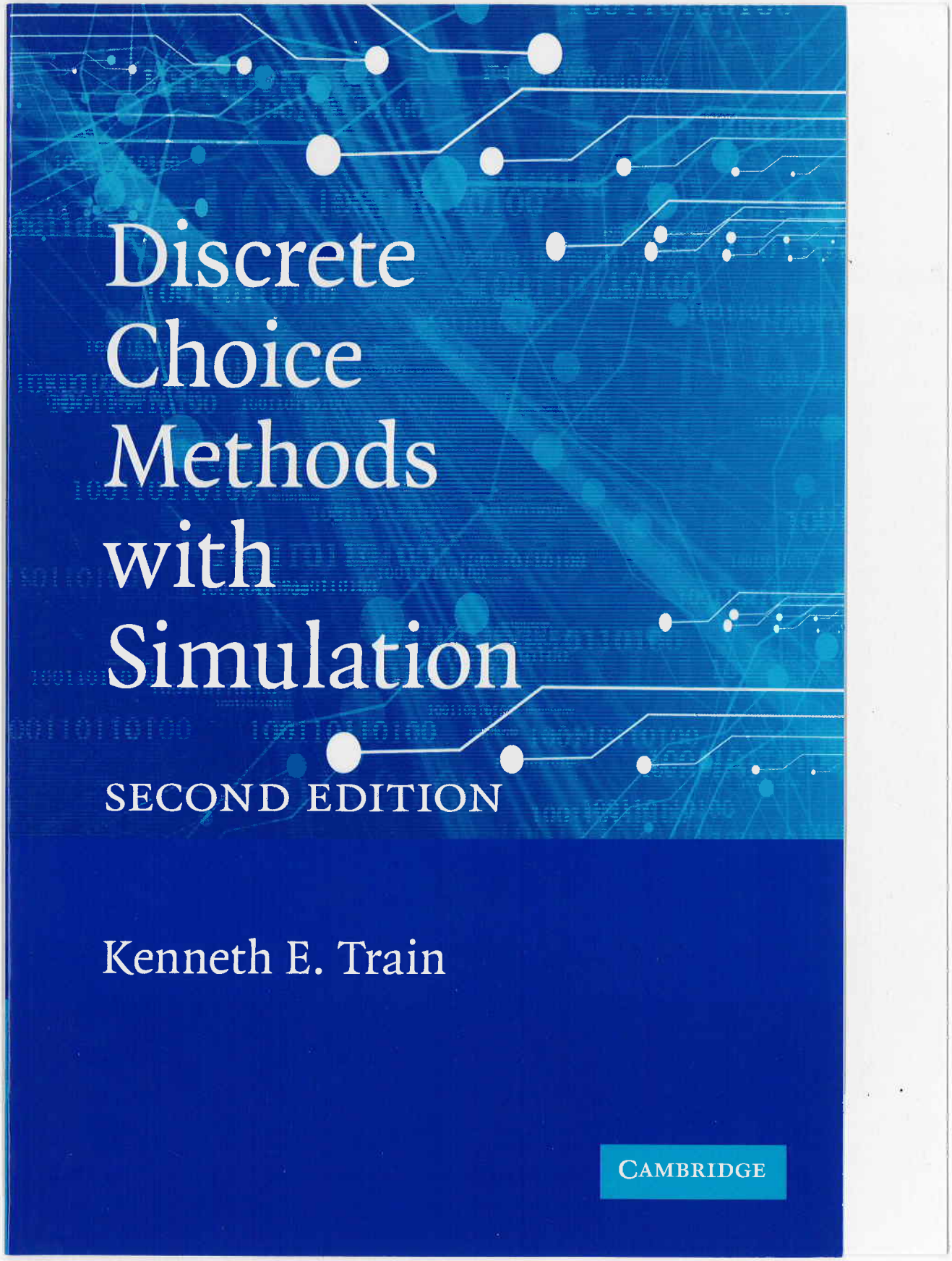
Discrete
Choice
Methods
with
Simulation
SECOND EDITION
Kenneth E. Train
CAMBRIDGE

Discrete Choice Methods with Simulation

Second Edition

This book describes the new generation of discrete choice methods, focusing on the many advances that are made possible by simulation. Researchers use these statistical methods to examine the choices that consumers, households, firms, and other agents make. Each of the major models is covered: logit, generalized extreme value (including nested and cross-nested logits), probit, and mixed logit, plus a variety of specifications that build on these basics. Simulation-assisted estimation procedures are investigated and compared, including maximum simulated likelihood, method of simulated moments, and method of simulated scores. Procedures for drawing from densities are described, including variance reduction techniques such as antithetics and Halton draws. Recent advances in Bayesian procedures are explored, including the use of the Metropolis–Hastings algorithm and its variant Gibbs sampling. This second edition adds chapters on endogeneity and expectation-maximization algorithms. No other book incorporates all these topics, which have arisen in the past 25 years. The procedures are applicable in many fields, including energy, transportation, environmental studies, health, labor, and marketing.

Professor Kenneth E. Train teaches econometrics, regulation, and industrial organization at the University of California, Berkeley. He also serves as Vice President of National Economic Research Associates (NERA), Inc., in San Francisco, California. The author of *Optimal Regulation: The Economic Theory of Natural Monopoly* (1991) and *Qualitative Choice Analysis* (1986), Dr. Train has written more than 60 articles on economic theory and regulation. He chaired the Center for Regulatory Policy at the University of California, Berkeley, from 1993 to 2000 and has testified as an expert witness in regulatory proceedings and court cases. He has received numerous awards for his teaching and research.

Additional Praise for the First Edition of *Discrete Choice Methods with Simulation*

"Ken Train's book provides outstanding coverage of the most advanced elements of the estimation and usage of discrete choice models that require simulation to take account of randomness in the population under study. His writing is clear and understandable, providing both the new and experienced reader with excellent insights into and understanding of all aspects of these new and increasingly important methods."

– Frank S. Koppelman, *Northwestern University*

"This is a masterful book, authored by one of the leading contributors to discrete choice methods and analysis. No other book covers this ground with such up-to-date detail in respect of theory and implementation. The chapters on simulation and recent developments such as mixed logit are most lucid. As a text or reference work this volume should have currency for a long time. It will appeal to the practitioner as much as to the specialist researcher who has been in this field for many years."

– David Hensher, *The University of Sydney*

"Simulation-based estimation is a major advance in econometrics and discrete choice modeling. The technique has revolutionized both classical and Bayesian analysis. Ken Train's many papers have made a large contribution to this literature. *Discrete Choice Methods with Simulation* collects these results in a comprehensive, up-to-date source, with chapters on behavioral foundations, theoretical and practical aspects of estimation, and a variety of applications. This book is a thoroughly enjoyable blend of theory, analysis, and case studies; it is a complete reference for developers and practitioners."

– William Greene, *New York University*

Discrete Choice Methods with Simulation

Second Edition

Kenneth E. Train

University of California, Berkeley, and NERA

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town, Singapore, São Paulo, Delhi

Cambridge University Press
32 Avenue of the Americas, New York, NY 10013-2473, USA

www.cambridge.org
Information on this title: www.cambridge.org/9780521747387

© Kenneth E. Train 2009

This publication is in copyright. Subject to statutory exception
and to the provisions of relevant collective licensing agreements,
no reproduction of any part may take place without the written
permission of Cambridge University Press.

First published 2009

Printed in the United States of America

A catalog record for this publication is available from the British Library.

Library of Congress Cataloging in Publication data

Train, Kenneth.

Discrete choice methods with simulation / Kenneth E. Train. – 2nd ed.

p. cm.

Includes bibliographical references and index.

ISBN 978-0-521-76655-5 (hardback) – ISBN 978-0-521-74738-7 (pbk.)

1. Decision making – Simulation methods. 2. Consumers' preferences – Simulation methods. I. Title.

HD30.23.T725 2009
003′.56–dc22 2009020438

ISBN 978-0-521-76655-5 hardback
ISBN 978-0-521-74738-7 paperback

Cambridge University Press has no responsibility for the persistence or accuracy of URLs for external or third-party Internet Web sites referred to in this publication and does not guarantee that any content on such Web sites is, or will remain, accurate or appropriate. Information regarding prices, travel timetables, and other factual information given in this work are correct at the time of first printing, but Cambridge University Press does not guarantee the accuracy of such information thereafter.

To
Daniel McFadden
and
in memory of
Kenneth Train, Sr.

Contents

1 Introduction *page* 1
- 1.1 Motivation 1
- 1.2 Choice Probabilities and Integration 3
- 1.3 Outline of Book 7
- 1.4 A Couple of Notes 8

Part I Behavioral Models

2 Properties of Discrete Choice Models 11
- 2.1 Overview 11
- 2.2 The Choice Set 11
- 2.3 Derivation of Choice Probabilities 14
- 2.4 Specific Models 17
- 2.5 Identification of Choice Models 19
- 2.6 Aggregation 29
- 2.7 Forecasting 32
- 2.8 Recalibration of Constants 33

3 Logit 34
- 3.1 Choice Probabilities 34
- 3.2 The Scale Parameter 40
- 3.3 Power and Limitations of Logit 42
- 3.4 Nonlinear Representative Utility 52
- 3.5 Consumer Surplus 55
- 3.6 Derivatives and Elasticities 57
- 3.7 Estimation 60
- 3.8 Goodness of Fit and Hypothesis Testing 67
- 3.9 Case Study: Forecasting for a New Transit System 71
- 3.10 Derivation of Logit Probabilities 74

4 GEV 76
- 4.1 Introduction 76
- 4.2 Nested Logit 77

4.3 Three-Level Nested Logit 86
4.4 Overlapping Nests 89
4.5 Heteroskedastic Logit 92
4.6 The GEV Family 93

5 Probit 97
5.1 Choice Probabilities 97
5.2 Identification 100
5.3 Taste Variation 106
5.4 Substitution Patterns and Failure of IIA 108
5.5 Panel Data 110
5.6 Simulation of the Choice Probabilities 114

6 Mixed Logit 134
6.1 Choice Probabilities 134
6.2 Random Coefficients 137
6.3 Error Components 139
6.4 Substitution Patterns 141
6.5 Approximation to Any Random Utility Model 141
6.6 Simulation 144
6.7 Panel Data 145
6.8 Case Study 147

7 Variations on a Theme 151
7.1 Introduction 151
7.2 Stated-Preference and Revealed-Preference Data 152
7.3 Ranked Data 156
7.4 Ordered Responses 159
7.5 Contingent Valuation 164
7.6 Mixed Models 166
7.7 Dynamic Optimization 169

Part II Estimation

8 Numerical Maximization 185
8.1 Motivation 185
8.2 Notation 185
8.3 Algorithms 187
8.4 Convergence Criterion 198
8.5 Local versus Global Maximum 199
8.6 Variance of the Estimates 200
8.7 Information Identity 202

9 Drawing from Densities 205
9.1 Introduction 205
9.2 Random Draws 205
9.3 Variance Reduction 214

10 Simulation-Assisted Estimation 237
10.1 Motivation 237
10.2 Definition of Estimators 238
10.3 The Central Limit Theorem 245
10.4 Properties of Traditional Estimators 247
10.5 Properties of Simulation-Based Estimators 250
10.6 Numerical Solution 257

11 Individual-Level Parameters 259
11.1 Introduction 259
11.2 Derivation of Conditional Distribution 262
11.3 Implications of Estimation of θ 264
11.4 Monte Carlo Illustration 267
11.5 Average Conditional Distribution 269
11.6 Case Study: Choice of Energy Supplier 270
11.7 Discussion 280

12 Bayesian Procedures 282
12.1 Introduction 282
12.2 Overview of Bayesian Concepts 284
12.3 Simulation of the Posterior Mean 291
12.4 Drawing from the Posterior 293
12.5 Posteriors for the Mean and Variance of a Normal Distribution 294
12.6 Hierarchical Bayes for Mixed Logit 299
12.7 Case Study: Choice of Energy Supplier 305
12.8 Bayesian Procedures for Probit Models 313

13 Endogeneity 315
13.1 Overview 315
13.2 The BLP Approach 318
13.3 Supply Side 328
13.4 Control Functions 334
13.5 Maximum Likelihood Approach 340
13.6 Case Study: Consumers' Choice among New Vehicles 342

14 EM Algorithms 347
14.1 Introduction 347
14.2 General Procedure 348
14.3 Examples of EM Algorithms 355
14.4 Case Study: Demand for Hydrogen Cars 365

Bibliography 371
Index 385