



# Applied Health Economics

Large-scale survey datasets, in particular complex survey designs such as panel data, provide a rich source of information for health economists. They offer the scope to control for individual heterogeneity and to model the dynamics of individual behaviour. However the measures of outcome used in health economics are often qualitative or categorical. These create special problems for estimating econometric models. The dramatic growth in computing power over recent years has been accompanied by the development of methods that help to solve these problems. This book provides a practical guide to the skills required to put these techniques into practice.

Jones *et al.* illustrate practical applications of these methods using data on health from, among others, the British Health and Lifestyle Survey (HALS), the British Household Panel Survey (BHPS), the European Community Household Panel (ECHP) and the WHO Multi-Country Survey Study (WHOMCS). Assuming a familiarity with the basic syntax and structure of Stata, this book presents and explains the statistical output using empirical case studies rather than general theory.

A distinctive feature of the text is the way that it brings together theory and practice. This book will be of great benefit to applied economists, as well as advanced undergraduate and post-graduate students in health economics and applied econometrics.

**Andrew M. Jones** is Director of the Graduate Programme in Health Economics at the University of York.

**Nigel Rice** is Reader in Health Economics at the University of York.

**Teresa Bago d’Uva** is Assistant Professor at Erasmus University.

**Silvia Balia** is a researcher at the Department of Economic and Social Research, University of Cagliari.





—

—

—



—



—



# Applied Health Economics

**Andrew M. Jones, Nigel Rice,  
Teresa Bago d'Uva and Silvia Balia**



 **Routledge**  
Taylor & Francis Group  
LONDON AND NEW YORK





First published 2007  
by Routledge  
2 Park Square, Milton Park, Abingdon OX14 4RN

Simultaneously published in the USA and Canada  
by Routledge  
270 Madison Ave, New York, NY 10016

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

© 2007 Andrew M. Jones, Nigel Rice, Teresa Bago d'Uva and  
Silvia Balia

Typeset in 10/12pt Times NR MT by Graphicraft Limited, Hong Kong  
Printed and bound in Great Britain by

All rights reserved. No part of this book may be reprinted or  
reproduced or utilized in any form or by any electronic,  
mechanical, or other means, now known or hereafter  
invented, including photocopying and recording, or in any  
information storage or retrieval system, without permission in  
writing from the publishers.

*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*

ISBN10: 0-415-39771-5 (hbk)  
ISBN10: 0-415-39772-3 (pbk)  
ISBN10: 0-203-97230-9 (ebk)

ISBN13: 978-0-415-39771-1 (hbk)  
ISBN13: 978-0-415-39772-8 (pbk)  
ISBN13: 978-0-203-97230-4 (ebk)



# Contents

|   |           |
|---|-----------|
| <i>Preface</i>  | viii      |
| <i>Acknowledgements</i>   | x         |
| <b>Introduction</b>   | <b>1</b>  |
| <b>PART I</b>   |           |
| <b>Data description</b>   | <b>3</b>  |
| <b>1 Data and survey design</b>                                     | <b>5</b>  |
| 1.1 <i>The Health and Lifestyle Survey (HALS)</i>                   | 5         |
| 1.2 <i>The British Household Panel Survey (BHPS)</i>                | 7         |
| 1.3 <i>The European Community Household Panel (ECHP)</i>            | 9         |
| 1.4 <i>The Canadian National Population Health Survey (NPHS)</i>    | 10        |
| 1.5 <i>The WHO Multi-Country Survey Study (Who-MCS)</i>             | 12        |
| 1.6 <i>Overview</i>   | 12        |
| <b>2 Describing the dynamics of health</b>                          | <b>13</b> |
| 2.1 <i>Introduction</i>   | 13        |
| 2.2 <i>Graphical analysis</i>                                       | 15        |
| 2.3 <i>Tabulating the data</i>                                      | 22        |
| <b>3 Inequality in health utility and self-assessed health</b>      | <b>29</b> |
| 3.1 <i>Introduction</i>   | 29        |
| 3.2 <i>Distributional analysis</i>                                  | 30        |
| 3.3 <i>Regression analysis of HUI: Ordinary Least Squares (OLS)</i> | 34        |
| 3.4 <i>Regression analysis of SAH: ordered probit model</i>         | 37        |
| 3.5 <i>Combined analysis of HUI and SAH: interval regression</i>    | 45        |
| 3.6 <i>Overview</i>   | 49        |



vi *Contents*

**PART II**

**Categorical data 51**

**4 Bias in self-reported data 53**

*4.1 Introduction 53*

*4.2 Vignettes 54*

*4.3 Standard ordered probit model 56*

*4.4 Using vignettes to control for heterogeneous reporting 58*

**5 Health and lifestyles 81**

*5.1 Introduction 81*

*5.2 HALS data and sample 82*

*5.3 Descriptive analysis 88*

*5.4 Estimation strategy and results 97*

*5.5 Discussion 121*

**PART III**

**Survival data 123**

**6 Smoking and mortality 125**

*6.1 Introduction 125*

*6.2 Basic concepts of survival analysis 126*

*6.3 The hals data 127*

*6.4 Survival data in HALS 129*

*6.5 Descriptive statistics 134*

*6.6 Duration models 136*

**7 Health and retirement 169**

*7.1 Introduction 167*

*7.2 Preparing and summarizing the data 169*

*7.3 The endogeneity of health 180*

*7.4 Empirical approach to duration modelling 183*

*7.5 Stock sampling and discrete-time hazard analysis 189*

*7.6 Overview 199*



|  |            |
|--|------------|
| <b>PART IV</b>                                 |            |
| <b>Panel data</b>                              | <b>201</b> |
| <b>8 Health and wages</b>                      | <b>203</b> |
| 8.1 <i>Introduction</i>                        | 203        |
| 8.2 <i>BHPS sample and variables</i>           | 203        |
| 8.3 <i>Empirical model and estimation</i>      | 209        |
| 8.4 <i>Discussion</i>                          | 225        |
| <b>9 Modelling the dynamics of health</b>      | <b>227</b> |
| 9.1 <i>Introduction</i>                        | 227        |
| 9.2 <i>Static models</i>                       | 233        |
| 9.3 <i>Dynamic models</i>                      | 249        |
| <b>10 Non-response and attrition bias</b>      | <b>265</b> |
| 10.1 <i>Introduction</i>                       | 265        |
| 10.2 <i>Testing for non-response bias</i>      | 268        |
| 10.3 <i>Estimation</i>                         | 272        |
| <b>11 Models for health-care use</b>           | <b>279</b> |
| 11.1 <i>Introduction</i>                       | 279        |
| 11.2 <i>The poisson model</i>                  | 280        |
| 11.3 <i>The negative binomial model</i>        | 283        |
| 11.4 <i>Zero inflated models</i>               | 286        |
| 11.5 <i>Hurdle models</i>                      | 291        |
| 11.6 <i>Finite mixture latent class models</i> | 293        |
| 11.7 <i>Latent class hurdle model</i>          | 299        |
| <br>   |            |
| <i>Bibliography</i>                            | 321        |
| <i>Index</i>                                   | 330        |





## Preface

Large-scale survey datasets, in particular complex survey designs such as panel data, provide a rich source of information for health economists. They offer the scope to control for individual heterogeneity and to model the dynamics of individual behaviour. However the measures of outcome used in health economics are often qualitative or categorical. These create special problems for estimating econometric models. The dramatic growth in computing power over recent years has been accompanied by the development of methods that help to solve these problems. The purpose of this book is to provide a practical guide to the skills required to put these techniques into practice.

Practical applications of the methods are illustrated using data on health from, among others, the British Health and Lifestyle Survey (HALS), the British Household Panel Survey (BHPS), the European Community Household Panel (ECHP) and the WHO Multi-Country Survey Study (WHO-MCS). There is a strong emphasis on applied work, illustrating the use of relevant computer software with code provided for Stata ([www.stata.com](http://www.stata.com)). Familiarity with the basic syntax and structure of Stata is assumed. The Stata code and extracts from the statistical output are embedded directly in the main text and explained as we go along. The commands lines appear in the same format that they are recorded in the Stata log file, prefixed by ‘.’, for example:

- `use "c:\stata\data\bhps.dta", clear`

The Stata output appears alongside in a smaller font. The code presented in this book can be downloaded from the web at the homepage of the Health-Econometrics and Data Group, <http://www.york.ac.uk/res/herc/hedg.html>.

We do not attempt to provide a review of the extensive health economics literature that makes use of econometric methods (for a survey of the pre-2000 literature see Jones (2000) and for a collection of papers see Jones and O’Donnell (2002)). Instead, the book is built around empirical case studies, rather than general theory, and the emphasis is on learning by example. We present a detailed dissection of methods and results of some recent research



*Preface* ix

papers written by the authors and our colleagues. Relevant methods are presented alongside the Stata code that can be used to implement them, and the empirical results are discussed as we go along. To our knowledge, no comparable text exists. There are health economics texts and there are econometrics texts but these tend to focus on theory rather than application and tend not to bring the two disciplines together for the benefit of applied economists. The emphasis is on hands-on empirical analysis: the kind of thing that econometric texts tend to neglect. The closest in spirit is Angus Deaton's (1997) excellent book on the analysis of household surveys, but that emphasizes issues in the economics of development, poverty and welfare rather than health. A general knowledge of microeconomic methods is assumed. For more details readers can refer to texts such as Baltagi (2005), Cameron and Trivedi (2005), Greene (2003) and Wooldridge (2002b).

As the book is built around case studies, and these reflect the particular interests of the authors, we do not claim to cover the full diversity of topics within applied health economics. However we hope that these examples will provide guidance and inspiration for those working on other topics within the field who want to make use of econometric methods. The book is primarily aimed at advanced undergraduates and postgraduates in health economics, along with health economics researchers in academic, government and private sector organizations who want to learn more about empirical research methods. In addition the book may be used by other applied economists, in areas such as labour and environmental economics, and by health and social statisticians.





## Acknowledgements

Data from the *British Household Panel Survey (BHPS)* were supplied by the UK Data Archive. Neither the original collectors of the data nor the archive bear any responsibility for the analysis or interpretations presented here. The *European Community Household Panel Users' Database (ECHP)*, version of December 2003, was supplied by Eurostat. Data from the *Health and Lifestyle Survey (HALS)* were supplied by the UK Data Archive. Neither the original collectors of the data nor the archive bear any responsibility for the analysis or interpretation presented here. We are grateful to Statistics Canada for access to the *National Population Health Survey (NPHS)* data. We thank the World Health Organization for providing access to the *WHO Multi-Country Survey Study (WHO-MCS)* data.

We are very grateful to all of the co-authors of the joint work that we use as case studies: Somnath Chatterji, Paul Contoyannis, Martin Forster, Cristina Hernández-Quevedo, Xander Koolman, Maarten Lindeboom, Owen O'Donnell, Jennifer Roberts and Eddy van Doorslaer. The specific papers that are adapted for the case studies are:

- Bago d'Uva, T. (2006) 'Latent class models for health care utilisation', *Health Economics*, 15: 329–343.
- Bago d'Uva, T., van Doorslaer, E., Lindeboom, M., O'Donnell, O. and Chatterji, S. (2006) 'Does reporting heterogeneity bias the measurement of health disparities?', HEDG Working Paper 06/03, University of York.
- Balia, S. and Jones, A.M. (2005) 'Mortality, Lifestyle and Socio-Economic Status', HEDG Working Paper 05/02, University of York.
- Contoyannis, P., Jones, A.M. and Rice, N. (2004) 'The dynamics of health in the British Household Panel Survey', *Journal of Applied Econometrics*, 19: 473–503.
- Contoyannis, P. and Rice, N. (2001) 'The impact of health on wages: Evidence from the British Household Panel Survey', *Empirical Economics*, 26: 599–622.
- Forster, M. and Jones, A.M. (2001) 'The role of tobacco taxes in starting and quitting smoking: duration analysis of British data', *Journal of the Royal Statistical Society Series A*, 164: 517–547.



*Acknowledgements* xi

- Jones, A.M., Koolman, X. and Rice, N. (2006) 'Health-related non-response in the BHPS and ECHP: using inverse probability weighted estimators in nonlinear models', *Journal of the Royal Statistical Society Series A*, 169: 543–569.
- van Doorslaer, E. and Jones, A.M. (2003) 'Inequalities in self-reported health: validation of a new approach to measurement', *Journal of Health Economics*, 22: 61–87.

A draft of the book was used as teaching material for a short course 'Applied Health Economics' that was hosted by the Health, Econometrics and Data Group (HEDG) at the University of York, 19<sup>th</sup>–30<sup>th</sup> June 2006. This course was part of the Marie Curie Training Programme in Applied Health Economics. We are grateful for the input from other members of HEDG who were involved with the course: Cristina Hernández Quevedo, Eugenio Zuccheli, Silvana Robone, Pedro Rosa Dias and Rodrigo Morena Serra. Also we should like to thank the course participants for their valuable feedback on the material.

Finally we should like to thank Rob Langham at Routledge for encouraging us to take on this project and for his patience and support in seeing it through.





—

—

—



—

—



—