

Econometric Foundations

Ron C. Mittelhammer

*Washington State
University*

George G. Judge

*University of
California, Berkeley*

Douglas J. Miller

Purdue University



CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS

The Edinburgh Building, Cambridge CB2 2RU, UK <http://www.cup.cam.ac.uk>

40 West 20th Street, New York, NY 10011-4211, USA <http://www.cup.org>

10 Stamford Road, Oakleigh, Melbourne 3166, Australia

Ruiz de Alarcón 13, 28014 Madrid, Spain

© Ron C. Mittelhammer, George G. Judge, Douglas J. Miller 2000

This book 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 2000

Printed in the United States of America

Typeface Times Roman 10.5/13 pt. *System* L^AT_EX 2_ε [TB]

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

Library of Congress Cataloging in Publication Data

Mittelhammer, Ron.

Econometric foundations / Ron C. Mittelhammer, George G. Judge,
Douglas J. Miller.

p. cm.

ISBN 0-521-62394-4 hb

1. Econometrics. I. Judge, George G. II. Miller, Douglas.

III. Title.

HB139.M575 2000

330'.01'5195 – dc21

99-40040

CIP

ISBN 0 521 62394 4 hardback

Contents

Preface	xxv
I Information Processing and Recovery	1
1 The Process of Econometric Information Recovery	3
1.1 Introduction	4
1.2 The Nature of Economic Data	4
1.3 The Probability Approach to Economics	5
1.4 The Process of Searching for Quantitative Economic Knowledge	6
1.4.1 Econometric Model Components	6
1.4.2 Econometric Analysis	8
1.5 The Inverse Problem	9
1.6 A Comment	10
1.7 Notation	10
1.8 Idea Checklist – Knowledge Guides	12
2 Probability–Econometric Models	13
2.1 Parametric, Semiparametric, and Nonparametric Models	14
2.1.1 Parametric Models	14
2.1.2 Nonparametric and Semiparametric Models	15
2.2 The Classical Linear Regression Model	17
2.2.1 Establishing a Linkage between Dependent and Explanatory Variables	17
2.2.2 The Distribution of \mathbf{Y} around the Systematic Component	20
2.2.3 Inverse Problems: Estimation, Inference, and Interpretation	21
2.2.4 Some Variants of the Linear Regression Model	22
2.3 A Class of Probability Models	26
2.4 Class of Inverse Problems and Solutions	28
2.5 Concluding Comments	29

CONTENTS

2.6	Exercises	30
2.6.1	Idea Checklist – Knowledge Guides	30
2.6.2	Problems	30
2.7	References	31
II Regression Model – Estimation and Inference		
3	The Multivariate Normal Linear Regression Model: ML Estimation	35
3.1	The Linear Regression Model	35
3.1.1	The Linearity Assumption and the Inverse Problem	36
3.1.2	Linearity and Beyond: Sampling Implication of Model Assumptions	37
3.1.3	The Parametric Model	39
3.2	Maximum Likelihood Estimation of β and σ^2	39
3.2.1	The Normal Linear Regression Model	40
3.2.2	The Maximum Likelihood (ML) Criterion	40
3.2.3	Maximum Likelihood Estimators for β and σ^2	41
3.2.4	Distribution, Moments, and Bias Properties of the ML Estimator	42
3.3	Efficiency of the Bias-Adjusted ML Estimators of β and σ^2	43
3.4	Consistency, Asymptotic Normality, and Asymptotic Efficiency of ML Estimators of β and σ^2	44
3.4.1	Consistency	44
3.4.2	Asymptotic Normality	46
3.4.3	Asymptotic Efficiency	47
3.5	Summary of the Finite Sample and Asymptotic Sampling Properties of the ML Estimator	48
3.6	Estimating ε and $\text{cov}(\hat{\beta})$	49
3.6.1	An Estimator for ε	50
3.6.2	An Estimator for $\text{cov}(\hat{\beta})$	51
3.7	Concluding Remarks	51
3.8	Exercises	53
3.8.1	Idea Checklist – Knowledge Guides	53
3.8.2	Problems	53
3.8.3	Computer Problems	54
3.9	Appendix: Admissibility of ML Estimator – Introduction to Biased Estimation	54
3.9.1	Is the ML Estimator Inadmissible?	55
3.9.2	Risk Comparisons for Restricted and Unrestricted ML Estimator	57
3.10	Appendix: Proofs	58
3.11	References	60
4	The Multivariate Normal Linear Regression Model: Inference	61
4.1	Introduction	61
4.2	Some Basic Sampling Distributions of Functions of $\hat{\beta}$ and S^2	62

CONTENTS

4.3	Hypothesis Testing	64
4.3.1	Test Statistics	65
4.3.2	Generalized Likelihood Ratio (GLR) Test	65
4.3.3	Properties of the GLR Test	66
4.3.4	Testing Hypotheses Relating to Linear Restrictions on β	67
4.3.5	Testing Linear Inequality Hypotheses	72
4.3.6	Bonferroni Joint Tests of Inequality and Equality Hypotheses about β	73
4.3.7	Testing Hypotheses about σ^2	74
4.4	Confidence Interval and Region Estimation	76
4.4.1	Meaning of “Confidence” Interval or Region	76
4.4.2	Rationale for the Use of Confidence Regions	77
4.4.3	Confidence Intervals, Regions, and Bounds for β	77
4.4.4	Confidence Bounds on σ^2	79
4.5	Pretest Estimators Based on Hypothesis Tests: Introduction	80
4.5.1	The Pretest Estimator	80
4.5.2	The Pretest Estimator Risk Function	81
4.6	Concluding Remarks	82
4.7	Exercises	84
4.7.1	Idea Checklist – Knowledge Guides	84
4.7.2	Problems	84
4.7.3	Computer Exercises	85
4.8	References	85
5 The Linear Semiparametric Regression Model: Least-Squares Estimation		86
5.1	The Semiparametric Linear Regression Model	86
5.2	The Problem of Estimating β	88
5.2.1	The Squared-Error Metric and the Least-Squares Principle	89
5.2.2	The Least-Squares Estimator	90
5.3	Statistical Properties of the LS Estimator	91
5.3.1	Mean, Covariance, and Unbiasedness	92
5.3.2	GAUSS Markov Theorem: $\hat{\beta}$ Is BLUE, Not Necessarily MVUE	93
5.3.3	$\hat{\beta}$ Is a Minimax Estimator of β under Quadratic Risk	94
5.3.4	Consistency	95
5.3.5	Asymptotic Normality	96
5.4	Estimating ε , σ^2 , and $\text{cov}(\hat{\beta})$	97
5.4.1	An Estimator for ε	98
5.4.2	An Estimator for σ^2 and $\text{cov}(\hat{\beta})$	99
5.5	Critique	100
5.6	Exercises	101
5.6.1	Idea Checklist – Knowledge Guides	101
5.6.2	Problems	101
5.6.3	Computer Exercises	102

CONTENTS

5.7	References	103
5.8	Appendix: Proofs	103
6	The Linear Semiparametric Regression Model: Inference	105
6.1	Asymptotics: Why, What Kind, and How Useful?	107
6.1.1	Why?	107
6.1.2	What Kind?	108
6.1.3	How Useful?	108
6.2	Hypothesis Testing: Linear Equality Restrictions on β	109
6.2.1	Wald (W) Tests	110
6.2.2	Lagrange Multiplier (LM) Tests	114
6.2.3	The W and LM Tests under Normality	118
6.2.4	W and LM Tests: Interrelationships and Extensions	120
6.3	Confidence Region Estimation	120
6.3.1	Wald-Based Confidence Regions	121
6.3.2	LM-Based Confidence Regions	122
6.4	Testing Linear Inequality Hypotheses and Generating Confidence Bounds on β	123
6.4.1	Linear Inequality Hypotheses	124
6.4.2	Confidence Bounds for $c\beta$	125
6.5	Critique	126
6.6	Comprehensive Computer Application	127
6.7	Exercises	128
6.7.1	Idea Checklist – Knowledge Guides	128
6.7.2	Problems	128
6.7.3	Computer Exercises	129
6.8	References	129
III	Extremum Estimators and Nonlinear and Nonnormal Regression Models	131
7	Extremum Estimation and Inference	133
7.1	Introduction	133
7.2	ML and LS Estimators Expressed in Extremum Estimator Form	136
7.3	Asymptotic Properties of Extremum Estimators	136
7.3.1	Consistency of Extremum Estimators	136
7.3.2	Asymptotic Normality of Extremum Estimators	138
7.4	Asymptotic Properties of Maximum Likelihood Estimators in an Extremum Estimator Context	139
7.4.1	Consistency of the ML-Extremum Estimator	140
7.4.2	Asymptotic Normality of the ML-Extremum Estimator	141
7.5	Asymptotic Properties of the Least-Squares Estimator in an Extremum Estimator Context	142
7.5.1	Consistency of the LS-Extremum Estimator	142
7.5.2	Asymptotic Normality of the LS-Extremum Estimator	143

CONTENTS

7.6	Inference Based on Extremum Estimation	144
	7.6.1 Lagrange Multiplier Tests and Confidence Regions	144
	7.6.2 Wald Tests and Confidence Regions	148
	7.6.3 Pseudo-Likelihood Ratio Tests and Confidence Regions	149
	7.6.4 Testing Linear Inequalities and Confidence Bounds	152
7.7	Critique	153
7.8	Exercises	154
	7.8.1 Idea Checklist – Knowledge Guides	154
	7.8.2 Problems	154
	7.8.3 Computer Problems	155
7.9	References	155
8	The Nonlinear Semiparametric Regression Model: Estimation and Inference	157
8.1	The Nonlinear Regression Model	157
	8.1.1 Assumed Probability Model Characteristics: Discussion	159
	8.1.2 The Inverse Problem	161
8.2	The Problem of Estimating β	162
	8.2.1 The Nonlinear Least-Squares Estimator	162
	8.2.2 Parameter Identification Relative to a Probability Model	163
	8.2.3 Parameter Identification Relative to the Least-Squares Criterion and Given Data	165
8.3	Sampling Properties of the NLS Estimator	167
8.4	The Problem of Estimating ε , σ^2 , and $\text{cov}(\hat{\beta})$	169
	8.4.1 An Estimator for ε	169
	8.4.2 An Estimator for σ^2 and $\text{cov}(\hat{\beta})$	170
8.5	Wald Statistics: Tests and Confidence Regions	171
	8.5.1 Hypotheses Relating to Differentiable Functions of β	171
	8.5.2 Wald (W) Tests	172
	8.5.3 Wald Statistic Distribution under H_a	173
	8.5.4 Test Application	174
	8.5.5 Confidence Region Estimation	175
8.6	LM Statistics: Tests and Confidence Regions	176
	8.6.1 Lagrange Multiplier Distribution	176
	8.6.2 Classical Form of LM Test	178
	8.6.3 Score Form of LM Test	178
	8.6.4 LM Statistic Distribution under H_a	179
	8.6.5 Test Application	179
	8.6.6 Confidence Region Estimation	180
8.7	Pseudo-Likelihood Ratio Statistic: Tests and Confidence Regions	180
	8.7.1 The Pseudo-Likelihood Ratio Statistic	180
	8.7.2 Distribution of PLR under H_0	181
	8.7.3 Test and Confidence Region Application and Properties	182

CONTENTS

8.8	Nonlinear Inequality Hypotheses and Confidence Bounds	183
8.8.1	Testing Nonlinear Inequalities: Z-Statistics	183
8.8.2	Confidence Bounds	185
8.9	Asymptotic Properties of the NLS-Extremum Estimator	186
8.9.1	Consistency	186
8.9.2	Asymptotic Normality	187
8.9.3	Asymptotic Linearity	190
8.9.4	Best Asymptotically Linear Consistent Estimator	191
8.10	Concluding Comments	192
8.11	Exercises	192
8.11.1	Idea Checklist – Knowledge Guides	192
8.11.2	Problems	193
8.11.3	Computer Exercises	193
8.12	References	194
8.13	Appendixes	195
8.13.1	Appendix A: Computation of NLS Estimates (with Ronald Schoenberg)	195
8.13.2	Newton–Raphson Method	195
8.13.3	Gauss–Newton Method	197
8.13.4	Quasi–Newton Methods	198
8.13.5	Computational Issues	199
8.13.6	Appendix B: Proofs	202
9	Nonlinear and Nonnormal Parametric Regression Models	204
9.1	The Normal Nonlinear Regression Model	204
9.1.1	Maximum Likelihood–Extremum Estimation	205
9.1.2	Maximum Likelihood–Extremum Inference	207
9.2	Nonnormality	209
9.2.1	A Nonnormal Dependent Variable, Normal Noise Component Case: The Box–Cox Transformation	209
9.2.2	A Nonnormal Noise Component Case	212
9.3	General Considerations in Applying the ML–Extremum Criterion	215
9.3.1	Point Estimation	216
9.3.2	Testing and Confidence-Region Estimation	217
9.4	A Final Remark	220
9.5	Exercises	220
9.5.1	Idea Checklist – Knowledge Guides	220
9.5.2	Problems	220
9.5.3	Computer Exercises	221
9.6	References	221

IV	Avoiding the Parametric Likelihood	223
10	Stochastic Regressors and Moment-Based Estimation	225
10.1	Introduction	225
10.2	Linear Model Assumptions, Estimation, and Inference Revisited	227
10.3	LS and ML Estimator Properties	229
10.3.1	LS Estimator Properties: Finite Samples	229
10.3.2	LS Estimator Properties: Asymptotics	230
10.3.3	ML Estimation of β and σ^2 under Conditional Normality	232
10.4	Hypothesis Testing and Confidence-Region Estimation	233
10.4.1	Semiparametric Case	233
10.4.2	Parametric Case	233
10.5	Summary: Statistical Implications of Stochastic \mathbf{X}	235
10.6	Method of Moments Concept	235
10.6.1	Asymptotic Properties	236
10.6.2	A Linear Model Formulation	238
10.6.3	Extensions to Nonlinear Models	239
10.7	Concluding Comments	241
10.8	Exercises	242
10.8.1	Idea Checklist – Knowledge Guides	242
10.8.2	Problems	242
10.8.3	Computer Exercises	242
10.9	References	243
10.10	Appendix: Proofs	244
11	Quasi-Maximum Likelihood and Estimating Equations	245
11.1	Quasi-Maximum Likelihood Estimation and Inference	247
11.2	QML–E Estimation and Inference	248
11.2.1	Consistency and Asymptotic Normality	248
11.2.2	True PDF Approximation Property and Asymptotic Normality of Inconsistent QML–E Estimators	249
11.2.3	Consistent Estimation of the Asymptotic Covariance Matrix	251
11.2.4	Necessary and Sufficient Conditions for Consistency and Asymptotic Normality	251
11.2.5	QML Inference	254
11.2.6	QML Generalizations	255
11.2.7	QML Summary Comments	256
11.3	Estimating Equations: LS, ML, QML–E, and Extremum Estimators	256
11.3.1	Linear and Nonlinear Estimating Functions	259
11.3.2	Optimal Unbiased Estimating Functions: Finite Sample Optimality of ML	261

CONTENTS

11.3.3	Consistency of the EE Estimator	265
11.3.4	Asymptotic Normality and Efficiency of EE Estimators	267
11.3.5	Inference in the Context of EE Estimation	268
11.4	Unifying-Linking OptEF and QML: QML-EE Estimation and Inference	270
11.4.1	The Best Linear-Unbiased QML-EE	272
11.4.2	General QML-EE Estimation and Inference	274
11.5	Final Remarks	275
11.6	Exercises	275
11.6.1	Idea Checklist – Knowledge Guides	275
11.6.2	Problems	276
11.6.3	Computer Problems	277
11.7	References	277
11.8	Appendix: Proofs	279
12	Empirical Likelihood Estimation and Inference	281
12.1	Empirical Likelihood: iid Case	282
12.1.1	The EL Concept	283
12.1.2	Nonparametric Maximum Likelihood Estimate of a Population Distribution	284
12.1.3	Empirical Likelihood Function for θ	286
12.2	Maximum Empirical Likelihood Estimation: iid Case	292
12.2.1	Maximum Empirical Likelihood Estimator	292
12.2.2	MEL Efficiency Property	293
12.2.3	MEL Estimation of a Population Mean	294
12.2.4	MEL Estimation Based on Two Moments	297
12.3	Hypothesis Tests and Confidence Regions: iid Case	298
12.3.1	Empirical Likelihood Ratio Tests and Confidence Regions for $\mathbf{c}(\theta)$	299
12.3.2	Wald Tests and Confidence Regions for $\mathbf{c}(\theta)$	300
12.3.3	Lagrange Multiplier Tests and Confidence Regions for $\mathbf{c}(\theta)$	300
12.3.4	Z-Tests of Inequality Hypotheses for the Value of $\mathbf{c}(\theta)$	301
12.3.5	Testing the Validity of Moment Equations	301
12.3.6	MEL Testing and Confidence Intervals for Population Mean	302
12.3.7	Illustrative MEL Confidence Interval Example	303
12.4	MEL in the Linear Regression Model with Stochastic \mathbf{X}	304
12.4.1	MEL Regression Estimation for Stochastic \mathbf{X}	304
12.4.2	EL-Based Testing and Confidence Regions When \mathbf{X} Is Stochastic	306
12.4.3	Extensions to the Nonlinear Regression Model for Stochastic \mathbf{X}	307
12.5	MEL in the Linear Regression Model with Nonstochastic \mathbf{X} : Extensions to the Non-iid Case	307
12.6	Concluding Comments	309

CONTENTS

12.7	Exercises	310
12.7.1	Idea Checklist – Knowledge Guides	310
12.7.2	Problems	311
12.7.3	Computer Exercises	311
12.8	References	312
13	Information Theoretic-Entropy Approaches to Estimation and Inference	313
13.1	Solutions to Systems of Estimating Equations and Kullback–Leibler Information	313
13.1.1	Kullback–Leibler Information Criterion (KLIC)	316
13.1.2	Relationship Between the MEL Objective and KL Information	317
13.1.3	Relationship Between the Maximum Entropy (ME) Objective and KL Information	319
13.2	The General MEEL Alternative Empirical Likelihood Formulation	321
13.2.1	The MEEL Estimator and Likelihood	321
13.2.2	MEEL Asymptotics	322
13.2.3	MEEL Inference	323
13.2.4	Contrasting the Use of Estimating Functions in EE and MEEL Contexts	326
13.3	A Cross-Entropy Formalism and Solution	326
13.4	α -Entropy: Unifying the MEL, MEEL, and CEEL Estimation Objectives	328
13.5	Application of the Maximum Entropy Principle to the Regression Model	329
13.5.1	Stochastic \mathbf{X} in the Linear Model	329
13.5.2	Fixed \mathbf{x} in the Linear Model	331
13.5.3	Extensions to Nonlinear Regression Models	331
13.5.4	Inference in Regression Models	331
13.6	Concluding Remarks – Which Criterion?	332
13.7	Exercises	334
13.7.1	Idea Checklist – Knowledge Guides	334
13.7.2	Problems	334
13.7.3	Computer Problems	334
13.8	References	335
13.9	Supplemental References	335
V	Generalized Regression Models	337
14	Regression Models with a Known General Noise Covariance Matrix	339
14.1	Applying LS–ML to the Linear Model and Untransformed Data: Ignoring that $\Psi \neq \mathbf{I}$	340
14.1.1	Point Estimation	341
14.1.2	Testing and Confidence-Region Estimation	342

CONTENTS

14.2	GLS–ML–Extremum Analysis of the Linear Model: Incorporating a Known Ψ	344
14.2.1	ML–Extremum Estimators for β and σ^2	344
14.2.2	LS–Extremum Estimators for β and σ^2 : Transformed Linear Model	345
14.2.3	Asymptotic Properties	347
14.2.4	Hypothesis Testing and Confidence Regions	349
14.3	GLS–ML–Extremum Analysis of the Nonlinear Model: Incorporating a Known Ψ	350
14.3.1	Estimator Properties	350
14.3.2	Hypothesis Testing and Confidence Regions	351
14.3.3	Applying NLS to Untransformed Data	352
14.4	Parametric Specifications of Noise Covariance Matrices	353
14.4.1	Estimation and Inference with AR(1) Noise	354
14.4.2	Heteroscedasticity – Structural Specification Known, Parameters Unknown	355
14.4.3	General Considerations	356
14.5	Sets of Regression Equations	356
14.5.1	Sets of Linear Equations	357
14.5.2	Sets of Nonlinear Equations	359
14.6	The Estimating Equations – Quasi-Maximum Likelihood View: A Unified Approach	361
14.6.1	EE Estimation of β and σ^2 in the Linear Model When Ψ Is Known	361
14.6.2	EE Estimation of β and σ^2 in the Nonlinear Model When Ψ Is Known	363
14.6.3	EE Estimation of β and σ^2 in the Linear or Nonlinear Normal Parametric Regression Model When Ψ Is Known	364
14.6.4	QML–EE Estimation of β and σ^2 in the Linear or Nonlinear Regression Model	365
14.6.5	EE Estimation of β and σ^2 in Systems of Regression Equations	366
14.7	Some Comments	367
14.8	Exercises	368
14.8.1	Idea Checklist – Knowledge Guides	368
14.8.2	Problems	368
14.8.3	Computer Exercises	369
14.9	References	369
15	Regression Models with an Unknown General Noise Covariance Matrix	370
15.1	Linear Regression Models with Unknown Noise Covariance	371
15.1.1	Single-Equation Semiparametric Linear Regression Model	372
15.1.2	Single-Equation Parametric Linear Regression Model – The ML Approach	377

CONTENTS

15.2	System of Linear Regression Equations	378
15.2.1	Estimation: Semiparametric Case	379
15.2.2	Testing and Confidence Regions: Semiparametric Case	380
15.2.3	ML Approach: Parametric Case	381
15.3	Nonlinear Regression Models with Unknown Noise Covariance	383
15.3.1	Single-Equation Semiparametric Nonlinear Regression Model	383
15.3.2	Estimation	383
15.3.3	Testing and Confidence Regions	384
15.3.4	Single-Equation Parametric Nonlinear Regression Model – ML Approach	385
15.3.5	Sets of Nonlinear Regression Equations	386
15.4	Robust Solution Methods: OLS and Robust Covariance Matrix Estimation	387
15.4.1	Heteroscedasticity	389
15.4.2	Heteroscedasticity and Autocorrelation	391
15.5	The Estimating Equations – Quasi-Maximum Likelihood View: A Unified Approach	393
15.5.1	A Unified EE Characterization of Inverse Problem Solutions	393
15.5.2	QML–EE Estimation	395
15.5.3	EE Extensions	396
15.5.4	MEL and MEEL Applications of EE	397
15.6	Some Comments	398
15.7	Exercises	400
15.7.1	Idea Checklist – Knowledge Guides	400
15.7.2	Problems	400
15.7.3	Computer Exercises	400
15.8	References	401
 VI	 Simultaneous Equation Probability Models and General Moment-Based Estimation and Inference	 403
16	Generalized Moment-Based Estimation and Inference	405
16.1	Parameter Estimation in Just-determined and Overdetermined Models with iid Observations: Back to the Future	406
16.1.1	OptEF Approach	409
16.1.2	Empirical Likelihood Approaches	410
16.1.3	Summary and Foreword	411
16.2	GMM Solutions for Unbiased Estimating Equations in the Overdetermined Case	412
16.2.1	GMM Concept	412
16.2.2	GMM Linear Model Estimation	413
16.2.3	GMM Estimators – General Properties	420
16.3	IV Solutions in the Just-determined Case When $E[\mathbf{X}'\varepsilon] \neq \mathbf{0}$	423
16.3.1	Traditional Instrumental Variable Estimator in the Linear Model	424

CONTENTS

16.3.2	GLS as an IV Estimator	426
16.3.3	Extensions: Nonlinear IV Formulations and Non-iid Sampling	427
16.3.4	Hypothesis Testing and Confidence Regions	428
16.3.5	Summary: IV Approach to Estimation and Inference	429
16.4	Solutions in the Overdetermined Case When $E[\mathbf{X}'\varepsilon] \neq \mathbf{0}$	429
16.4.1	Unbiased Estimating Equations Basis for Inverse Problem Solutions When $E[\mathbf{X}'\varepsilon] \neq \mathbf{0}$	430
16.4.2	GMM Approach	431
16.4.3	MEL and MEEL Approaches	432
16.4.4	OptEF–Quasi–ML Approach: Asymptotic Unification of Inverse Solution Methods	436
16.4.5	Asymptotic Sampling Properties and Inference	437
16.4.6	Testing Moment Equation Validity	438
16.4.7	Relationship Between Estimator Efficiency and the Number and Type of Estimating Equations	439
16.5	Concluding Comments	441
16.6	Exercises	443
16.4.1	Idea Checklist – Knowledge Guides	443
16.4.2	Problems	443
16.4.3	Computer Exercises	444
16.7	References	444
17	Simultaneous Equations Econometric Models: Estimation and Inference	446
17.1	Linear Simultaneous Equations Models	447
17.1.1	An Equivalent Vectorized System of Equations	450
17.1.2	The Reduced-Form Regression Model	451
17.1.3	Estimating the Reduced-Form Coefficients	453
17.1.4	The Identification Problem	455
17.2	Least-Squares and GMM Estimation: The Semiparametric Case	458
17.2.1	Estimators of Parameters for a Just-Identified Structural Equation	459
17.2.2	GMM Estimator of Parameters for an Overidentified Structural Equation	460
17.2.3	Estimation of a Complete System of Equations	462
17.3	Maximum Likelihood Estimation in the Linear Model: The Parametric Case	465
17.3.1	Full-Information Maximum Likelihood (FIML)	465
17.3.2	Limited Information Maximum Likelihood (LIML)	467
17.4	Nonlinear Simultaneous Equations	468
17.4.1	Single-Equation Estimation: Semiparametric Case	469
17.4.2	Complete System Estimation: Semiparametric Case	472

CONTENTS

17.4.3	Nonlinear Maximum Likelihood Estimation: Parametric Case	473
17.4.4	Identification in Nonlinear Systems of Equations	474
17.5	Information Theoretic Procedures	475
17.5.1	Minimum KLIC Approach: MEEL and MEL	476
17.5.2	MEEL Estimation and Inference	479
17.5.3	MEL Estimation	484
17.5.4	The KLIC–MEEL–MEL Procedures: Critique	485
17.6	OptEF Estimation and Inference	486
17.6.1	The OptEF Estimator in Simultaneous Equations Models	486
17.6.2	OptEF Asymptotic Unification of Inverse Problem Solutions	487
17.7	Concluding Comments	488
17.8	Exercises	489
17.8.1	Idea Checklist – Knowledge Guides	489
17.8.2	Problems	489
17.8.3	Computer Problems	490
17.9	References	491
17.10	Appendix: Historical Perspective	492
VII Model Discovery		495
18 Model Discovery: The Problem of Variable Selection and Conditioning		497
18.1	Introduction	497
18.1.1	Experiments in Nonexperimental Model Building	498
18.1.2	The Chapter Format	499
18.2	Variable Selection Problem in a Loss or MSE Context	500
18.2.1	Bias-Variance Trade-Off in Incorrect Variable Selection	501
18.2.2	Statistical Implications under an MSE Matrix Measure	503
18.2.3	Statistical Implications under a Quadratic Risk Function Measure	505
18.2.4	Extensions to Nonlinear Models, General Covariance Structures, and Systems of Equations	506
18.3	Fisherian Testing and Model Choice: Critique	507
18.4	Estimated Risk Criteria and Model Choice: Mallows C_p Criterion	509
18.5	An Information Theoretic Model Selection Criterion: Akaike Information	511
18.5.1	Basic Rationale for the AIC and Variants	511
18.5.2	A Comment	513
18.6	Shrinkage as a Basis for Dealing with Variable Uncertainty	514
18.6.1	Stein-Like Shrinkage Estimators	514
18.6.2	Multiple Shrinkage Estimator	515
18.7	The Problem of an Ill-Conditioned Explanatory Variable Matrix	517
18.7.1	Penalized Estimation	518
18.7.2	Finite Sample Performance	520

CONTENTS

18.8	Final Comments and Critique	521
18.9	Exercises	524
	18.9.1 Idea Checklist – Knowledge Guides	524
	18.9.2 Problems	524
	18.9.3 Computer Problems	524
18.10	References	525
19	Model Discovery: The Problem of Noise Covariance Matrix Specification	528
19.1	Introduction	528
19.2	Specific Parametric Specifications of the Noise Covariance Matrix: Estimation and Inference	530
	19.2.1 AR(1) Noise	530
	19.2.2 Heteroscedastic Noise: $\sigma_i^2 = (\mathbf{z}_i \cdot \boldsymbol{\alpha})^2$	533
19.3	Tests for Heteroscedasticity: Rationale and Application	535
	19.3.1 Types of Heteroscedasticity Tests	536
	19.3.2 Motivation for Tests Based on $\hat{\varepsilon}_i^2$	536
	19.3.3 Motivation for the Test Based on $\ln(\hat{\varepsilon}_i^2)$	540
	19.3.4 Motivation for Test Based on $ \hat{\varepsilon}_i $	541
	19.3.5 More Tests	542
19.4	Tests for Autocorrelation: Rationale and Application	543
	19.4.1 Autocorrelation Processes	543
	19.4.2 Estimation	546
	19.4.3 Autocorrelation Tests	547
19.5	Pretest Estimators Defined by Heteroscedasticity or Autocorrelation Testing	551
	19.5.1 Two-Equation Linear System with Potential Heteroscedasticity	552
	19.5.2 A Heteroscedasticity Pretest Estimator	553
	19.5.3 An Autocorrelation Pretest Estimator	555
19.6	Concluding Comments	556
19.7	Exercises	557
	19.7.1 Idea Checklist – Knowledge Guides	557
	19.7.2 Problems	557
	19.7.3 Computer Problems	557
19.8	References	558
VIII	Special Econometric Topics	561
20	Qualitative–Censored Response Models	563
20.1	Binary–Discrete Choice Response Models	565
	20.1.1 A Linear Probability Model	566
	20.1.2 A Reformulated Binary Response Model	567

CONTENTS

20.2	Maximum Likelihood Estimation and Inference for the Discrete Choice Model	570
20.2.1	Logit Model	571
20.2.2	Probit Model	573
20.2.3	Probit or Logit?	575
20.3	Multinomial Discrete Choice	575
20.3.1	Multinomial Logit	576
20.3.2	Information Theoretic Estimation of the Multinomial Decision Model	581
20.3.3	Ordered Multinomial Choice	584
20.4	Censored Response Data	585
20.4.1	Censoring Versus Truncation	585
20.4.2	Tobit Model	587
20.4.3	Extensions of the Tobit Formulation	591
20.5	Concluding Remarks	592
20.6	Exercises	594
20.6.1	Idea Checklist – Knowledge Guides	594
20.6.2	Problems	594
20.6.3	Computer Problems	595
20.7	References	595
20.8	Appendix	597
21	Introduction to Nonparametric Density and Regression Analysis	599
21.1	Density Estimation via Kernels	601
21.1.1	Kernels	602
21.1.2	IMSE, AIMSE, and Kernel Choice	604
21.1.3	Bandwidth Choice	606
21.1.4	Other Properties and Issues	608
21.1.5	Multivariate Extensions	609
21.2	Kernel Regression Estimators	611
21.2.1	Nonparametric Regression Model Specification	612
21.2.2	Nadaraya–Watson Kernel Regression: Alias Zero-Order Local Polynomial Regression	613
21.3	Local Polynomial Regression	622
21.4	Prequel of Fundamental Concepts: Local Weighted Averages, Local Linear Regressions, and Histograms	624
21.4.1	Nonparametric Regression under Repeated Sampling: Simple Sample Means	625
21.4.2	Local Sample Means: Using Data Neighborhoods	626
21.4.3	Local Weighted Averages within Data Neighborhoods: Local Least-Squares Regressions	628
21.4.4	Local Polynomial Regressions	629
21.4.5	Histograms: Precursor to Kernel Density Estimation	630
21.5	Concluding Comments	636

CONTENTS

21.6	Exercises	637
21.6.1	Idea Checklist – Knowledge Guides	637
21.6.2	Problems	637
21.6.3	Computer Problems	638
21.7	References	638
21.8	Appendix: Derivation of Equation (21.4.31)	640
IX	Bayesian Estimation and Inference	643
22	Bayesian Estimation: General Principles with a Regression Focus	645
22.1	Introduction	645
22.1.1	Bayes Theorem	646
22.1.2	A Format for Bayesian Reasoning	647
22.2	Bayesian Probability Models and Posterior Distributions	649
22.2.1	Prior Distributions	650
22.2.2	Posterior Probability Distributions	651
22.3	The Bayesian Linear–Regression-Based Probability Model	652
22.3.1	Bayesian Regression Analysis under Normality and Uninformative Priors	653
22.3.2	Uninformative Priors and Proper Marginal Posteriors	654
22.3.3	Posterior Distribution under an Uninformative Prior	655
22.3.4	Marginal Posteriors	656
22.4	Bayesian Regression Analysis under Normality and Conjugate Informative Priors	658
22.4.1	The Joint Posterior Distribution under the Conjugate Prior	660
22.4.2	The Marginal Posteriors	660
22.5	Bayesian Point Estimates	661
22.5.1	Minimum Expected Risk	663
22.5.2	Admissibility	663
22.5.3	Consistency and Asymptotic Normality	663
22.5.4	Comments on Bayes Estimator Properties	664
22.6	On the Use of Conjugate Priors and Coincidence of Classical and Bayesian Estimates	665
22.7	Concluding Remarks	666
22.8	Exercises	667
22.8.1	Idea Checklist – Knowledge Guides	667
22.8.2	Problems	667
22.8.3	Computer Problems	668
22.9	References	669
22.10	Appendix: Bayesian Asymptotics	669
22.10.1	Bayesian Asymptotics: Specific Cases	670
22.10.2	Bayesian Asymptotics: General Considerations	672
22.10.3	Appendix References	674

23	Alternative Bayes Formulations for the Regression Model	675
23.1	<i>g</i> -Priors	676
23.1.1	A Family of <i>g</i> -priors and Associated Posteriors	676
23.1.2	Rationalizing and Specifying <i>g</i> -priors	678
23.2	An Empirical Bayes Estimator	679
23.3	General Bayesian Regression Analysis with Nonconjugate Informative–Uninformative Prior	682
23.3.1	Normal Noise Component	682
23.3.2	Nonnormal Noise Component	687
23.3.3	Summary Comments	688
23.4	Bayesian Method of Moments	688
23.4.1	Continuous Entropy Formulation	691
23.4.2	A Posterior Density Function	692
23.5	Concluding Remarks	693
23.6	Exercises	694
23.6.1	Idea Checklist – Knowledge Guides	694
23.6.2	Problems	694
23.6.3	Computer Problems	695
23.7	References	696
24	Bayesian Inference	698
24.1	Credible Regions	699
24.1.1	General Principles	699
24.1.2	Highest Posterior Density Credible Regions	700
24.1.3	Credible Regions in the Regression Model	701
24.2	Hypothesis Evaluation and Decision	702
24.2.1	General Principles	703
24.3	General Hypothesis Testing and Credible Sets in the Regression Model	704
24.3.1	Evaluating Composite Hypotheses about β and Bayes Factors	704
24.3.2	Evaluating Simple Hypotheses about β	706
24.4	A Comment	708
24.5	Exercises	708
24.5.1	Idea Checklist – Knowledge Guides	708
24.5.2	Problems	708
24.5.3	Computer Problems	709
24.6	References	709
X	Epilogue	711
	Appendix: Introduction to Computer Simulation and Resampling Methods	713
A.1	Pseudorandom Number Generation	713
A.1.1	Generating U(0,1) Pseudorandom Numbers	714

CONTENTS

A.1.2	Generating Continuous Nonuniform Pseudorandom Numbers	715
A.1.3	Generating Discrete Pseudorandom Numbers	717
A.1.4	Evaluating the Performance of Pseudorandom Number Generators	718
A.2	Monte Carlo Simulation	719
A.2.1	Background and Conceptual Motivation	720
A.2.2	Key Assumptions	721
A.2.3	Basic Properties of Monte Carlo Simulation Estimators	722
A.3	Bootstrap Resampling	724
A.3.1	Basic Properties of the Bootstrap	725
A.3.2	Bootstrap Simulation Procedures for Regression Models	728
A.4	Numerical Tools for Evaluating Posterior Distributions	730
A.4.1	The Gibbs Sampling Algorithm	731
A.4.2	The Metropolis–Hastings Algorithm	733
A.5	Concluding Remarks	734
A.6	Exercises	735
A.6.1	Idea Checklist	735
A.6.2	Problems	735
A.6.3	Computer Problems	735
A.7	References	736
	Author Index	739
	Subject Index	743

CD-ROM

	Foundations Review
E1	Probability Theory
E2	Classical Estimation and Inference Principles
E3	Ill-Posed Inverse Problems
	ELECTRONIC MANUALS
	Matrix Review Manual
	Econometric Examples Manual
	GAUSS Tutorial and Technical Archive
	GAUSS Software

The Process of Econometric Information Recovery

MOST econometric problems begin with several fundamental questions. One basic question is, How does one develop a plausible basis for reasoning in situations involving partial–incomplete information? Another basic question relates to how one goes about learning from a sample of data.

For the theoretical econometrician, questions tend to be of a nonempirical and hypothetical what-if type: What if a sample of data is described by a particular imagined sampling process? This leads to the question of how one characterizes the sampling process in terms of a probability model that properly identifies the stochastic characteristics of the sampling process as well as the data-restricting constraints, the knowns and unknowns of the problem, and the observable and unobservable components in the model. Given that the data-sampling process can be described by a probability model that expresses the state of knowledge about possible real-world outcomes, another question then arises relating to how one devises effective estimation and hypothesis-testing procedures that will allow the recovery of estimates of the unknowns and provide a basis for making inferences. The theoretical econometrician may, by a process of interpretation, ultimately associate the conceptualized sampling process with a set of *observable economic data*. At this point in the theoretical econometrician's investigation the probability model is interpretable as an econometric model having economic meaning with both real-world economic and statistical implications.

For the applied econometrician, the econometric problem begins with a real-world economic question, perhaps involving the implications of scarcity and choice or perhaps the allocative or distributive impacts, resulting from an action or decision. The next step involves restating the real-world question within a theoretical–conceptual economic model framework in which real-world components are identified to facilitate drawing logic-based conclusions about the question. This step exposes structure and defines the explicit economic model to be used in the empirical analysis of the economic question.

1.1. Introduction

Given a theoretical–conceptual economic playing field, a basis is needed for connecting the real-world data outcomes with their counterparts in the economic model. By visualizing some imagined sampling process by which the outcomes may have evolved and then characterizing this sampling process by a probability model, an econometric model is born. This model then acts as a vehicle for expressing knowledge about real-world outcomes and identifies knowns, unknowns, and observed and unobservable model components. If the applied econometrician is fortunate, the resulting econometric model may be consistent with a probability model that already exists in the literature and for which a well-defined basis for estimation and inference is already available. In this case the applied econometrician will use established statistical procedures to address research questions. On the other hand, the econometric model may not be consistent with a commonly specified and evaluated data-sampling process. Consequently, the applied econometrician must assume the role of the theoretical econometrician in first developing effective estimation and hypothesis-testing procedures and then carrying through the estimation and inference stages needed to answer research questions.

As one reads through this chapter and the chapters ahead, it will at times be necessary to assume the roles of both a theoretical and an applied econometrician to derive maximum benefit from the econometric venture. One goal of the exercises in each chapter is to lead and inspire the reader in this direction. Before going on to consider the question of how to specify a probability–econometric model to provide a basis for learning from a sample of observations, we focus some attention on the real-world component referred to as economic data.

1.2. The Nature of Economic Data

Why do we have books on econometrics? Why not just have books devoted to statistics for economists? What is it that makes economics unique relative to other fields of science?

One thing that tends to make economics and econometrics unique is the nature of economic data and the special characteristics of the sampling processes by which economic data are obtained. In providing an answer to the opening questions of this section, William Barnett, in private correspondence, points us to a classic article by Schumpeter in *Econometrica*, Vol. 1, No. 1, 1933, “The Common Sense of Econometrics.” In this article Schumpeter wrote, “Econometrics is the most quantitative . . . of all sciences, physics not excluded. . . . Every economist is an econometrician whether he wants to be or not.” His rationale is that the economy produces and inherently depends upon numbers. Indeed, the very act of transacting in markets depends explicitly upon the numerical values of such variables as quantities and prices. But in physics, for example, the physical world can and will operate without dependence upon numerical measurement of variables. Scientists have to construct devices to measure temperature, pressure, speed, weight, and the like because nature does not “quote” these numbers in

markets or even identify a need to know these numbers. So in this sense, as Schumpeter observed, economics is inherently more quantitative than any other scientific field.

Economies and markets carry out experiments and produce numerical data through the very nature of their operation. However, they do so in a manner that is not usually in accordance with statistically designed experiments. Although physical scientists can be viewed, as they are in Schumpeter's article, as being disadvantaged by the need to measure data, in comparison with economists, who need only record the numerical data that the economy produces, physical scientists have the advantage of being able to run controlled experiments and to generate their data in a manner consistent with established and understood experimental designs. Because the economy is not a statistically designed experiment, economists must in many cases utilize ill-conditioned data. This is a principal reason why econometrics requires special tools for probability model formulation, estimation, and inference and why econometrics is characterized as an experiment in nonexperimental model building. The uncertain nature of economic outcomes goes a long way to explaining why almost everyone, at one time or another, has felt comfortable assuming the role of an economist on certain issues (How often have you heard the phrase "I am not an economist, but. . . ?") and why all economists are in one way or another econometricians.

1.3. The Probability Approach to Economics

The probability theory that you encountered in your courses in theoretical statistics (and that is reviewed in an electronic document on the CD that accompanies this text) has important implications for how one should organize, incorporate, and utilize data and prior information in quantitative economic analyses. In economic problems characterized by incomplete knowledge and uncertainty, this theory, through a process of abstraction and interpretation by analysts, defines a reasoning process for expressing our knowledge about real-world outcomes, for recovering information from data, and for assessing its validity. The calculus of reasoning defined by probability theory facilitates learning and problem resolution and defines a logical basis for evaluating decisions and making choices.

Like Mozart's *Don Giovanni*, conceptual tools such as random variables and stochastic processes, as well as models of economic systems, have been invented or created by theoreticians and empirical analysts rather than discovered. The participants or players in a postulated economic system are presumed to define economic processes that result in measurable outcomes and, by a process of interpretation on the part of the econometrician, these outcomes are viewed in probabilistic form. In most econometric problems, at least a portion of the information available for analysis will be in the form of a sample of data that has been generated as an outcome of some real-world economic process. In addition, the analyst generally has some prior knowledge about the relevant economic processes and institutions that may have conditioned the sample outcomes. If one views the outcomes of the economic system as having come from an imagined sampling process, concepts such as random variables and probability distributions can be used as conceptual tools to characterize the full complement of existing knowledge

about these economic observables. In particular, a probability model associated with the imagined sampling process can be defined that serves as a vehicle for describing our state of knowledge relating to how the observable economic data were obtained. Given this information repository, the fundamental econometric problem is concerned with transforming the conceptual probability model and the sample information associated with it into more specific knowledge about the unknown model components and parameters that represent characteristics of real-world economic processes.

1.4. The Process of Searching for Quantitative Economic Knowledge

In searching for quantitative economic knowledge contained in a sample of data, one must begin with some understanding of the economic process to which the data relate as well as some conception of the underlying sampling process by which the data were obtained. Otherwise, a sample of data is merely a collection of numbers with no contextual meaning or information value. Thus, to have a basis for interpreting the observed data, one needs a conceptual model of the process to which the data refer or some basis for specifying a *data-sampling process* (DSP) that links the sample of observations to our state of knowledge about how these observations were obtained.

The first step in this search for economic knowledge is for the analyst to identify an *economic process* that the analyst seeks to understand and about which there is incomplete knowledge and uncertainty (henceforth refer to Figure 1.1). By *process* we mean a particular method of doing something that generally involves several steps, operations, and interacting components and that leads to an observable outcome or result. For example, this might be the method by which, given prices and a budget constraint, consumers decide on market purchases (consumer decision process) or the method by which a commodity market leads to a product price (a market price equilibrium process).

The next step is a *process of abstraction* whereby *predata* theories, facts, assumptions, and an imagined sampling process whose outcomes are related to random variables are logically molded together into a *probability–econometric model* of the economic process. In a formal and often idealized way, the econometric model summarizes the analyst’s state of knowledge about the mechanisms that are thought to underlie the workings of the economic process under study and the sampling process by which observed data are obtained. The model, which is an abstraction, may be expressed in a variety of ways such as mathematical equations, algorithms, behavioral rules, diagrams, or all of these.

1.4.1. Econometric Model Components

It may be useful to think of an econometric model as being composed of components that include an economic model, a sampling model, and a probability model (Figure 1.1). The *economic model* component distinguishes an *econometric model* from a biological, physical, psychometric, or sociometric model. Models in other disciplines are defined

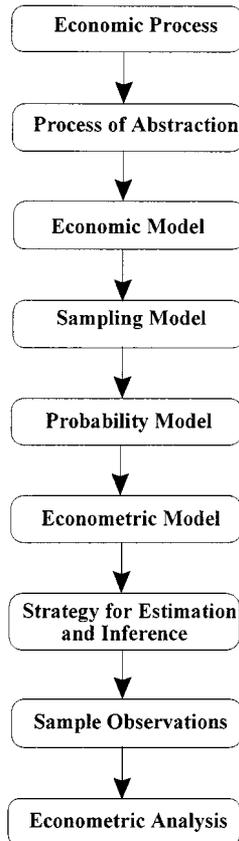


Figure 1.1: Process of Economic Information Recovery.

when appropriate discipline-specific theories, concepts, and knowledge are substituted in place of the economic model component. The economic model is based on a combination of the analyst's understanding of the institutions and mechanisms operating within the economic process being modeled and the economic theory thought to be relevant for explaining data outcomes produced by the economic process.

Once the economic model has been postulated, interpretations and questions relating to the workings of the economic process may be deduced from it. In this way the economic model provides a basis for defining relevant economic variables, forming tentative explanations, and suggesting hypotheses. However, this process of deduction tells us nothing *per se* about the actual truth or falsity of any explanations, hypotheses, or conclusions. It only ensures that conclusions *deductively* generated from the economic model are internally consistent with the definitions and postulates on which the model is based *provided that* the rules of logic have been applied correctly.

The *sampling model* characterizes a sampling process linkage between observable economic data and the postulated real-world components of the economic model. In particular, the sampling model identifies an imagined sampling process postulating that the observed data are the outcome of a collection of random variables

$\mathbf{Y} = \{Y_1, Y_2, \dots, Y_n\}$. At this point, general assumptions regarding the sampling characteristics of the random variables enter the analysis. For example are the random variables independent and identically distributed (iid), independent or dependent.

Moving farther in the direction of a formal basis for stating specific stochastic characteristics of the random variables in the imagined sampling process, the *probability model* postulates that the economic data are the outcome of some random variable or vector \mathbf{Y} having a joint probability distribution that belongs to some set of potential probability distributions, such as $\{F(\mathbf{y}; \boldsymbol{\theta}), \boldsymbol{\theta} \in \Omega\}$. If the elements in the set of probability distributions cannot be identified or indexed by a finite vector of parameter values $\boldsymbol{\theta}$, we may more generally denote the collection of probability distributions as $F(\mathbf{y}) \in \Psi$. By postulating such a probability model, the analyst effectively defines the range of possibilities for the joint probability distribution thought to characterize the behavior of potential sample outcomes. Unknown, uncontrolled, and unobservable components of the probability model are represented by parameters, random variables, or both. Together, the probability and sampling models identify both the candidates for the joint probability distribution of the observed data and the degree of interdependence, or lack thereof, among the individual data observations.

The combination of the probability, sampling, and economic models results in an *econometric model* that links a specified sampling process to the data. The adjective *econometric* arises from the realization, identification, and incorporation of an economic component into the formation and interpretation of the model. The econometric model represents our knowledge of the sampling of economic data in terms of a collection of random variables, $\mathbf{Y} = \{Y_1, Y_2, \dots, Y_n\}$, that have a certain economic interpretation, a certain dependence structure, and a joint probability distribution that belongs to some set of probability distributions $\{F(\mathbf{y}; \boldsymbol{\theta}), \boldsymbol{\theta} \in \Omega\}$ or $F(\mathbf{y}) \in \Psi$. Having defined the econometric model, the analyst has effectively specified a complete model of the sampling of the economic data under investigation. This means that, if values for the unknown and unobservable components of the perceived econometric model were known or assumed, the analyst would expect, or hope, that data consistent with the economic process being analyzed could be simulated from the econometric model. To the extent that the econometric model represents an accurate depiction of the true data-sampling process, the simulated outcomes could be used to produce additional samples of economic data relating to the economic process under study.

1.4.2. Econometric Analysis

Given a fully specified econometric model, the analyst has created a complete probabilistic and economic description and interpretation of the imagined sampling process for the economic data being analyzed. The model thus provides a complete picture of the analyst's state of knowledge about a set of economic outcomes and identifies *what is assumed* and *what is left to be discovered in the research process*.

An analyst's econometric model of sample outcomes for the random variables \mathbf{Y} is usually specified in terms of a *systematic* or *signal component* and an unobservable random *error, disturbance, or noise component* $\boldsymbol{\varepsilon}$. The two components are assumed to combine in a way that determines the *exact* values of observed sample outcomes.

In particular, the extent to which the value of \mathbf{Y} cannot be functionally represented in terms of the systematic component is accounted for by some function of ε , the content of which in some sense reflects the development of economics as a science. An example is the additive formulation $\mathbf{Y} = \mathbf{g}(\mathbf{X}, \boldsymbol{\theta}) + \varepsilon$ in which the sum of the systematic and noise components represents the sample outcome, where $\boldsymbol{\theta}$ is a vector of unknown and unobservable parameters and \mathbf{X} is a set of conditioning variables. More specific examples of the characterization of \mathbf{Y} in terms of systematic and noise components will be examined in Chapter 2.

Once the econometric model has been specified, the applied econometrician's objective is to proceed to the econometric analysis of the model. A necessary ingredient for such an analysis is the collection of *sample observations* of economic data relating to the economic process under study. The analyst must then also devise a *strategy for estimation and inference* in which appropriate statistical procedures for information discovery and recovery are identified within the context of the model being analyzed.

Given the sample observations and identified statistical procedures, the analyst then conducts an *econometric analysis* by applying the statistical procedures to the sample data and generating estimates and inferences. The analyst then provides a statistical and economic interpretation of the results obtained to complete the econometric analysis of the economic process.

1.5. The Inverse Problem

A challenge in econometric analyses is that unknown and uncontrolled components of the econometric model cannot generally be observed directly, and thus the analyst must use indirect observations based on observable data to recover information on these components. This challenge is associated with a concept in systems and information theory called the *inverse problem*, which is the problem of recovering information about unknown and uncontrolled components of a model from indirect observations on these components. The adjective *indirect* refers to the fact that, although the observed data are considered to be directly influenced by the values of model components, the observations are not themselves the direct values of these components but only indirectly reflect the influence of the components. Thus, the relationship characterizing the effect of unobservable components on the observed data must be somehow inverted to recover information about the unobservable model components from the data observations. Because econometric relations generally contain a systematic and a noise component, the problem of recovering information about unknowns and unobservables ($\boldsymbol{\theta}$, ε) from sample observations (\mathbf{y} , \mathbf{x}) within the context of an econometric model $\mathbf{Y} = \boldsymbol{\eta}(\mathbf{X}, \varepsilon, \boldsymbol{\theta})$ is referred to as an *inverse problem with noise*. A solution to this inverse problem is of the general form $(\mathbf{y}, \mathbf{x}) \Rightarrow (\boldsymbol{\theta}, \varepsilon)$.

Motivation for viewing a problem in econometric analysis as an inverse problem can be provided by a familiar illustration involving the theory of the firm. Firm managers need to make decisions concerning the profit-maximizing mix of inputs and levels of outputs under fixed prices. To make these optimizing decisions, information is needed on unknown components of the real-world production process such as the marginal

products of inputs. The marginal products are not directly observable, but we can observe the levels of outputs that result when various levels of inputs are used. These observable outcomes of the production process are *indirect* observations on the marginal products that, although not equal to the values of the marginal products themselves, are influenced by their values. Thus, we are confronted with an inverse problem: How can we best use the observed levels of input and output to recover information about the unobservable marginal products? At this point, in the absence of effective methods of information recovery, it must be clear that few, if any, rational or informed bets could be devised relative to the values of unknown, unobserved, and unobservable components of our econometric models. The principal objective of this book is to provide a foundation for the development of effective information recovery methods for this and other inverse problems.

1.6. A Comment

We view an economic–probability–econometric model as a starting point that lets us state, for all to see, what we are maintaining, or willing to assume, is known and what we consider unknown and seek to discover relative to an economic process under investigation. One of our econometric friends once remarked that he would rather be asked by a curious 3-year-old where babies come from than to try to answer the question, Where do econometric models come from? Perhaps it is less important where they come from than what the models represent, which is a starting place – a postulational base that leads to questions, experimentation, data collection, estimation, and finally inference and conclusions. In other words, they are the basis for a research process in which the model, the data, and the method of information recovery are interdependent links in the knowledge search and recovery chain. In Chapter 2, we review some interesting econometric questions and begin to examine the process of progressing from an economic model to a probability–econometric model. Our intention is to start the reader thinking about the estimation and inference methods for solving inverse problems with noise.

1.7. Notation

Before moving to Chapter 2 to begin our conceptualization of some alternative econometric models, we review here some notational conventions that give meaning to Figure 1.1 and the formulations we use in this book. A scalar random variable is denoted by a capital letter such as X or Y . A multivariate random variable in the form of a vector or matrix is denoted by a bold capital letter such as \mathbf{X} or \mathbf{Y} . A subscripted index distinguishes between different random variables. For example, we will use Y_i to indicate one representative of a collection of random variables, (Y_1, Y_2, \dots, Y_n) . Random variables whose outcomes we seek to explain will be referred to as *dependent variables*. We will also be interested in *explanatory variables*, whose values are used to help explain the values of dependent variables.

It is most often the case that an econometric model will contain more than one explanatory variable. An index is then needed if we wish to distinguish the explanatory variables from one another for a given observation. Depending on the circumstances, explanatory variables may either be fixed or random. In either case we will use a double subscript, where the first subscripted index denotes the observation number and the second specifies the particular explanatory variable number. We normally use \mathbf{X} or \mathbf{x} for a matrix of random or fixed explanatory variables, respectively. We emphasize that in either the random or fixed case, boldface denotes a vector or a matrix, whereas a nonboldfaced symbol will denote a scalar. Thus, if the explanatory variables are random, then X_{ij} represents the i th potential observation on the j th random explanatory variable. If the explanatory variables are fixed, then x_{ij} is the i th value of the j th fixed explanatory variable. An alternative notation for indicating the (i, j) th element of the explanatory variable matrix will be the standard matrix element notation $\mathbf{X}[i, j]$ or $\mathbf{x}[i, j]$. We will also represent the i th row of the explanatory variable matrix by \mathbf{X}_i or \mathbf{x}_i , or in standard matrix notation, by $\mathbf{X}[i, \cdot]$ or $\mathbf{x}[i, \cdot]$. The corresponding notation for designating the j th column of the explanatory variable matrix will be $\mathbf{X}_{\cdot j}$ or $\mathbf{x}_{\cdot j}$, and in standard matrix notation $\mathbf{X}[\cdot, j]$ or $\mathbf{x}[\cdot, j]$.

As an example of the preceding notation, we may develop a model of individual incomes Y_i using observations on explanatory variables representing scores on intelligence tests, years of schooling, highest degree obtained, grade point average (GPA), gender, race, and geographical region. In general functional notation we may then specify that $Y_i = g(\mathbf{X}_i)$ or $Y_i = g(\mathbf{X}[i, \cdot])$, for $i = 1, \dots, n$.

In representing general functions of variables, we will on occasion need to distinguish between *scalar* and *vector* functions of variables. The general notation will be $g(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$, respectively, where again, boldface denotes a vector. Furthermore, it is often the case in representing the systematic part of econometric models that the same functional definition is applied to each data observation. In this case, we will use the notation $g(\mathbf{x}_i)$ to denote the function applied to the observation \mathbf{x}_i and then continue to use $\mathbf{g}(\mathbf{x})$ to denote the vector of all of the observations, that is, $\mathbf{g}(\mathbf{x}) = [g(\mathbf{x}_1), g(\mathbf{x}_2), \dots, g(\mathbf{x}_n)]'$.

In some circumstances, we will want to characterize the values of dependent variables over time. In cases where we want to emphasize the temporal nature of the observations we will use a t subscript to denote the time index. For example, one may be interested in the values of a dependent and associated explanatory variables at n distinct time periods. The dependent and explanatory variables at time t will be denoted by Y_t and \mathbf{X}_t or $\mathbf{X}[t, \cdot]$, respectively. A data set of n observations over time would then consist of y_t and \mathbf{x}_t or $\mathbf{x}[t, \cdot]$ for $t = 1, \dots, n$, where y_t is the observed value of the random variable Y_t , and \mathbf{x}_t or $\mathbf{x}[t, \cdot]$ is the observed value of \mathbf{X}_t at time t .

Note that there will be a few exceptions to the conventions introduced above when precedent in the literature is so strong as to warrant an exception. An exception already encountered in the text is the use of ε to denote the *random* variable representing the noise component of an econometric model. Because we will later use \mathbf{e} to denote an outcome of the noise component, as is very often done in the literature, we avoid confusion with the letter \mathbf{E} , which is the conventional notation for mathematical expectation, and instead choose ε to be the random variable whose outcome is \mathbf{e} . We will be careful to

identify notational exceptions when they are first introduced in the text, and, regarding exceptions to the capital letter–random variable convention, we will endeavor to use Greek-letter alternatives.

Now that we have a context for discussing econometric models and the notation to represent them formally, in Chapter 2 we identify and classify a range of econometric models that will be of major interest as we work our way through the chapters to come.

1.8. Idea Checklist – Knowledge Guides

- 1.** Assume you are a theoretical econometrician. Identify a general format that you might use in developing a research project or reporting a working paper or journal article.
- 2.** Assume you are an applied econometrician. Identify a general format that you might use in developing a research project or reporting a working paper or journal article.
- 3.** To use later as a basis of comparison for how much your understanding of econometric analysis has matured, write a short essay on the topic: Where do econometric models come from?
- 4.** To use later as a basis of comparison for how much your understanding of econometric analysis has matured, write a short essay on the topic: Is econometrics necessary?
- 5.** Test your ability to specify a simple linear statistical model that involves a set of data from which you want to recover information on a mean-location level and a variance-scale parameter.