Historians have suggested that the purchase of second-hand items represented the trickle-down or emulation elements of consumerism (Lemire, 1988). Fashions moved on and the cast-offs of the rich became inherited style for the poor. There was a huge market for second-hand clothing in eighteenth-century England in shops, at pawnbrokers, fairs, markets and inns, and involving itinerant traders. It often overlapped with new sales. Specific areas in London specialised in its sale, for instance, Monmouth Street, Rosemary Lane and Petticoat Lane. The working poor frequently bought their garments at pawnbrokers, market stalls and slop shops rather than the smart retail establishment frequented by the upper class. The middling sorts were more inclined to engage in alterations, remakes and accessorising to remain chic. Items other than clothing followed somewhat different routes and reached wider markets. There was a significant market for second-hand household goods, such as linen, china and furniture. Some shops sold both new and second-hand, sometimes accepting old goods as a trade-in for new. Furniture and books reached the market via auctions or dedicated book sellers. However, such outlets did not just cater for the poorer classes pressed by necessity, but were also visited by the middling sorts looking for a 'bargain', and exhibiting 'clever' consumption (Stobart, 2006, p. 233). The elite even purchased the durables of the era, such as carriages, second-hand, typically through specialist markets. Thrift may have featured in such purchases but clever consumption also played its part. The destination of other items, such as quantities of lead and timber and tea and sugar stolen in bulk, is less certain, but the discussion of perquisites and non-monetary forms of payment suggests that much would have been exchanged for cash with manufacturers and shopkeepers (Lemire, 2006). Most goods could find an outlet and were sold to most classes of ordinary people.

The characteristics of stolen goods might reflect not just desirability but also portability. To some extent whatever was stolen depended on the type of theft, for instance a pick-pocket could only take what was carried on the person, but housebreakers and burglars had more options. Surprisingly perhaps they were rarely limited to what could be easily moved. Gangs of thieves, organised crime and even opportunistic behaviour could lead to whole houses being stripped. One victim reported how everything, including the curtains from his windows, had been removed, another how she was locked in her cellar by five men while they removed the entire contents of her shop, and the Archbishop of Canterbury's Palace was robbed of a fortune in silver plate in 1788 (Beattie, 1986, p. 163 n. 52, p. 164 n. 53). Burglary particularly attracted professional thieves because of the promise of large rewards (Beattie, 1986, pp. 161–7). Therefore beds, mattresses, desks and carpets all found their way from people's homes onto the resale market alongside smaller items such as clothing, household linen, jewellery and glassware. Mayhew's (only slightly later) accounts of housebreaking and burglary provide a vivid picture of thieves at work that is entirely consistent with the idea of a systematic sifting of property for its attractiveness and value. Thieves are described as well-equipped and organised, often starting on the top floor of premises and working their way through the building, not hesitating to break open the servant's money box, readily evaluating plate and other valuables and calmly packing up large numbers of shirts and silk handkerchiefs (Mayhew, 1861, pp. 366–73). The size of attractive booty was no deterrent: 'Should the plunder be bulky, they will have a cart or cab, or a costermonger's barrow, ready on a given signal to carry it away' (Mayhew, 1861, p. 372).

Overall, stolen property reflects the goods that people had as well as those they wanted and changes over time capture both availability and desirability. The Old Bailey records can be used to document change in goods in the possession of people at various levels of society and can map transitions from the exceptional and novel to the common and outmoded.

10.3 Preliminary analysis of the data

Our focus in this chapter is to consider the booty of housebreakers and burglars.[3] We concentrate on these forms of theft because, as argued above, indictments for burglary and housebreaking were less likely to have been affected by changes in the legal system than other forms of theft. Furthermore, what was taken from people's houses suggests the range of goods owned and coveted. We can also avoid some of the pitfalls occasioned by change in definition for other types of theft or by uncertainty about the nature of the crime. For instance, pocket picking required the thief to 'privately' steal goods valued at a shilling or over from a person, without that person's knowledge. This latter requirement was dropped in 1808, yielding many more cases of pick-pocketing. Shoplifting occurred when goods valued at five shillings or more were stolen from a shop but a decline in prosecutions in 1820–1 reflected a disinclination to meet out the statutory death sentence. The law was changed in 1823. Until 1827, when the distinction was removed, grand larceny (a capital offence) covered the theft of goods valued at one shilling or more, petty larceny (non-capital) thefts under one shilling, which possibly encouraged prosecutors to downgrade the value of the item stolen. In cases of 'theft from a specified place' there are sometimes doubts about illegality. For instance, traditional perquisites were reconstructed as pilfering around the 1790s and dockers' acquisition, for example, of sweepings of sugar and tobacco after the unloading of ships became criminalised. Similarly, lodgings were often rented fully furnished. Pledging the bedding and curtains at the local pawnbrokers may have been theft but, if the intention was (eventually) to redeem they might not have been permanently alienated (Styles, 2006). Housebreaking and burglary together are usually just below 10 per cent of all prosecuted thefts in the late eighteenth and early nineteenth centuries, but lower earlier (see Table 10.1).

Table 10.1 Economic conditions and prosecutions for theft

Years	*Total no. thefts (all types)*	*No. housebreaking and burglary cases*	*House/burglary as percentage of total thefts*	*Population of London[1] ('000)*	*House/burglary as percentage population (×1,000)*	*Price of bread[2]*	*War/peace time?*
1750–1	849	51	6.0	675	7.6	4.9	Demobilisation following 1739–48 War of Austrian Succession
1760–1	576	15	2.6	732	2.0	4.7	War 1756–63 Seven Years' War
1770–1	1,145	123	10.7	789	15.6	5.75	Peace
1780–1	875	75	8.6	846	8.8	5.6	War 1775–83 War with American Colonies
1790–1	995	94	9.4	903	10.4	6.5	Peace
1800–1	1,555	127	8.2	960	13.2	12.45	War 1793–1815 Napoleonic war
1810–11	1,624	155	9.5	1,017	15.2	14.2	War 1793–1815
1820–1	2,633	140	5.3	1,074	13.0	10.1	Peace

Notes
1 Taken from Wrigley (2011, p. 61) for 1750 and 1800. Interpolated for intermediate years.
2 Taken from Mitchell (1988, pp. 769–71), price in d. per 4 lb loaf, average price for previous two years used to capture extent of hardship faced.

We focus on thefts from homes between 1750–1 and 1820–1. The start date relates to the need to identify consumption patterns in early industrialisation when existing probate evidence fades away. The end is dictated by the discontinuities in classification and coverage that followed the restructuring of criminal law by Sir Robert Peel in the 1820s (Beattie, 1986, p. 13). By the 1850s, Old Bailey business was serious crime.[4] For each decade, data for two consecutive years were taken to yield a sufficiently large sample size of housebreaking and burglary cases (see Table 10.1). In total this resulted in 780 crimes involving some 4,542 individual stolen items. By computerising the information contained in all cases of housebreaking and burglary for the selected years we have information on every item stolen for which the theft was prosecuted, thus ensuring we record the full range of purloined property.

The detail in the Old Bailey records on the goods taken in each theft is considerable. For instance, on 26th May 1790 the Court heard that

> Elizabeth Asker was indicted for burglariously and feloniously breaking and entering the dwelling house of Thomas English, about the hour of nine in the night, on the 18th March last, and burglariously stealing therein, three cotton gowns, value 14s., four cotton petticoats, value 10s., two black silk cloaks, value 10s. and one child's dimity cloak, value 2s. his property.[5]

Thomas English, of Great Earl Street, Seven Dials, and his wife Ann recounted how they had discharged Elizabeth Asker, their servant of three weeks, in the early evening and how she had then returned at night and removed their possessions. Two officers, who took her into custody, and two pawnbrokers, to whom she tried to sell the clothes, gave evidence. Found guilty, Elizabeth Asker was sentenced to death by Mr Justice Ashurst.

For each of the cases, we recorded in the database the year in which the trial took place, the Old Bailey identifying number, whether the theft occurred through housebreaking or burglary and whether more than one item was stolen. We then recorded some details of the accused; age and sex; and of the erstwhile owner of the property; age, sex, address of crime, status or occupation of owner, and any additional detail given. For each item stolen we noted what was stolen, the material it was made of, the number of each item stolen (for instance, one gown or six spoons) and the value, from which value per item in decimal (£) was subsequently computed, for up to 30 items in each case of burglary or housebreaking.

Thus, for the theft from Thomas English above, the details of the items stolen would be recorded against the case identifier as follows:

Item	*Number*	*Material*	*Value*	*Value per item*
Gown	3	cotton	14s	0.233
Petticoat	4	cotton	10s	0.125
Cloak	2	silk	10s	0.25
Cloak, child's	1	dimity	2s	0.10

Note that thefts of multiple identical items in the same crime are counted only once within the case. Thus the four petticoats stolen above are given the same weight in the data set as one stolen in another break-in. That is, in the total count of items stolen multiple identical items count as just one item.

In addition to recording all details of the theft against the specific crime case, we also separately itemise the individual items stolen in the data set. Thus we also include sub-case lines for: case identifier, second (third, etc.) item stolen, type of good stolen, number, material, and value per item, as above. This repetition enables us to choose the level at which we wish to analyse the data. We can either select crime cases and focus on, for example, who goods were stolen from and the types of goods stolen together, or, using the sub-case information, select a type of good, for instance, gowns, and look at the value and material it was made from for every gown stolen within our data set.[6]

The accounts vary in the extent to which they state the material from which stolen items were made. In the eighteenth century descriptions were available for between three-fifths and four-fifths of all stolen items, but by the 1820s this has declined to just over 10 per cent. The reports continued to state whether items were made of gold or silver or contained gem stones, but other metals, cloth and other materials dropped off substantially. This may have reflected greater difficulty in accurately identifying the item's constitution with the advent of new materials, but could also reflect a desire to compress the proceedings. Styles (2007, p. 329) observes this change in relation to reports of stolen clothing in the Old Bailey papers but continues to find materials stated in the Northern quarter sessions reports which would support the latter explanation.

Some historians have argued that the level of criminal prosecutions and perhaps even the level of crime itself is related to underlying economic circumstances (Beattie, 1974, 1986, pp. 199–236; Hay, 1982; Shoemaker, 1991). Year-to-year fluctuations in indictments for theft, they allege, correlate positively with high grain prices, and involve increases in petty theft. They also correlate negatively with wartime, it is suggested, because war stimulates industrial expansion and high employment, while lulls in hostilities lead to demobilisation and saturated labour markets, particularly in urban areas, and hence to higher crime. Others have been sceptical about the economic relationship (Innes and Styles, 1986). They have argued that, given the low level of thefts actually prosecuted, changes in the behaviour of prosecutors could cause the observed fluctuations as easily as changes in the behaviour of offenders (Innes and Styles, 1986, p. 393). For instance, low levels of prosecution in wartime may have resulted from an ability to offer military service as an alternative to punishment and, if applied to habitual criminals, may have had a significant effect on the number of indictments actually observed. Conversely, the return of criminally inclined soldiers could be causally connected with observed post-war crime waves. The purported relationship between high prices and crime has also been criticised: again it may well be a statistical artefact caused by perpetrators having less money with which to buy off prosecution (King, 1984). Whether indictments reflect the underlying level of crime and how this relates to economic conditions remains uncertain, but our concern is less with this set of links than with whether the putative effects of economic conditions would introduce any systematic biases into the goods we observe being stolen and the social class of those from whom they were stolen. High prices and demobilisation in our sample period could be marked by increased theft. If this involved opportunistic stealing by poor people from those further down the social scale it might bring lower value goods into purview. However, contemporaries thought that many burglars and housebreakers sought high-value swag and worked in professional gangs. Similarly, the return of the habitual criminal from the army might raise the number of high-value thefts that occurred and counterbalance the first effect. *A priori*, it is difficult to identify a systematic bias.

According to Table 10.1, burglaries and housebreaking show a high correlation with the population of London,[7] but other factors are also important. With the exception of the Napoleonic Wars, instances of burglary and housebreaking clearly follow the pattern of lower levels of crime in wartime and higher levels in peacetime.[8] High levels of theft are evident in the early nineteenth-century years of poor harvests and high wheat prices.[9] Using regression analysis to disentangle these effects shows only bread prices and a dummy variable for 1770–1 to be significant determinants of thefts from homes. This would imply that variations in prosecuted thefts, in part, have an economic basis, and this might then affect the range of goods we observe being stolen over time. In particular, we might expect more low-value thefts to enter the proceedings from 1800 onwards, occasioned by high bread prices, and more high-value thefts in peacetime years, when more hardened criminals might have been at work. These two effects might be expected to raise the number of thefts but counterbalance each other in terms of the average value of goods stolen in 1820–1. We can investigate this by considering the average value of items stolen in each year.[10]

As already illustrated, valuations are readily available in the Old Bailey records. In each indictment, the value of the stolen property had to be given, usually agreed between the owner and the court clerk. Invariably, these were second-hand values (representing both the current value and the price usually paid by the majority who purchased second-hand) but they are not thought to have involved any systematic bias. There was no obvious benefit to the claimant in either inflating or deflating the value, although the defendant may have preferred it reduced below the threshold that determined the specific form of theft and hence the resultant penalty. Styles (2007, app. 2 t.17–19) has shown that the values given for a wide range of clothing items in the Proceedings were similar to those recorded elsewhere. It is reasonable to query whether second-hand valuations are an adequate proxy for prices. Presumably they reflect the prices that thieves anticipated receiving and, indeed, may be better indications of the prices usually paid than those that economic historians are usually constrained to use, such as wholesale or institutional prices. Of the 4,542 items stolen, only 423 had no specific value attached.[11]

We relate the above observations on value to the expected trends in value of goods stolen over time given the prevailing economic circumstances. We

Table 10.2 Relationship between expected and actual value of items stolen

Years	*Expected value of items stolen*	*Average value of all stolen items (£ per item)*
1750–1	High	0.22
1760–1	Medium	0.24
1770–1	High	4.55
1780–1	Medium	1.11
1790–1	High	0.82
1800–1	Low	2.18
1810–11	Low	5.31
1820–1	Medium	2.14

Note

High expected value of goods stolen given value 3, medium 2 and low 1. Pearson correlation coefficient –0.353 (not significant, 0.39, two-tailed test); Spearman's rank correlation coefficient –0.504 (0.20, not significant, two-tailed test).

compare two values for each decade: the expected values of goods stolen based on whether the sample observations are from war or peacetime years and the prevailing price of bread; and the average value of all stolen items (Table 10.2). Correlation coefficients show no relationship between the actual value of goods stolen and the expected value.[12] Although economic factors may have affected the number of thefts, they did not have a significant effect on the value of the goods stolen. We find this test reassuring; it appears that the data can capture trends in consumerism and goods available and that other factors do not swamp this picture.

10.4 Insights into consumer preferences from thefts of clothing

For our analysis of these records for the insights they reveal into consumerism and the respective roles of technology and tastes in expanding demand we focus on one large subset of items taken: clothing. In our data set 1,942 individual types of clothing were reported as taken;[13] this represents 42 per cent of all items stolen in incidences of housebreaking and burglary in the Old Bailey records in the years selected. Clothing has been deemed the single most important category of material culture in the explanation of expanding consumption in the third quarter of the eighteenth century (Riello, 2006, pp. 5–9). This is a category about which much is already known so our purpose is to illustrate how thefts documented in the Old Bailey Proceedings follow fashions in style and material, and to examine these trends in a quantitative way, not to replicate the comprehensive study of clothing, fashions and culture epitomised by the work of John Styles (2007).

Fashions were evident (Table 10.3). While traditional items, shifts, aprons, stockings, were taken in 1750–1, we can also observe the theft of more fashionable items, such as handkerchiefs. Thieves may have been attracted to stockings as they changed from home-spun to workshop-made and hence became better quality and more easily marketed, while the lure of handkerchiefs, the fashion accessory of the day, probably lay in their bright colours and fancy materials (Fine and Leopold, 1993; Styles, 2007). Note that 'fashion' tended to centre on changed accessorising – ribbons, scarves, laces and buckles – rather than wholesale change of dress. In the 1770s, we note a marked transition in the types of items taken. Stolen clothing follows fashions with cloaks, waistcoats, sleeves, shawls, petticoats and small accessories such as collars, veils, frills and ruffles, growing in importance. Boots and shoes too disappeared more often, perhaps because people could afford more than one pair (Riello, 2006, pp. 22–3). By the century's closing years a more austere economic environment prevailed in the capital (Allen, 2009). Although thefts show some retrenchment into the traditional and mundane, fashion still played a role. Shawls and scarves continued to be taken and, now-unfashionable, items, such as perukes and sleeves, fell out of favour. The nineteenth century saw new fashions in clothing appearing on lists of stolen property: trousers, jackets, drawers (underwear), umbrellas, pelisses, spencers and fashionable accessories.

The value of items stolen (Table 10.4) generally shows a remarkable stability through time, particularly for the unchanging items, such as stockings, caps, shifts and shirts. Even items such as waistcoats, aprons, breeches, cloaks and shoes are fairly consistent. There is some evidence of the valuation given to items becoming lower as they fall out of fashion, for instance perukes and sleeves, more variation where the item is heterogeneous, for instance handkerchiefs and gowns, and increases where quality may have improved, for instance coats and shoes. Thus, trends in valuations appear consistent and reasonable.

The materials from which clothing was made are also illuminating (Table 10.5). Shirts and shifts remained made out of linen throughout, cotton was not making inroads into undergarment production.[14] Aprons too were predominantly linen, although instances of other fabrics, including the gradual ascendancy of fashionable muslin, is evident. Petticoats were made from a variety of materials but stuff, dimity, flannel, calico and cotton were perennial favourites. New materials and fashions are most evident for stockings, gowns and handkerchiefs. While worsted stockings were stolen throughout, silk made an appearance in the fashion-conscious late-eighteenth century, but both were quickly superseded by cotton in popularity. Similarly, gowns followed the fashion from linens to (presumably printed) silks and cottons with cotton, calico and muslin gowns in ascendancy from 1790 onwards.[15] Men's clothes showed more stability although here too fashionable materials make an appearance. Breeches were typically made of standard woollen cloth or leather but velvet, nankeen, corduroy, kersey and fustian all make appearances from 1780 onwards. Waistcoats too shifted towards linen from cloth in 1770–1 and then to cottons from 1780–1. Coats, however, remained made of cloth. Handkerchiefs showed great variety but here again silk and linen were gradually superseded by muslin, and cotton began to make inroads from the 1770s.

Table 10.3 Items stolen in housebreaking and burglary: clothing

	1750–1	*1760–1*	*1770–1*	*1780–1*	*1790–1*	*1800–1*	*1810–11*	*1820–1*
Skirt	–	–	4	1	2	–	1	2
Shirt	7	2	30	13	17	20	23	25
Apron	7	7	29	24	19	14	8	13
Stockings	11	5	24	27	26	29	22	21
Breeches	3	2	18	5	15	13	10	17
Cap/hood/bonnet	4	4	10	10	17	8	11	9
Hat	3	2	20	2	4	6	10	11
Coat	5	4	26	9	12	25	17	25
Waistcoat	2	2	24	10	9	12	45	18
Shift	7	8	11	10	10	5	10	11
Gown/frock	5	6	38	30	22	41	31	34
Cape/cloak	1	1	13	18	9	6	4	5
Stays	2	2	2	1	3	2	1	–
Shoes/boots	2	1	16	6	8	21	14	19
Petticoats	2	8	23	28	17	22	14	12
Gloves	2	–	1	2	1	7	2	1
Peruke/wig	1	–	2	–	1	–	–	–
Jacket	–	–	–	–	1	5	5	4
Sleeves	1	1	10	4	1	1	–	–
Handkerchief	13	9	45	24	36	34	37	28
Pantaloons/trousers	–	–	–	–	–	5	5	10
Drawers	–	–	1	1	–	1	1	3
Parasol/umbrella	–	–	–	–	–	–	–	5
Pelisse (women’s jacket)	–	–	–	–	–	–	4	7
Spencer (men’s short tail coat)	–	–	–	–	–	–	1	5
Shawl/scarf	–	1	10	1	11	12	8	14
Accessories: collar, veil, ruffle, frill	3	1	10	1	1	3	3	7
Other accessories	–	1	3	2	6	7	7	4
Night gown	1	2	7	4	8	4	8	2
Other clothing	–	2	4	16	–	7	5	6

Table 10.4 Value of some clothing items commonly stolen

	1750–1	*1760–1*	*1770–1*	*1780–1*	*1790–1*	*1800–1*	*1810–11*	*1820–1*
Shirts	0.15 (2)	0.11 (2)	0.41 (25)	0.14 (12)	0.15 (17)	0.17 (20)	0.20 (23)	0.31 (25)
Aprons	– (–)	0.06 (2)	0.11 (20)	0.13 (24)	0.08 (19)	0.10 (14)	0.05 (8)	0.03 (13)
Stockings	0.07 (2)	0.06 (5)	0.12 (20)	0.11 (26)	0.09 (26)	0.07 (28)	0.11 (22)	0.08 (21)
Breeches	0.34 (2)	0.15 (2)	0.67 (16)	0.60 (5)	0.28 (15)	0.54 (13)	0.47 (10)	0.45 (16)
Cap, hood	0.02 (1)	0.13 (1)	0.05 (5)	0.04 (10)	0.06 (17)	0.04 (8)	0.08 (11)	0.11 (9)
Hat	0.13 (2)	0.25 (1)	0.25 (18)	0.01 (2)	0.17 (3)	0.07 (6)	0.19 (10)	0.25 (11)
Coat	0.38 (2)	0.48 (2)	1.06 (19)	0.33 (9)	0.61 (12)	0.57 (25)	1.11 (17)	1.16 (25)
Waistcoat	– (–)	0.25 (2)	1.86 (20)	0.21 (10)	0.20 (8)	0.12 (12)	0.21 (15)	0.35 (18)
Shift	0.17 (3)	0.11 (2)	0.10 (9)	0.14 (10)	0.09 (10)	0.10 (5)	0.26 (10)	0.14 (11)
Gown/frock	0.03 (2)	0.60 (2)	0.37 (25)	0.56 (30)	0.35 (22)	0.40 (41)	0.34 (31)	0.48 (34)
Cape, cloak	– (–)	– (–)	0.36 (9)	0.71 (17)	0.54 (9)	0.34 (6)	0.38 (4)	0.32 (5)
Stays	– (–)	1.00 (1)	0.40 (2)	1.50 (1)	0.27 (3)	0.54 (2)	1.00 (1)	– (–)
Shoes, boots	0.14 (1)	– (–)	0.15 (15)	0.26 (6)	0.22 (8)	0.24 (21)	0.42 (14)	0.32 (19)
Petticoats	– (–)	0.05 (1)	0.45 (12)	0.27 (28)	0.22 (17)	0.20 (22)	0.09 (14)	0.13 (12)
Gloves	0.05 (1)	– (–)	0.20 (1)	0.05 (2)	0.02 (1)	0.05 (7)	0.08 (2)	0.03 (1)
Peruke, wig	– (–)	– (–)	0.43 (2)	– (–)	0.25 (1)	– (–)	– (–)	– (–)
Jackets	– (–)	– (–)	– (–)	– (–)	0.25 (1)	0.18 (1)	0.31 (5)	0.18 (4)
Sleeves	– (–)	– (–)	0.04 (7)	0.03 (4)	0.01 (1)	0.01 (1)	– (–)	– (–)
Handkerchiefs	0.17 (7)	0.18 (6)	0.08 (31)	0.08 (24)	0.06 (36)	0.05 (34)	0.15 (36)	0.11 (0.28)

Note
In each cell the value per item
(£) appears first followed by the number of valuations in brackets. A dash signifies no cases.

Table 10.5 Material from which some commonly stolen clothes were made

	1750–1	*1760–1*	*1770–1*	*1780–1*	*1790–1*	*1800–1*
shirts	linen (cotton)	linen	linen (cotton, calico)	linen	linen (flannel)	linen (cotton, calico)
aprons	linen (cotton, wool)	linen (lawn)	linen, muslin (cloth, lawn, flounce)	linen, muslin (cloth, gauze, lawn)	linen, muslin (camblet)	muslin (linen, cotton)
stockings	worsted (wool, yarn, thread, cotton)	(worsted, silk, cotton)	worsted, silk, cotton, thread (linen, gauze)	cotton, silk, linen, worsted (thread)	cotton, silk, worsted (thread, muslin)	cotton, worsted, silk thread
shift	linen	linen (dowlas)	linen	linen	linen	linen
gowns	linen (camblet)	linen (camblet)	cotton, linen, silk stuff (camblet, muslin, satin, bombazine)	linen, silk, cotton, stuff (calico, crape)	linen, cotton, calico muslin (silk, stuff, dimity)	cotton, muslin, silk (linen, satin)
petticoats	–	calico (wool, dimity satin)	dimity, stuff, flannel, silk, calico	dimity, stuff, flannel, cotton, calico linen (silk, woollen)	dimity, stuff, flannel, cotton, satin (calico)	dimity (flannel, satin, muslin, flannel, calico cotton)
handkerchiefs	silk linen (gauze, lace cotton, muslin)	linen silk (camblet)	linen, silk, muslin (gauze, lace cotton, lawn)	linen, muslin, silk (gauze, cotton lawn)	muslin, silk, linen (lawn, cotton)	silk, muslin, cotton
breeches	cloth (leather)	cloth	cloth, silk, leather	leather, velvet	cloth, velvet, corduroy (leather, kersey, nankeen, fustian)	leather (velvet, corduroy, fustian)
waistcoats	(cloth)	(cloth)	cloth, silk, linen (velvet, flannel)	cotton, woollen dimity (linen, serge)	(silk, linen, cloth, velvet, cotton, dimity, nankeen)	(silk)
coats	(cloth, velvet)	cloth	cloth, woollen, (line, velvet)	cloth	cloth (dimity)	cloth (fustian)

Notes
The material from which things are made tend not to be recorded in the proceedings in 1820–1.
Materials ordered by frequency of occurrence in each year. Items in parentheses indicate only one item made out of this type of material recorded.

The ascendancy of cotton obviously relates to contemporaneous changes in production. Cromptons' Mule of 1779 enabled a cheaper and finer cotton yarn to be spun and this led to the production of an all cotton cloth at lower price (Mokyr, 1990, p. 98). But the desirability of cotton was aided by its amenability to printing and thus the replication of the bright colours and intricate patterns found on silk (Styles, 2007, p. 127). Here too inventions helped. Berthollet changed bleaching technology in 1784 and this was improved by Charles Tennant with the invention of bleaching powder in 1799. Thomas Bell's metal printing cylinders (1783) enabled patterns to be printed on finished cloth (Mokyr, 1990, p. 99). But the repeal of the Acts prohibiting the wearing of pure cottons in 1774 must also have played their part and were, presumably, a response to demand pressures. Linen also underwent some production changes, power spinning was introduced in the Napoleonic Wars, around 1810 (Mokyr, 1985, p. 9), but there is no real evidence of this occasioning a fall in the price of shifts and shirts, if anything their price increased (Table 10.6). Indeed, considering the valuation given to articles made from different materials enables us to highlight the relative importance of demand compared with production factors in some of the shifts observed (Table 10.6). For instance, the shift towards cotton stockings after 1780 was not accompanied by price declines. Similarly cotton petticoats remained more expensive than those made of stuff, flannel or dimity, and cotton

Table 10.6 Value given to various items of clothing classified by the material from which they were made

Material	*1750–1*	*1760–1*	*1770–1*	*1780–1*	*1790–1*	*1800–1*
Stockings						
Worsted	0.067	–	0.059	–	0.051	–
Cotton	–	–	0.033	–	0.068	–
Gowns						
Linen	0.03	0.80	0.22	0.40	0.28	0.25
Silk	–	–	0.51	0.76	0.33	0.93
Cotton	–	–	0.35	0.76	0.32	0.23
Muslin	–	–	–	–	0.93	0.87
Calico	–	–	–	0.10	0.52	–
Petticoats						
Stuff	–	–	0.33	0.26	0.28	0.11
Dimity	–	–	0.18	0.24	0.13	0.17
Flannel	–	–	0.05	0.08	0.04	0.07
Cotton	–	–	–	0.36	0.56	0.20
Handkerchiefs						
Silk	0.11	0.08	0.11	0.11	0.11	0.10
Linen	0.05	0.04	0.07	0.05	0.07	–
Muslin	0.25	–	0.12	0.14	0.06	0.06
Cotton	–	–	0.03	0.02	0.02	0.03
Coats						
Cloth	0.25	0.48	1.30	0.35	0.71	0.72

gowns were more expensive than linen ones.[16] Only in the case of handkerchiefs do we see a clear cheapening of the product relative to the other materials with the advent of new technology.

10.5 Conclusions

Thefts from people's houses as documented in the Old Bailey records from 1750–1 to 1820–1 have been established as revealing much about the popularity and spread of particular goods. The shift in items of clothing stolen followed clear trends in fashions, new styles in clothing appeared and old ones faded from sight. Witness the rise of the pelisse and the demise of separate sleeves. Changes in the materials from which clothing items were made were also evident, for many items woollen garments were replaced by cotton and muslin variants. But what drove these changes? Both shifts in consumer tastes and changed production technologies which lowered prices played a role. Mokyr (1990, p. 111) argues forcefully that new technology dramatically reduced the prices of many products and so increased demand: consumers themselves played a passive role. Certainly, many of the shifts in desirable clothing noted here were associated with the introduction of new technology. Yet the valuations of stolen property belie the idea that price declines always drove consumption. Cotton was not the cheaper material for gowns or petticoats despite huge improvements in spinning, bleaching and printing processes. It would seem that the 'comfort revolution' driven by cupidity had a role to play too. On the evidence presented here, the expansion of consumer goods available in the latter part of the eighteenth century reflected the interplay of consumer revolution with industrial development.

Notes

1 See, for example, Styles (2007) on clothing.
2 Beattie (1986, pp. 252–63, 515) notes large gangs operating in London in the 1750s, 1760s and 1780s, (see p. 257, p. 259 in particular).
3 We have included all cases where theft was accused regardless of the verdict. Many trials were abandoned as the putative owner of the property either was not known or did not turn up at the trial. For our purposes, the fact that a particular good was supposedly stolen remains an indication of its desirability, regardless of culpability.
4 Non-violent theft had been more than 80 per cent of the Courts' business in the eighteenth century; this reduced to around 5 per cent by the 1900s (Old Bailey, 2012).
5 Old Bailey Proceedings Online (www.oldbaileyonline.org) reference t17900526–5, accessed 12 March 2013.
6 The data set was constructed in SPSS and was non-rectangular.
7 Pearson correlation coefficient thefts and population 0.823, two-tailed test significant at 1 per cent level.
8 The level of prosecutions for theft in 1770–1 appears high and may be a response to the purported crime waves identified in 1763 and 1780s, but is also consistent with the view that crime in the capital was multiplying towards the end of the eighteenth century (Beattie, 1986, pp. 14–15).
9 Pearson correlation coefficient thefts and bread price 0.799, two-tailed test significant at 2 per cent level.

10 The average value of an item is determined per item stolen. That is, we take the average value of an individual spoon if 20 identical spoons were taken in a burglary.

11 Individual values were not attached to a large proportion of the items stolen in 1750–1 (69 per cent) and 1760–1 (48 per cent) but this omission reduced to less than 1 per cent of items from 1780–1 onwards.

12 This remains true when the median value of goods stolen, rather than the mean value, is used.

13 Note that this does not represent the total number of clothing items stolen as we count a number of items of the same type taken in a theft only once, for example the theft of five pairs of stockings counts as only one item in our database. See discussion above.

14 See Styles (2007, pp. 127–31) for a full discussion and refutation of the argument that cotton was replacing linen for undergarments.

15 This change is described fully in Styles (2007).

16 Styles (2007, p. 126) also notes that cotton gowns were not cheaper than linen ones in the 1770s and 1780s therefore considers they must have had a quality advantage.

References

Allen, Robert (2009) *The British Industrial Revolution in Global Perspective*, Cambridge: Cambridge University Press.

Beattie, John M. (1974) The pattern of crime in England 1660–1800, *Past and Present*, 62, 47–95.

Beattie, John M. (1986) *Crime and the Courts in England, 1660–1800*, Oxford: Clarendon Press.

Beattie, John M. (2001) *Policing and Punishment in London, 1660–1750*, Oxford: Oxford University Press.

Berg, Maxine (2005) *Luxury and Pleasure in Eighteenth-Century Britain*, Oxford: Oxford University Press.

Boulton, Jeremy (2000) London 1540–1700, in P. Clark (ed.) *The Cambridge Urban History of Britain*, *Volume 2*, Cambridge: Cambridge University Press, pp. 315–46.

Cox, Jeff and Nancy Cox (2000) Probate 1500–1800: a system in transition, in T. Arkell, N. Evans and N. Goose (eds) *When Death Us Do Part*, Oxford: Leopards Head Press, pp. 14–71.

Davies, David (1795) *The Case of Labourers in Husbandry*, London.

De Vries, Jan (2008) *The Industrious Revolution*, Cambridge: Cambridge University Press.

Devereaux, Simon (1996) The City and the Sessions papers; 'Public justice' in London, 1770–1800, *Journal of British Studies*, 35, 466–503.

Eden, Sir. F. M. (1797) *The State of the Poor: A History of the Labouring Classes of England. Volumes I–III*, London.

Engels, Friedrich (1845) *The Condition of the Working Class in England* (repr. 1987), Harmondsworth: Penguin.

Fine, Ben and Ellen Leopold (1993) *The World of Consumption*, Abingdon: Routledge.

French, Henry (2007) *The Middle Sort of People in Provincial England 1600–1750*, Oxford: Oxford University Press.

Hay, Douglas (1982) War, dearth and theft in the eighteenth century: the record of the English courts, *Past and Present*, 95, 117–60.

Innes, Joanna and John Styles (1986) The crime wave: recent writing on crime and criminal justice in eighteenth-century England, *Journal of British Studies*, 25, October, 380–435.

King, Peter (1984) Decision-makers and decision making, *Historical Journal*, 27, 25–58.

King, Peter (1997) Pauper inventories and the material lives of the poor in the eighteenth and early nineteenth centuries, in T. Hitchcock, P. King and P. Sharpe (eds)

Chronicling Poverty. The Voices and Strategies of the English Poor, 1640–1840, Basingstoke: Macmillan, pp. 155–91.

Langbein, John (2003) *The Origins of Adversary Criminal Trial*, Oxford: Oxford University Press.

Lemire, Beverly (1988) Consumerism in early Industrial England: the trade in second hand clothes, *Journal of British Studies*, 27, 1–24.

Lemire, Beverly (2006) Plebian commercial circuits and everyday material exchange in England, *c.*1600–1900, in B. Blonde, P. Stabel, J. Stobart and I. Van Damme (eds) *Buyers and Sellers: Retail Circuits and Practices in Medieval and Early Modern Europe*, Studies In European Urban History, 9, BREPOLS, Belgium, pp. 245–351.

Mayhew, Henry (1861) *London Labour and the London Poor: The Condition and Earnings of those that will Work, cannot Work and will not Work*, London.

Mitchell, Brian (1988) *British Historical Statistics*, Cambridge: Cambridge University Press.

Mokyr, Joel (1985) Industrial Revolution, in J. Mokyr (ed.) *The Economics of the Industrial Revolution*, London: Allen & Unwin.

Mokyr, Joel (1990) *The Lever of Riches. Technical Creativity and Economic Progress*, Oxford: Oxford University Press.

Office of National Statistics (2002) *Social Trends* No. 32, ed. Jill Matheson and Penny Babb, London: HMSO.

Old Bailey (2012) www.oldbaileyonline.org (accessed, most recently, 12 March 2013)

Overton, Mark, Jane Whittle, D. Dean and A. Hann (2004) *Production and Consumption in English Households 1600–1750*, Abingdon: Routledge.

Riello, George (2006) *A Foot in the Past: Consumers, Producers and Footwear in the Long Eighteenth Century*, Oxford: Oxford University Press.

Schwartz, Leonard (2000) London 1700–1840, in P. Clark (ed.) *The Cambridge Urban History of Britain, Volume 3*, Cambridge: Cambridge University Press, pp. 641–71.

Shammas, Carole (1990) *The Pre-Industrial Consumer in England and America*, Oxford: Clarendon Press.

Shoemaker, R. B. (1991) *Prosecution and Punishment: Petty Crime and the Law in London and Rural Middlesex,* c.*1660–1725*, Cambridge: Cambridge University Press.

Shoemaker, Robert Brink (2008) The Old Bailey Proceedings and the representation of crime and criminal justice in eighteenth century London, *Journal of British Studies*, 47(3), 559–80.

Sneath, Ken (2009) *Consumption, wealth, indebtedness and social structure in early modern England*, unpublished Ph.D. thesis, University of Cambridge.

Stobart, Jon (2006) Clothes, cabinets and carriages: second-hand dealing in eighteenth century England, in B. Blonde, P. Stabel, J. Stobart and I. Van Damme (eds) *Buyers and Sellers: Retail Circuits and Practices in Medieval and Early Modern Europe*, Studies In European Urban History, 9, BREPOLS, Belgium, pp. 225–44.

Styles, John (2006) Lodging at the Old Bailey: lodgings and their furnishings in eighteenth century London, in John Styles and Amanda Vickery (eds) *Gender, Taste and Material Culture in Britain and North America 1700–1830*, New Haven, CT: Yale University Press, pp. 61–80.

Styles, John (2007) *The Dress of the People: Everyday Fashion in Eighteenth-century England*, New Haven: Yale University Press.

Weatherill, Lorna (1988) *Consumer Behaviour and Material Culture 1660–1760*, London: Routledge.

Wrigley, E. A. (2011) *Energy and the English Industrial Revolution*, Cambridge: Cambridge University Press.

Index

Page numbers in *italics* denote tables, those in **bold** denote figures.