

PETER G. MOFFATT

EXPERIMETRICS

ECONOMETRICS FOR
EXPERIMENTAL
ECONOMICS

Experiments

Econometrics for Experimental Economics

Peter G. Moffatt

© Peter G. Moffatt 2016

All rights reserved. No reproduction, copy or transmission of this publication may be made without written permission.

No portion of this publication may be reproduced, copied or transmitted save with written permission or in accordance with the provisions of the Copyright, Designs and Patents Act 1988, or under the terms of any licence permitting limited copying issued by the Copyright Licensing Agency, Saffron House, 6–10 Kirby Street, London EC1N 8TS.

Any person who does any unauthorized act in relation to this publication may be liable to criminal prosecution and civil claims for damages.

The author has asserted his right to be identified as the author of this work in accordance with the Copyright, Designs and Patents Act 1988.

First published 2016 by
PALGRAVE

Palgrave in the UK is an imprint of Macmillan Publishers Limited, registered in England, company number 785998, of 4 Crinan Street, London N1 9XW.

Palgrave Macmillan in the US is a division of St Martin's Press LLC, 175 Fifth Avenue, New York, NY 10010.

Palgrave is a global imprint of the above companies and is represented throughout the world.

Palgrave® and Macmillan® are registered trademarks in the United States, the United Kingdom, Europe and other countries.

ISBN 978–0–230–25022–2 hardback

ISBN 978–0–230–25023–9 paperback

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources. Logging, pulping and manufacturing processes are expected to conform to the environmental regulations of the country of origin.

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

Printed and bound by CPI Group (UK) Ltd, Croydon, CR0 4YY

Praise for *Experimetrics*

'Experimental data seem deceptively easy. This book not only shows why and when estimation is involved. It also pushes the frontier, e.g. by introducing methods to capture patterned heterogeneity, random utility, or two-part decision rules.'

Christoph Engel, Max Planck Institute, Germany

'A long-awaited, systematic treatise of the econometric modelling of experimental data brilliantly accomplished. A work of art!' **Anna Conte, University of Westminster, UK**

'*Experimetrics* provides an excellent overview of the issues concerning the econometric analysis of experimental data. Numerous STATA codes enrich the book and make the methods very accessible.' **Charles Bellemare, Université Laval, Canada**

'I believe *Experimetrics* should become an indispensable part of every experimentalist's toolkit.' **David Butler, Murdoch University, Australia**

'This text provides guidance and reference for the increasingly large number of experimental economists who want to do some serious econometric analysis of their experimental data. Not only does it explain the concepts cleanly and precisely, it also provides numerous examples of applications. It will become compulsory reading for all experimental economists.' **John Hey, University of York, UK**

'*Experimetrics* is a very well-written book, providing an outstanding in-depth analysis of the econometric methodology into economics experiments. It is structured in a balanced way covering the most topical issues in the experimental literature. Superb writing style and coverage!' **Michalis Drouvelis, University of Birmingham, UK**

Experimetrics

Acknowledgements

This book is, in large part, the end-product of a sequence of masterclasses in Experimentics that I have delivered in various locations since 2006. I thank various people for arranging these courses and for providing hospitality during the visits, including Catherine Eckel, Tobias Uske, Anna Conte, Sylvia Arnold, Vittoria Levati, Angela Fletcher, Lionel Page, Uwe Dulleck, David Butler, Steven Schilizzi, Arne Weiss, and Gina Neff. I also thank the participants on these masterclasses for providing the opportunity to test the material on live audiences. I am grateful to the editors at Palgrave, Jaime Marshall, Aléa Bezuidenhout, Kirsty Reade, Jo Endell-Cooper, and Georgia Walters for their persistent encouragement towards completion of the book, to Alex Antidius Arpoudam of Integra Software Services for managing the book's production so professionally, and to the proofreader, Sarah Perkins, for her meticulous reading of every page of the manuscript and for finding typos and infelicities that I would never have found myself. I am grateful to Phil Bacon and Linford Bacon for providing expert assistance in converting the manuscript to L^AT_EX. I am also grateful to a number of (partially) anonymous reviewers who have provided invaluable, highly expert feedback and recommendations at various stages of the project – the feedback of Ryan Oprea and John Hey has been particularly influential. Finally, I am grateful to Robert Sugden for suggesting to me – all those years ago – that it would be a good thing if more econometricians were interested in the analysis of experimental data.

Contents

Acknowledgements	xv
1 Introduction and Overview	1
1.1 What is Experiments?	1
1.2 Experimental Design	2
1.3 The Experiments of Theory Testing	3
1.4 Dependence in Experimental Data	5
1.5 Parametric versus Non-parametric Approaches	6
1.6 Structural Experiments	7
1.7 Modelling Subject Heterogeneity	9
1.8 Experiments of Other-regarding Preferences	10
1.9 Experiments of Bounded Rationality	11
1.10 Experiments of Learning	12
1.11 What is in this Book?	13
2 Statistical Aspects of Experimental Design in Experimental Economics	17
2.1 Introduction	17
2.2 The Average Treatment Effect	17
2.3 Randomisation Techniques	18
2.3.1 Completely randomised designs	19
2.3.2 Factorial designs	19
2.3.3 Block designs	20
2.3.4 Within-subject designs	20
2.3.5 Crossover designs	21
2.3.6 ABA designs	21
2.4 How Many Subjects? A Primer in Power Analysis	21
2.4.1 The case of one sample	21
2.4.2 Choosing the sample size in a treatment test	24
2.4.3 Treatments with unequal costs	26
2.4.4 Sample size in cluster designs	27
2.5 Four Very Popular Experiments	28

2.5.1	The ultimatum game	28
2.5.2	The dictator game	29
2.5.3	The trust game	29
2.5.4	The public goods game	29
2.6	Other Aspects of Design	30
2.6.1	The random lottery incentive (RLI) mechanism	30
2.6.2	The strategy method	31
2.6.3	“One-shot”, “partners”, and “strangers” designs	31
2.7	Summary and Further Reading	32
3	Treatment Testing	35
3.1	Introduction	35
3.2	The Mechanics of Treatment Testing	36
3.3	Testing with Discrete Outcomes	37
3.3.1	The binomial test	37
3.3.2	Fisher’s exact test	38
3.3.3	The chi-squared test	41
3.3.4	The chi-squared test on a real data set	43
3.4	Testing for Normality	46
3.5	Treatment Testing	48
3.5.1	Parametric tests of treatment effects	48
3.5.2	Non-parametric tests of treatment effects: the Mann-Whitney test	51
3.5.3	The bootstrap	52
3.5.4	Tests comparing entire distributions	55
3.6	Testing for Gender Effects	57
3.7	Within-subject Tests	60
3.7.1	The Allais paradox	61
3.7.2	Preference reversals	64
3.7.3	Continuous outcome	65
3.8	Summary and Further Reading	68
4	Theory Testing, Regression, and Dependence	71
4.1	Introduction	71
4.2	Experimental Auctions	73
4.2.1	Overview of auction theory	73
4.2.2	Carrying out an experimental auction	75
4.2.3	The simulated auction data	75
4.3	Tests of Auction Theory	78
4.3.1	A test of RNNE in a second-price common value auction	78
4.4	Tests of Comparative Static Predictions	80
4.4.1	Standard treatment tests	80
4.4.2	Treatment testing using a regression	82
4.4.3	Accounting for dependence: the ultra-conservative test	82
4.4.4	Accounting for dependence in a regression	84
4.4.5	Accounting for dependence: the block bootstrap	85

4.5	Multiple Regression with Auction Data	86
4.5.1	Introducing the effect of uncertainty	86
4.5.2	Introducing the effect of experience	88
4.5.3	Introducing the effect of cash balance	89
4.6	Panel Data Estimators	90
4.7	Multi-level Modelling	92
4.8	Modelling Data from Contest Experiments	97
4.8.1	The Tullock contest	97
4.8.2	A contest experiment	98
4.8.3	Analysis of data from contest experiment	99
4.9	Meta-analysis	102
4.10	Summary and Further Reading	105
5	Modelling of Decision Times Using Regression Analysis	107
5.1	Introduction	107
5.2	The Decision-time Data	108
5.3	A Theoretical Model of Effort Allocation	109
5.4	An Econometric Model of Effort Allocation	111
5.5	Panel Data Models of Effort Allocation	116
5.6	Discussion of Results	120
5.7	Post-estimation	121
5.8	Summary and Further Reading	123
6	Dealing with Discreteness in Experimental Data	125
6.1	Introduction	125
6.2	Binary Data	126
6.2.1	Likelihood and log-likelihood explained	126
6.2.2	Modelling choices between lotteries (the “house money effect”)	129
6.2.3	Marginal effects	133
6.2.4	Wald tests and LR tests	134
6.2.5	Analysis of ultimatum game data	135
6.2.6	The strategy method	138
6.3	The <code>m1</code> Routine in STATA	140
6.4	Structural Modelling	141
6.5	Further Structural Modelling	144
6.5.1	The heterogeneous agent model	144
6.5.2	The delta method	147
6.5.3	Another example using the delta method	148
6.6	Other Data Types	149
6.6.1	Interval data: the interval regression model	149
6.6.2	Continuous (exact) data	152
6.6.3	Censored data: the Tobit model	155
6.7	The Ultimatum Game: Further Analysis	160
6.7.1	Further tests of gender effects	160

6.7.2	The proposer's decision as a risky choice problem	161
6.8	Summary and Further Reading	164
7	Ordinal Data in Experiments	167
7.1	Introduction	167
7.2	Ordered Outcomes: The Case for Special Treatment	168
7.3	The Ordered Probit Model: Theory	169
7.4	Interpreting the Cut-point Parameters	171
7.5	Application to Data on Emotions	172
7.6	Application to Data on Strength of Preference	177
7.7	Summary and Further Reading	181
8	Dealing with Heterogeneity: Finite Mixture Models	183
8.1	Introduction	183
8.2	Mixture of Two Normal Distributions	184
8.2.1	Data and model	184
8.2.2	Posterior type probabilities	186
8.2.3	The estimation program	186
8.2.4	Results	187
8.3	The <code>fmm</code> Command in STATA	188
8.4	A Mixture Model for the "Acquiring a Company" Task	190
8.5	A Mixture Model for Giving in a Public Goods Experiment	193
8.5.1	Background	193
8.5.2	Experiment	194
8.5.3	The data	195
8.5.4	The finite mixture two-limit Tobit model with tremble	197
8.5.5	Program	201
8.5.6	Results	206
8.5.7	Posterior type probabilities	208
8.6	Summary and Further Reading	210
9	Simulating Experimental Data, and the Monte Carlo Method	211
9.1	Introduction	211
9.2	Random Number Generation in STATA	212
9.3	Simulating Data Sets	214
9.3.1	Simulating data from a linear model	214
9.3.2	Simulating panel data	216
9.3.3	Simulating dynamic panel data	217
9.3.4	Simulating binary panel data	219
9.4	Monte Carlo Investigation of the Hausman Test	220
9.5	Summary and Further Reading	224

10	Introduction to the Method of Maximum Simulated Likelihood (MSL)	227
10.1	Introduction	227
10.2	The Principle of Maximum Simulated Likelihood (MSL)	228
10.3	Halton Draws	229
10.3.1	The case for using Halton draws	229
10.3.2	Generating Halton draws in STATA	230
10.3.3	Halton draws for panel data estimation	233
10.4	The Random Effects Probit Model	234
10.4.1	Model	234
10.4.2	Simulation	236
10.4.3	Estimation by MSL	238
10.4.4	Preparation of data for estimation	239
10.4.5	The likelihood evaluator	240
10.4.6	Complete annotated code	244
10.4.7	Results	246
10.5	The Random Effects Two-limit Tobit Model	248
10.5.1	Construction of log-likelihood function	249
10.5.2	Estimation by MSL	250
10.5.3	STATA code	251
10.5.4	Results from random effects two-limit Tobit model	253
10.6	Summary and Further Reading	255
11	Dealing with Zeros: Hurdle Models	257
11.1	Introduction	257
11.2	Review of Tobit and Random Effects Tobit	258
11.3	The Need for Hurdle Models	260
11.4	The Double Hurdle Model and Variants	261
11.4.1	p -Tobit	261
11.4.2	Double hurdle	261
11.4.3	The single hurdle model	263
11.5	The Panel Hurdle Model	264
11.5.1	The basic model	264
11.5.2	The panel single hurdle model	265
11.5.3	Construction of likelihood function	265
11.5.4	Panel hurdle with upper censoring	266
11.5.5	Panel hurdle model with tremble	266
11.5.6	Panel hurdle model with dependence	267
11.5.7	Obtaining posterior probabilities	268
11.5.8	Estimation	269
11.5.9	STATA code and simulation	269
11.6	A Panel Hurdle Model of Dictator Game Giving	275
11.6.1	The experiment	275
11.6.2	Estimation	278
11.6.3	Results	279

11.7	A Panel Hurdle Model of Contribution in a Public Goods Game	282
11.8	Summary and Further Reading	288
12	Choice under Risk: Theoretical Issues	291
12.1	Introduction	291
12.2	Utility Functions and Risk Aversion	291
12.3	Lottery Choice	295
12.4	Stochastic Dominance	298
12.5	Non-expected Utility Models	299
12.5.1	Weighting functions	299
12.6	Stochastic Models of Choice under Risk	302
12.7	Summary and Further Reading	304
13	Choice under Risk: Econometric Modelling	307
13.1	Introduction	307
13.2	The Choice Models	310
13.2.1	The framework	310
13.2.2	Incorporating violations of EU	311
13.2.3	The Fechner model	312
13.2.4	The Random Preference (RP) model	312
13.2.5	Dominance problems in the RP model	313
13.2.6	The tremble parameter	313
13.2.7	The role of experience	314
13.2.8	Between-subject variation and the sample log-likelihood ...	315
13.2.9	Posterior estimation of risk attitude	316
13.3	Simulation and Estimation	316
13.3.1	Data generating process	316
13.3.2	The STATA code	317
13.3.3	The simulated data	325
13.3.4	Output from estimation routines	326
13.4	Results and Post-estimation	327
13.4.1	The model estimates	327
13.4.2	Vuong's non-nested likelihood ratio test	328
13.4.3	Clarke's non-parametric non-nested test	329
13.4.4	Obtaining individual risk attitudes	331
13.4.5	Obtaining closeness to indifference	331
13.4.6	Simulation of decision times	332
13.5	Summary and Further Reading	333
14	Optimal Design in Binary Choice Experiments	335
14.1	Introduction	335
14.2	Rudiments of Experimental Design Theory	337
14.2.1	The principle of D-optimal design	337
14.2.2	Simple linear regression	338
14.2.3	Simple probit and simple logit	339

14.3	The Random Preference (RP) Model Revisited	341
14.4	Application of D-optimal Design Theory to a Risky Choice Experiment	344
14.5	Optimal Design in the Presence of a Tremble Parameter	345
14.6	Optimal Design for a Sample of Subjects	347
14.6.1	Choice of design used in Chapter 13	348
14.7	Summary and Further Reading	349
15	Social Preference Models	351
15.1	Introduction	351
15.2	Estimation of Preference Parameters from Dictator Game Data ...	352
15.2.1	The framework	352
15.2.2	The Andreoni-Miller data	352
15.2.3	Estimating the parameters of a CES utility function	357
15.3	A Model of Altruism with Binding Non-negativity Constraints ...	360
15.3.1	Background	360
15.3.2	The model	360
15.3.3	Estimation	363
15.3.4	Results	366
15.4	Finite Mixture Models of Altruism	367
15.4.1	The experiment	368
15.4.2	The data	368
15.4.3	A mixture model of fairness	368
15.4.4	The “fairness ideals”	370
15.4.5	The econometric model	370
15.4.6	The program and results	371
15.5	Estimation of Social Preference Parameters Using Discrete Choice Models	374
15.5.1	The setting	374
15.5.2	Formalising the criteria for choosing between allocations ...	375
15.5.3	Data	376
15.5.4	The conditional logit model (CLM)	376
15.5.5	Results	378
15.5.6	The effect of subject characteristics	378
15.6	Summary and Further Reading	379
16	Repeated Games and Quantal Response Models	383
16.1	Introduction	383
16.2	Analysis of Repeated Game Data	384
16.2.1	Computing a mixed strategy Nash equilibrium	384
16.2.2	Non-parametric tests on repeated game data	385
16.3	Quantal Response Equilibrium (QRE)	389
16.3.1	Theory of QRE	389
16.3.2	Computing the probabilities in the QRE model	391
16.4	Estimation of the QRE Model	392
16.5	The Risk-averse QRE Model	397

16.6	QRE Applied to Contest Data	400
16.7	Summary and Further Reading	406
17	Depth of Reasoning Models	407
17.1	Introduction	407
17.2	A Level-k Model for the Beauty Contest Game	408
17.3	The Cognitive Hierarchy Model	413
17.4	Summary and Further Reading	418
18	Learning Models	419
18.1	Introduction	419
18.2	Directional Learning (DL)	420
18.3	Data Used for Estimation of RL, BL and EWA	423
18.4	Notation Used in RL, BL and EWA	423
18.5	Reinforcement Learning (RL)	424
18.5.1	Program and results for RL	425
18.6	Belief learning (BL)	427
18.6.1	Program and results for BL	429
18.7	The Experience Weighted Attraction (EWA) Model	431
18.7.1	Introduction to EWA	431
18.7.2	Simulation of data set using EWA	432
18.7.3	Estimation of EWA Model	434
18.7.4	Results from EWA Model	436
18.8	Summary and Further Reading	439
19	Summary and Conclusion	441
19.1	Experimental Design Issues	441
19.2	Experimetrics and Theory Testing	443
19.3	Data Features	443
19.4	Experimetrics of Social Preferences	444
19.5	Risk Experimetrics	445
19.6	The Experimetrics of Games	447
19.7	Heterogeneity	448
A	List of Data Files and Other Files	451
B	List of STATA Commands	453
C	Choice Problems Used in Chapters 5 and 13	461
	References	463
	Index	473

Chapter 1

Introduction and Overview

1.1 What is Experimetrics?

*Experimetrics*¹ comprises the body of econometric techniques that are customized to experimental applications. A wide variety of such techniques appear in the experimental economics literature. The aim of this textbook is to assemble this body of techniques, to demonstrate their use in a hands-on style, drawing on as wide a range of examples as possible, and to interpret each set of results in ways that are most useful to experimental economists. The target audience is mainly researchers in experimental economics. It is also conceivable that the book might be of interest to econometricians who are curious to know what sort of techniques are being used by experimental economists.

The experimetric techniques that already appear in the experimental economics literature range from the very basic to the highly sophisticated. At the basic end of this spectrum we see the class of techniques known as *treatment tests*, that is, tests which compare outcomes with and without a treatment, or before and after a treatment. At the sophisticated end of the spectrum, we see highly complex structural models, with a deterministic core corresponding to an underlying behavioural theory, with possibly many structural parameters, and a stochastic specification including possibly many dimensions of variation, both within and between subject. Needless to say, the type of econometric approach that is chosen is often, and justifiably, dictated by the type of experiment that has been conducted, and by the types of research questions being addressed.

In some experiments, the “home-grown” characteristics of the experimental subjects are the focus, and the objectives are usually simply to investigate how individuals make decisions, or interact with each other, in particular settings. These

¹ The word “Experimetrics” was (to the best of my knowledge) coined by Camerer (2003, p. 42). Houser’s (2008) entry in the *New Palgrave Dictionary of Economics* on “Experiments and Econometrics” commences with the line “‘Experimetrics’ refers to formal procedures used in designed investigations of economic hypotheses.” Bardsley & Moffatt (2007) are apparently the first authors to have used the word in the title of a published paper.

studies typically rely on relatively simple experimental designs (e.g. choosing between lotteries; splitting a pie), with the ultimate objective of measuring subjects' characteristics, especially preference parameters. It is normal to expect substantial variation in these measured characteristics. Indeed, it is often the precise features of this variation in which we are most interested; for example, the proportion of the population who are "selfish", or the proportion who are EU maximisers. When data are from this type of experiment, it is often seen as appropriate to model the decision-making process using structural estimation methods, for example methods that simultaneously estimate all of the parameters appearing in the individuals' objective function, as well as distributional parameters, some of which capture preference heterogeneity.

In other experiments, the focus is on the functioning of an economic institution, rather than the characteristics of the individual participants within it, and the objective may be to test a particular theory as applying to that institution. In these settings, *induced value methodology* is commonly used. This technique is based on the idea that the appropriate use of a reward medium allows the experimenter to *induce* pre-specified characteristics (e.g. preferences) in the subjects so that their innate characteristics become irrelevant. Having essentially eliminated the impact of subjects' characteristics in this way, it is clearly much easier to apply close scrutiny to the theory under test. In these settings, the experimental designs are relatively complex, since key features of the economic institutions need to be captured in convincing ways. However, the econometric techniques required are often very simple. The level of control is normally such that straightforward treatment testing is often seen to be the natural framework, and the most suitable means of obtaining answers to the research questions of interest.

The following subsections provide brief overviews of the types of econometric techniques that are best suited to each of these two broad areas. As such, this chapter provides a broad overview of the contents of the remainder of the book.

1.2 Experimental Design

The topic of experimental design is clearly much more important in experimental economics than in other areas of economics. This is because, in other areas, typically the data generation process is out of the control of the investigator. In experimental economics, the data generation process is very much within the control of the investigator. Hence, design issues such as the choice of sample size, the sampling process, and the process for assignment of subjects to treatments, all take centre stage.

A central concept is randomization. If randomization is correctly applied in the process of selecting subjects for an experiment, then identification of treatment effects, which has always been a central problem in mainstream econometrics, is not a problem. The other side of this coin is that, since the data have not been collected in a natural environment, experimental results do not necessarily carry over to the world outside the lab. We therefore see that the advantages of experimental research

in terms of identification may be seen to be countered by the disadvantage of non-generalisability (see Al-Ubaydli & List, 2013).

There are several different types of design, including completely randomised designs, within-subject designs, crossover designs, and factorial designs. When groups of subjects are playing against each other, a choice also needs to be made between “partner” and “stranger” matching. Each design has both advantages and disadvantages, and the decision of which to use is often a delicate issue.

The choice of sample size is another key design decision. The question to be addressed here is how many subjects are required for the experimenter to be confident of reliable conclusions. More precisely, how many are required in each treatment? A useful framework for addressing these questions is power analysis (Cohen, 2013); that is, using probability theory to find what sample is required to provide a given “power” of the test being conducted.

A particularly interesting problem in experimental design is how to specify binary choice problems (e.g. lottery choice) in such a way as to generate a data set from which subjects’ preference parameters may be estimated with maximal precision. A chapter of the text is dedicated to this problem.

1.3 The Experiments of Theory Testing

It is often claimed the purpose of an experiment is to test an economic theory. From the econometrician’s perspective, a natural way to perceive such a test is as an assessment of whether the *predictions* from the theory provide good approximations to actual behaviour (i.e. the behaviour of subjects in the lab). From this perspective, the role of the experimenter is to find regularities in observed behaviour, and then to ask which theories are best able to account for these regularities.

Competitive equilibrium is the central concept in many theories. The objective of an experiment (a market experiment, say) may be simply to observe how close behaviour is to the competitive equilibrium, and we shall refer to this as testing the *fundamental prediction* of the theory. This type of experiment is usually performed using *induced value* methodology; that is, a system in which each subject is exogenously assigned a valuation of the object being traded, so that the complete demand and supply schedules, and therefore the equilibrium, are known by the experimenter. In this setting it is clearly a simple matter to use experimental data to assess how close actual behaviour is to the fundamental (equilibrium) prediction.

An important aspect of the fundamental prediction of the theory is that it often amounts to a “point prediction”, that is, it simply tells us that a particular decision variable will, given the known fixed values of the exogenous variables, take a particular value. There are no *free parameters* in the model that generates the fundamental prediction. To see how free parameters enter a model, consider the following example. Start with a benchmark model that is built on the assumption of risk neutrality (RN). Such a model is likely to lead to a “risk-neutral equilibrium” prediction, $y = \text{constant}$, where y is the decision variable. This benchmark model has no free parameters. Now consider what happens when we adopt the assumption of expected