Structural Analysis of Discrete Data with Econometric Applications

edited by
Charles F. Manski
and Daniel McFadden

The MIT Press
Cambridge, Massachusetts
London, England

Fourth printing, 1990

© 1981 by
The Massachusetts Institute of Technology

All rights reserved. No part of this book may be reproduced in any form or by any means, electronic or mechanical, including photocopying, recording, or by any information storage and retrieval system, without permission in writing from the publisher.

Printed and bound in the United States of America by Maple-Vail, Inc.

Library of Congress Cataloging in Publication Data
Main entry under title:

Structural analysis of discrete data with econometric applications.

Includes bibliographies and indexes.
1. Econometrics. I. Manski, Charles F. II. McFadden, Daniel.
III. Title: Discrete data with econometric applications.
HB139.S79 330'.028 81-1450
ISBN 0-262-13159-5 AACR2

Contents

	List of Contributors	xiii
	Preface	xv
	Editors' Introduction	xvii
I	**Statistical Analysis of Discrete Probability Models**	
1	**Alternative Estimators and Sample Designs for Discrete Choice Analysis** Charles F. Manski and Daniel McFadden	
1.1	Introduction	2
1.2	The Likelihood of an Observation under Alternative Stratified Sampling Processes	8
1.3	Estimation of the Choice Model Parameters	11
1.4	Estimation with p and **Q** Both Known	13
1.5	Estimation with p Known and **Q** Unknown	16
1.6	Estimation with p Unknown and **Q** Known	17
1.7	Estimation with p and **Q** Both Unknown	24
1.8	Estimation in a General Stratified Sample	28
1.9	Selection of a Sample Design and Estimation Method	31
1.10	Conclusion	34
1.11	Appendix: Consistency of the Estimators	36
1.12	Appendix: Asymptotic Normality	45
	References	49
2	**Efficient Estimation of Discrete-Choice Models** Stephen R. Cosslett	
2.1	Introduction	51
2.2	Discrete Choice Models	52
2.3	Stratified Sampling and Choice-Based Sampling	54
2.4	Generalized Choice-Based Sample	57
2.5	Sample with Known Aggregate Shares	59
2.6	Aggregate Shares Estimated from an Auxiliary Sample	60
2.7	Supplemented Sample	60
2.8	General Considerations in Maximum Likelihood Estimation	61
2.9	Notation for a General Choice-Based Sample	63
2.10	The Likelihood Function for Choice-Based Samples	64

2.11	Maximization of the Likelihood	66
2.12	Asymptotic Properties of the Unconstrained Estimator	68
2.13	Estimation of Aggregate Shares	71
2.14	The Unconstrained Maximum Likelihood Estimator	73
2.15	The Logit Model as a Special Case	75
2.16	Estimation with Known Aggregate Shares	76
2.17	Consistency of the Constraint Equations	79
2.18	Asymptotic Properties with Known Aggregate Shares	80
2.19	The Constrained Maximum Likelihood Estimator	82
2.20	Estimation of the Logit Model with Known Aggregate Shares	84
2.21	Estimation with Aggregate Shares Inferred from an Auxiliary Sample	85
2.22	Asymptotic Variance of the Auxiliary Sample Estimator	87
2.23	Special Cases of the Auxiliary Sample Estimator	89
2.24	Estimation with a Supplementary Sample	90
2.25	Comparison of Estimators and Sample Designs	93
2.26	Appendix: Conditions on the Choice Probability Model	105
2.27	Appendix: Derivation of Asymptotic Properties	107
2.28	Appendix: Alternative Estimators for Generalized Choice-Based Samples with Known Aggregate Shares	109
	References	110

II Dynamic Discrete Probability Models

3 Statistical Models for Discrete Panel Data
James J. Heckman

3.1	Introduction	114
3.2	A Framework for Analyzing Dynamic Choice	119
3.3	The General Model	121
3.4	An Independent Trials Bernoulli Model	124
3.5	A Random Effect Bernoulli Model and One-Factor Schemes	127
3.6	A Fixed Effect Bernoulli Model	133
3.7	Models with General Correlation in the Errors: The Concept of Heterogeneity Extended	134
3.8	Models with Structural State Dependence	138
3.9	A Renewal Model	144

3.10	A Model with Habit Persistence	145
3.11	Computation in the General Model	147
3.12	A Summary of Sections 3.2 through 3.11	148
3.13	Heterogeneity versus Structural State Dependence: An Application of the Preceding Models	150
3.14	Testing for Heterogeneity versus State Dependence	154
3.15	Analogies with Time-Series Models	161
3.16	Examples of Models that Generate Structural State Dependence	163
3.17	Summary and Conclusion	166
3.18	Appendix: Factor Analytic Probit Models	167
	References	175

4 The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process
James J. Heckman

4.1	Introduction	179
4.2	The Problem of Initial Conditions and Some Formal Solutions	181
4.3	Simpler Solutions and the Problem of Incidental Parameters	185
4.4	Some Monte Carlo Evidence	189
4.5	Conclusions	194
	References	195

III Structural Discrete Probability Models Derived from Theories of Choice

5 Econometric Models of Probabilistic Choice
Daniel McFadden

5.1	Economic Man	198
5.2	Discrete Choice	199
5.3	Probabilistic Consumer Theory	201
5.4	Probabilistic Choice Systems	201
5.5	The Random Utility Maximization Hypothesis	202
5.6	Stochastic Revealed Preference	204

5.7	Aggregation of Preferences	206
5.8	The Williams-Daly-Zachary Theorem	210
5.9	Criteria for Parametric Probabilistic Choice Systems	217
5.10	Specification of Variables	217
5.11	Functional Form	218
5.12	The Luce Model	221
5.13	Thurstone's Model V	223
5.14	Tversky Elimination Models	225
5.15	Generalized Extreme Value Models	226
5.16	Preference Trees	230
5.17	Estimation of Tree Extreme Value Models	238
5.18	Sequential Estimation	241
5.19	An Application	242
5.20	Appendix: Normalization in MNL and MNP Models	249
5.21	Appendix: Computational Formulas for a Simple Model	250
5.22	Appendix: Computational Formulas for the Nested Multinomial Logit Model	252
5.23	Appendix: Proof of Theorem 5.1	260
5.24	Appendix: The Elimination-by-Strategy Model	268
	References	269

6	**Random versus Fixed Coefficient Quantal Choice Models** Gregory W. Fischer and Daniel Nagin	
6.1	Introduction	273
6.2	Quantal Choice Theory and Variation in Tastes	273
6.3	An Empirical Comparison of the LPIID and RCCD Models	280
6.4	Details of the Experiment	280
6.5	Results	281
6.6	Analysis of Individual Respondents	282
6.7	A Comparison of LPIID Probit and RCCD Probit Estimation	287
6.8	Conclusions	297
6.9	Appendix: The Unestimable Models	299
6.10	Appendix: Mean Taste Estimates in the LPIID and RCCD Models	302
	References	304

7	**On the Use of Simulated Frequencies to Approximate Choice Probabilities** Steven R. Lerman and Charles F. Manski	
7.1	Introduction	305
7.2	The Simulated Frequency Method	306
7.3	Bayesian Approach	307
7.4	Estimation of a Function of a Collection of Probabilities	307
7.5	Application to the Calculation of Multinomial Probit Choice Probabilities	309
7.6	The Simulation Routine	311
7.7	The Clark Method	312
7.8	Numerical Test Objectives and Design	313
7.9	Test Results and Analysis	315
7.10	Conclusions	318
	References	319
8	**Application of a Continuous Spatial Choice Logit Model** Moshe Ben-Akiva and Thawat Watanatada	
8.1	Introduction	320
8.2	Basic Definitions	321
8.3	Spatial Aggregation	322
8.4	The Discrete Logit Model	322
8.5	Spatial Aggregation Using Continuous Functions	325
8.6	The Continuous Logit Model	327
8.7	A Parametric Example of Spatial Aggregation	329
8.8	Continuous Logit with Featureless Plane	332
8.9	Basic Operations of the MIT-TRANS Model	336
	References	342
IV	**Simultaneous Equations Models with Discrete Endogenous Variables**	
9	**Simultaneous Equations Models with Discrete and Censored Variables** Lung-Fei Lee	
9.1	Introduction	346

9.2	Two-Stage Methods and Amemiya's Principle	347
9.3	Structural Equations with Probit Structure	349
9.4	Structural Equations with Observable Continuous Endogenous Variables	351
9.5	Structural Equations with Censored Dependent Variables	352
9.6	Structural Equations with Tobit Structure	355
9.7	Switching and Censored Models with Sample Separation Information	358
9.8	Conclusion	362
	References	363

10 Stratification on Endogenous Variables and Estimation: The Gary Income Maintenance Experiment
Jerry A. Hausman and David A. Wise

10.1	Introduction	365
10.2	The Problem of Endogenous Sampling and Estimation Methods	367
10.3	Relative Efficiencies of Weighted Least Squares versus Maximum Likelihood Estimates	374
10.4	Empirical Results of the Selection Bias in the Gary Income Maintenance Experiment	379
10.5	Alternative Sampling Procedures	383
10.6	Conclusion	387
10.7	Appendix: Extension of the Analysis to Two Time Periods and Two Equations	388
	References	390

11 A Switching Simultaneous Equations Model of Physician Behavior in Ontario
Dale J. Poirier

11.1	Introduction	392
11.2	Econometric Model	393
11.3	Estimation	397
11.4	Estimation of the Switching (Option) Equation	402
11.5	Empirical Results	404
11.6	Estimated Option Equation	406
11.7	Estimated Referral Equation	411

11.8	Concluding Remarks	420
	References	421
12	**Constraints on the Parameters in Simultaneous Tobit and Probit Models** Peter Schmidt	
12.1	Introduction	422
12.2	Simultaneous Tobit Models	422
12.3	All Endogenous Variables Truncated	423
12.4	Some Endogenous Variables Truncated	424
12.5	Both Y and Y^* as Explanatory Variables	426
12.6	Simultaneous Probit Models	426
12.7	All Endogenous Variables Truncated	427
12.8	Some Endogenous Variables Truncated	430
12.9	Both Y and Y^* as Explanatory Variables	431
12.10	Conclusions	433
	References	434
13	**Estimating Credit Constraints by Switching Regressions** Robert B. Avery	
13.1	Introduction	435
13.2	The Supply of Debt	437
13.3	The Model and Data	439
13.4	Simultaneous Switching Regression and Linear Equations	446
13.5	The Evidence	453
13.6	Qualifications and Evaluations	461
13.7	Appendix: Proof of Theorem 13.1	462
13.8	Appendix: Empirical Reduced Form Equations	464
	References	471
	Index	473

List of Contributors

Robert B. Avery
Carnegie-Mellon University
Pittsburgh, Pa.

Moshe Ben-Akiva
Massachusetts Institute
of Technology
Cambridge, Mass.

Stephen R. Cosslett
Northwestern University
Evanston, Ill.

Gregory W. Fischer
Carnegie-Mellon University
Pittsburgh, Pa.

Jerry A. Hausman
Massachusetts Institute
of Technology
Cambridge, Mass.

James J. Heckman
University of Chicago
Chicago, Ill.

Lung-Fei Lee
University of Minnesota
Minneapolis, Minn.

Steven R. Lerman
Massachusetts Institute
of Technology
Cambridge, Mass.

Charles F. Manski
Hebrew University
Jerusalem, Israel

Daniel McFadden
Massachusetts Institute
of Technology
Cambridge, Mass.

Daniel Nagin
Carnegie-Mellon University
Pittsburgh, Pa.

Dale J. Poirier
University of Toronto
Toronto, Canada

Peter Schmidt
Michigan State University
East Lansing, Mich.

Thawat Watanatada
International Bank for
Reconstruction and Development
Washington, D.C.

David A. Wise
Harvard University
Cambridge, Mass.

Preface

The chapters in this volume are all original and previously unpublished major research contributions made by econometricians to the structural analysis of discrete data. Two factors led to our decision to organize this volume.

First, we feel that the piecemeal publication in the journals of new research in this field has made it difficult for econometricians not actively working on discrete data problems to overview the existing state of knowledge and the present frontiers of research. Coordinated publication of the basic findings in this new subject should lower the cost of entry into the field and speed dissemination of recent research into the graduate econometrics classroom.

Second, as the econometric literature on discrete data analysis has grown, and its contributions have matured, we have increasingly wished to communicate the concerns and results of this literature to the wider community of researchers involved in the structural analysis of discrete data, both in applied economics and in disciplines outside economics. As the editors' introduction to this volume emphasizes, there exist important interconnections between the econometric literature on discrete data and the work in statistics, biometrics, psychometrics, sociometrics, and other disciplines on discrete data analysis. We have tried to organize this volume so that readers from outside economics as well as applied economists will recognize the connections between the problems they face and the issues addressed in econometric analysis of discrete data and will find the results obtained here useful in their work.

Submissions of papers for possible publication in the volume were solicited by the editors in the spring of 1978. We are grateful to R. Avery, M. Ben-Akiva, S. Cosslett, G. Duncan, D. Gillen, J. Hausman, J. Heckman, L. Lee, S. Lerman, D. Nagin, D. Poirier, P. Schmidt, B. Singer, R. Westin, and D. Wise, each of whom reviewed one or more papers.

The editors and many of the contributors in this volume have benefited greatly from exchanges of results at a series of workshops on the econometric analysis of discrete data sponsored by the National Science Foundation. The editors wish to thank Carnegie-Mellon University, the University of California, Berkeley, and the University of Chicago for providing facilities for these workshops.