ESSENTIALS OF ECONOMETRICS

FOURTH EDITION

Damodar N. Gujarati

Professor Emeritus of Economics, United States Military Academy, West Point

Dawn C. Porter

University of Southern California

McGraw-Hill Irwin

Boston Burr Ridge, IL Dubuque, IA New York San Francisco St. Louis
Bangkok Bogotá Caracas Kuala Lumpur Lisbon London Madrid Mexico City
Milan Montreal New Delhi Santiago Seoul Singapore Sydney Taipei Toronto

ABOUT THE AUTHORS

DAMODAR N. GUJARATI

After teaching for more than 25 years at the City University of New York and 17 years in the Department of Social Sciences, U.S. Military Academy at West Point, New York, Dr. Gujarati is currently Professor Emeritus of economics at the Academy. Dr. Gujarati received his M.Com. degree from the University of Bombay in 1960, his M.B.A. degree from the University of Chicago in 1963, and his Ph.D. degree from the University of Chicago in 1965. Dr. Gujarati has published extensively in recognized national and international journals, such as the *Review of Economics and Statistics,* the *Economic Journal,* the *Journal of Financial and Quantitative Analysis,* and the *Journal of Business.* Dr. Gujarati was a member of the board of editors of the *Journal of Quantitative Economics,* the official journal of the Indian Econometric Society. Dr. Gujarati is also the author of *Pensions and the New York City Fiscal Crisis* (the American Enterprise Institute, 1978), *Government and Business* (McGraw-Hill, 1984), and *Basic Econometrics* (McGraw-Hill, 5th ed., 2009). Dr. Gujarati's books on econometrics have been translated into several languages.

Dr. Gujarati was a Visiting Professor at the University of Sheffield, U.K. (1970–1971), a Visiting Fulbright Professor to India (1981–1982), a Visiting Professor in the School of Management of the National University of Singapore 1985–1986), and a Visiting Professor of Econometrics, University of New South Wales, Australia (summer of 1988). Dr. Gujarati has lectured extensively on micro- and macroeconomic topics in countries such as Australia, China, Bangladesh, Germany, India, Israel, Mauritius, and the Republic of South Korea.

DAWN C. PORTER

Dawn Porter has been an assistant professor in the Information and Operations Management Department at the Marshall School of Business of the University of Southern California since the fall of 2006. She currently teaches undergraduate, M.B.A., and graduate elective statistics courses in the business school. Prior to joining the faculty at USC, from 2001–2006, Dawn was an assistant professor at the McDonough School of Business at Georgetown University and also served as a Visiting Professor in the Psychology Department at the Graduate School of Arts and Sciences at NYU. At NYU she taught a number of advanced statistical methods courses and was also an instructor at the Stern School of Business. Her Ph.D. is from the Stern School in Statistics, and her undergraduate degree is in mathematics from Cornell University.

Dawn's areas of research interest include categorical analysis, agreement measures, multivariate modeling, and applications to the field of psychology. Her current research examines online auction models from a statistical perspective. She has presented her research at the Joint Statistical Meetings, the Decision Sciences Institute meetings, the International Conference on Information Systems, several universities including the London School of Economics and NYU, and various e-commerce and statistics seminar series. Dawn is also a co-author on *Essentials of Business Statistics,* 2nd edition and *Basic Econometrics,* 5th edition, both from McGraw-Hill.

Outside academics, Dawn has been employed as a statistical consultant for KPMG, Inc. She also has worked as a statistical consultant for many other major companies, including Ginnie Mae, Inc.; Toys R Us Corporation; IBM; Cosmaire, Inc; and New York University (NYU) Medical Center.

CONTENTS

PREFACE xix

1 The Nature and Scope of Econometrics 1
1.1 WHAT IS ECONOMETRICS? 1
1.2 WHY STUDY ECONOMETRICS? 2
1.3 THE METHODOLOGY OF ECONOMETRICS 3
Creating a Statement of Theory or Hypothesis 3
Collecting Data 4
Specifying the Mathematical Model of Labor Force Participation 5
Specifying the Statistical, or Econometric, Model of Labor Force Participation 7
Estimating the Parameters of the Chosen Econometric Model 9
Checking for Model Adequacy: Model Specification Testing 9
Testing the Hypothesis Derived from the Model 11
Using the Model for Prediction or Forecasting 12
1.4 THE ROAD AHEAD 12
KEY TERMS AND CONCEPTS 13
QUESTIONS 14
PROBLEMS 14
APPENDIX 1A: ECONOMIC DATA ON THE WORLD WIDE WEB 16

PART I THE LINEAR REGRESSION MODEL 19

2 Basic Ideas of Linear Regression: The Two-Variable Model 21
2.1 THE MEANING OF REGRESSION 21
2.2 THE POPULATION REGRESSION FUNCTION (PRF): A HYPOTHETICAL EXAMPLE 22

2.3 STATISTICAL OR STOCHASTIC SPECIFICATION OF THE POPULATION REGRESSION FUNCTION 25
2.4 THE NATURE OF THE STOCHASTIC ERROR TERM 27
2.5 THE SAMPLE REGRESSION FUNCTION (SRF) 28
2.6 THE SPECIAL MEANING OF THE TERM "LINEAR" REGRESSION 31
Linearity in the Variables 31
Linearity in the Parameters 32
2.7 TWO-VARIABLE VERSUS MULTIPLE LINEAR REGRESSION 33
2.8 ESTIMATION OF PARAMETERS: THE METHOD OF ORDINARY LEAST SQUARES 33
The Method of Ordinary Least Squares 34
2.9 PUTTING IT ALL TOGETHER 36
Interpretation of the Estimated Math S.A.T. Score Function 37
2.10 SOME ILLUSTRATIVE EXAMPLES 38
2.11 SUMMARY 43
KEY TERMS AND CONCEPTS 44
QUESTIONS 44
PROBLEMS 45
OPTIONAL QUESTIONS 51
APPENDIX 2A: DERIVATION OF LEAST-SQUARES ESTIMATES 52

3 The Two-Variable Model: Hypothesis Testing 53
3.1 THE CLASSICAL LINEAR REGRESSION MODEL 54
3.2 VARIANCES AND STANDARD ERRORS OF ORDINARY LEAST SQUARES ESTIMATORS 57
Variances and Standard Errors of the Math S.A.T. Score Example 59
Summary of the Math S.A.T. Score Function 59
3.3 WHY OLS? THE PROPERTIES OF OLS ESTIMATORS 60
Monte Carlo Experiment 61
3.4 THE SAMPLING, OR PROBABILITY, DISTRIBUTIONS OF OLS ESTIMATORS 62
3.5 HYPOTHESIS TESTING 64
Testing H_0:$B_2 = 0$ versus H_1:$B_2 \neq 0$: The Confidence Interval Approach 66
The Test of Significance Approach to Hypothesis Testing 68
Math S.A.T. Example Continued 69
3.6 HOW GOOD IS THE FITTED REGRESSION LINE: THE COEFFICIENT OF DETERMINATION, r^2 71
Formulas to Compute r^2 73
r^2 for the Math S.A.T. Example 74
The Coefficient of Correlation, r 74
3.7 REPORTING THE RESULTS OF REGRESSION ANALYSIS 75

3.8 COMPUTER OUTPUT OF THE MATH S.A.T. SCORE EXAMPLE 76
3.9 NORMALITY TESTS 77
Histograms of Residuals 77
Normal Probability Plot 78
Jarque-Bera Test 78
3.10 A CONCLUDING EXAMPLE: RELATIONSHIP BETWEEN WAGES AND PRODUCTIVITY IN THE U.S. BUSINESS SECTOR, 1959–2006 79
3.11 A WORD ABOUT FORECASTING 82
3.12 SUMMARY 85
KEY TERMS AND CONCEPTS 86
QUESTIONS 86
PROBLEMS 88

4 Multiple Regression: Estimation and Hypothesis Testing 93
4.1 THE THREE-VARIABLE LINEAR REGRESSION MODEL 94
The Meaning of Partial Regression Coefficient 95
4.2 ASSUMPTIONS OF THE MULTIPLE LINEAR REGRESSION MODEL 97
4.3 ESTIMATION OF THE PARAMETERS OF MULTIPLE REGRESSION 99
Ordinary Least Squares Estimators 99
Variance and Standard Errors of OLS Estimators 100
Properties of OLS Estimators of Multiple Regression 102
4.4 GOODNESS OF FIT OF ESTIMATED MULTIPLE REGRESSION: MULTIPLE COEFFICIENT OF DETERMINATION, R^2 102
4.5 ANTIQUE CLOCK AUCTION PRICES REVISITED 103
Interpretation of the Regression Results 103
4.6 HYPOTHESIS TESTING IN A MULTIPLE REGRESSION: GENERAL COMMENTS 104
4.7 TESTING HYPOTHESES ABOUT INDIVIDUAL PARTIAL REGRESSION COEFFICIENTS 105
The Test of Significance Approach 105
The Confidence Interval Approach to Hypothesis Testing 106
4.8 TESTING THE JOINT HYPOTHESIS THAT $B_2 = B_3 = 0$ OR $R^2 = 0$ 107
An Important Relationship between F and R^2 111
4.9 TWO-VARIABLE REGRESSION IN THE CONTEXT OF MULTIPLE REGRESSION: INTRODUCTION TO SPECIFICATION BIAS 112
4.10 COMPARING TWO R^2 VALUES: THE ADJUSTED R^2 113

4.11 WHEN TO ADD AN ADDITIONAL EXPLANATORY VARIABLE TO A MODEL 114
4.12 RESTRICTED LEAST SQUARES 116
4.13 ILLUSTRATIVE EXAMPLES 117
Discussion of Regression Results 118
4.14 SUMMARY 122
KEY TERMS AND CONCEPTS 123
QUESTIONS 123
PROBLEMS 125
APPENDIX 4A.1: DERIVATIONS OF OLS ESTIMATORS GIVEN IN EQUATIONS (4.20) TO (4.22) 129
APPENDIX 4A.2: DERIVATION OF EQUATION (4.31) 129
APPENDIX 4A.3: DERIVATION OF EQUATION (4.50) 130
APPENDIX 4A.4: EVIEWS OUTPUT OF THE CLOCK AUCTION PRICE EXAMPLE 131

5 Functional Forms of Regression Models 132
5.1 HOW TO MEASURE ELASTICITY: THE LOG-LINEAR MODEL 133
Hypothesis Testing in Log-Linear Models 137
5.2 COMPARING LINEAR AND LOG-LINEAR REGRESSION MODELS 138
5.3 MULTIPLE LOG-LINEAR REGRESSION MODELS 140
5.4 HOW TO MEASURE THE GROWTH RATE: THE SEMILOG MODEL 144
Instantaneous versus Compound Rate of Growth 147
The Linear Trend Model 148
5.5 THE LIN-LOG MODEL: WHEN THE EXPLANATORY VARIABLE IS LOGARITHMIC 149
5.6 RECIPROCAL MODELS 150
5.7 POLYNOMIAL REGRESSION MODELS 156
5.8 REGRESSION THROUGH THE ORIGIN 158
5.9 A NOTE ON SCALING AND UNITS OF MEASUREMENT 160
5.10 REGRESSION ON STANDARDIZED VARIABLES 161
5.11 SUMMARY OF FUNCTIONAL FORMS 163
5.12 SUMMARY 164
KEY TERMS AND CONCEPTS 165
QUESTIONS 166
PROBLEMS 167
APPENDIX 5A: LOGARITHMS 175

6 Dummy Variable Regression Models 178
6.1 THE NATURE OF DUMMY VARIABLES 178
6.2 ANCOVA MODELS: REGRESSION ON ONE QUANTITATIVE VARIABLE AND ONE QUALITATIVE VARIABLE WITH TWO CATEGORIES: EXAMPLE 6.1 REVISITED 185

6.3 REGRESSION ON ONE QUANTITATIVE VARIABLE AND ONE QUALITATIVE VARIABLE WITH MORE THAN TWO CLASSES OR CATEGORIES 187
6.4 REGRESSION ON ONE QUANTIATIVE EXPLANATORY VARIABLE AND MORE THAN ONE QUALITATIVE VARIABLE 190
Interaction Effects 191
A Generalization 192
6.5 COMPARING TWO REGESSIONS 193
6.6 THE USE OF DUMMY VARIABLES IN SEASONAL ANALYSIS 198
6.7 WHAT HAPPENS IF THE DEPENDENT VARIABLE IS ALSO A DUMMY VARIABLE? THE LINEAR PROBABILITY MODEL (LPM) 201
6.8 SUMMARY 204
KEY TERMS AND CONCEPTS 205
QUESTIONS 206
PROBLEMS 207

PART II REGRESSION ANALYSIS IN PRACTICE 217

7 Model Selection: Criteria and Tests 219
7.1 THE ATTRIBUTES OF A GOOD MODEL 220
7.2 TYPES OF SPECIFICATION ERRORS 221
7.3 OMISSON OF RELEVANT VARIABLE BIAS: "UNDERFITTING" A MODEL 221
7.4 INCLUSION OF IRRELEVANT VARIABLES: "OVERFITTING" A MODEL 225
7.5 INCORRECT FUNCTIONAL FORM 227
7.6 ERRORS OF MEASUREMENT 229
Errors of Measurement in the Dependent Variable 229
Errors of Measurement in the Explanatory Variable(s) 229
7.7 DETECTING SPECIFICATION ERRORS: TESTS OF SPECIFICATION ERRORS 230
Detecting the Presence of Unnecessary Variables 230
Tests for Omitted Variables and Incorrect Functional Forms 233
Choosing between Linear and Log-linear Regression Models: The MWD Test 235
Regression Error Specification Test: RESET 237
7.8 SUMMARY 239
KEY TERMS AND CONCEPTS 240
QUESTIONS 240
PROBLEMS 241

8 Multicollinearity: What Happens If Explanatory Variables are Correlated? 245
8.1 THE NATURE OF MULTICOLLINEARITY: THE CASE OF PERFECT MULTICOLLINEARITY 246
8.2 THE CASE OF NEAR, OR IMPERFECT, MULTICOLLINEARITY 248
8.3 THEORETICAL CONSEQUENCES OF MULTICOLLINEARITY 250
8.4 PRACTICAL CONSEQUENCES OF MULTICOLLINEARITY 251
8.5 DETECTION OF MULTICOLLINEARITY 253
8.6 IS MULTICOLLINEARITY NECESSARILY BAD? 258
8.7 AN EXTENDED EXAMPLE: THE DEMAND FOR CHICKENS IN THE UNITED STATES, 1960 TO 1982 259
Collinearity Diagnostics for the Demand Function for Chickens (Equation [8.15]) 260
8.8 WHAT TO DO WITH MULTICOLLINEARITY: REMEDIAL MEASURES 261
Dropping a Variable(s) from the Model 262
Acquiring Additional Data or a New Sample 262
Rethinking the Model 263
Prior Information about Some Parameters 264
Transformation of Variables 265
Other Remedies 266
8.9 SUMMARY 266
KEY TERMS AND CONCEPTS 267
QUESTIONS 267
PROBLEMS 268

9 Heteroscedasticity: What Happens If the Error Variance Is Nonconstant? 274
9.1 THE NATURE OF HETEROSCEDASTICITY 274
9.2 CONSEQUENCES OF HETEROSCEDASTICITY 280
9.3 DETECTION OF HETEROSCEDASTICITY: HOW DO WE KNOW WHEN THERE IS A HETEROSCEDASTICITY PROBLEM? 282
Nature of the Problem 283
Graphical Examination of Residuals 283
Park Test 285
Glejser Test 287
White's General Heteroscedasticity Test 289
Other Tests of Heteroscedasticity 290
9.4 WHAT TO DO IF HETEROSCEDASTICITY IS OBSERVED: REMEDIAL MEASURES 291
When σ_i^2 Is Known: The Method of Weighted Least Squares (WLS) 291
When True σ_i^2 Is Unknown 292
Respecification of the Model 297

9.5 WHITE'S HETEROSCEDASTICITY-CORRECTED STANDARD ERRORS AND *t* STATISTICS 298
9.6 SOME CONCRETE EXAMPLES OF HETEROSCEDASTICITY 299
9.7 SUMMARY 302
KEY TERMS AND CONCEPTS 303
QUESTIONS 304
PROBLEMS 304

10 Autocorrelation: What Happens If Error Terms Are Correlated? 312
10.1 THE NATURE OF AUTOCORRELATION 313
Inertia 314
Model Specification Error(s) 315
The Cobweb Phenomenon 315
Data Manipulation 315
10.2 CONSEQUENCES OF AUTOCORRELATION 316
10.3 DETECTING AUTOCORRELATION 317
The Graphical Method 318
The Durbin-Watson *d* Test 320
10.4 REMEDIAL MEASURES 325
10.5 HOW TO ESTIMATE ρ 327
$\rho = 1$: The First Difference Method 327
ρ Estimated from Durbin-Watson *d* Statistic 327
ρ Estimated from OLS Residuals, e_t 328
Other Methods of Estimating ρ 328
10.6 A LARGE SAMPLE METHOD OF CORRECTING OLS STANDARD ERRORS: THE NEWEY-WEST (NW) METHOD 332
10.7 SUMMARY 334
KEY TERMS AND CONCEPTS 335
QUESTIONS 335
PROBLEMS 336
APPENDIX 10A: THE RUNS TEST 341
Swed-Eisenhart Critical Runs Test 342
Decision Rule 342
APPENDIX 10B: A GENERAL TEST OF AUTOCORRELATION: THE BREUSCH-GODFREY (BG) TEST 343

PART III ADVANCED TOPICS IN ECONOMETRICS 345

11 Simultaneous Equation Models 347
11.1 THE NATURE OF SIMULTANEOUS EQUATION MODELS 348
11.2 THE SIMULTANEOUS EQUATION BIAS: INCONSISTENCY OF OLS ESTIMATORS 350

11.3 THE METHOD OF INDIRECT LEAST SQUARES (ILS) 352
11.4 INDIRECT LEAST SQUARES: AN ILLUSTRATIVE EXAMPLE 353
11.5 THE IDENTIFICATION PROBLEM: A ROSE BY ANY OTHER NAME MAY NOT BE A ROSE 355
Underidentification 356
Just or Exact Identification 357
Overidentification 359
11.6 RULES FOR IDENTIFICATION: THE ORDER CONDITION OF IDENTIFICATION 361
11.7 ESTIMATION OF AN OVERIDENTIFIED EQUATION: THE METHOD OF TWO-STAGE LEAST SQUARES 362
11.8 2SLS: A NUMERICAL EXAMPLE 364
11.9 SUMMARY 365
KEY TERMS AND CONCEPTS 366
QUESTIONS 367
PROBLEMS 367
APPENDIX 11A: INCONSISTENCY OF OLS ESTIMATORS 369

12 Selected Topics in Single Equation Regression Models 371
12.1 DYNAMIC ECONOMIC MODELS: AUTOREGRESSIVE AND DISTRIBUTED LAG MODELS 371
Reasons for Lag 372
Estimation of Distributed Lag Models 374
The Koyck, Adaptive Expectations, and Stock Adjustment Models Approach to Estimating Distributed Lag Models 377
12.2 THE PHENOMENON OF SPURIOUS REGRESSION: NONSTATIONARY TIME SERIES 380
12.3 TESTS OF STATIONARITY 382
12.4 COINTEGRATED TIME SERIES 383
12.5 THE RANDOM WALK MODEL 384
12.6 THE LOGIT MODEL 386
Estimation of the Logit Model 390
12.7 SUMMARY 396
KEY TERMS AND CONCEPTS 397
QUESTIONS 397
PROBLEMS 398

INTRODUCTION TO APPENDIXES A, B, C, AND D: BASICS OF PROBABILITY AND STATISTICS 403

Appendix A: Review of Statistics: Probability and Probability Distributions 405
A.1 SOME NOTATION 405
The Summation Notation 405
Properties of the Summation Operator 406

A.2 EXPERIMENT, SAMPLE SPACE, SAMPLE POINT, AND EVENTS 407
Experiment 407
Sample Space or Population 407
Sample Point 408
Events 408
Venn Diagrams 408
A.3 RANDOM VARIABLES 409
A.4 PROBABILITY 410
Probability of an Event: The Classical or A Priori Definition 410
Relative Frequency or Empirical Definition of Probability 411
Probability of Random Variables 417
A.5 RANDOM VARIABLES AND THEIR PROBABILITY DISTRIBUTIONS 417
Probability Distribution of a Discrete Random Variable 417
Probability Distribution of a Continuous Random Variable 419
Cumulative Distribution Function (CDF) 420
A.6 MULTIVARIATE PROBABILITY DENSITY FUNCTIONS 422
Marginal Probability Functions 424
Conditional Probability Functions 425
Statistical Independence 427
A.7 SUMMARY AND CONCLUSIONS 428
KEY TERMS AND CONCEPTS 428
REFERENCES 429
QUESTIONS 429
PROBLEMS 430

Appendix B: Characteristics of Probability Distributions 434
B.1 EXPECTED VALUE: A MEASURE OF CENTRAL TENDENCY 434
Properties of Expected Value 436
Expected Value of Multivariate Probability Distributions 437
B.2 VARIANCE: A MEASURE OF DISPERSION 438
Properties of Variance 439
Chebyshev's Inequality 441
Coefficient of Variation 442
B.3 COVARIANCE 443
Properties of Covariance 444
B.4 CORRELATION COEFFICIENT 445
Properties of Correlation Coefficient 445
Variances of Correlated Variables 447
B.5 CONDITIONAL EXPECTATION 447
Conditional Variance 449
B.6 SKEWNESS AND KURTOSIS 449
B.7 FROM THE POPULATION TO THE SAMPLE 452
Sample Mean 452

	Sample Variance	453
	Sample Covariance	454
	Sample Correlation Coefficient	455
	Sample Skewness and Kurtosis	456
B.8	SUMMARY	456
	KEY TERMS AND CONCEPTS	457
	QUESTIONS	457
	PROBLEMS	458
	OPTIONAL EXERCISES	460
Appendix C: Some Important Probability Distributions		**461**
C.1	THE NORMAL DISTRIBUTION	462
	Properties of the Normal Distribution	462
	The Standard Normal Distribution	464
	Random Sampling from a Normal Population	468
	The Sampling or Probability Distribution of the Sample Mean $\bar{X}$	468
	The Central Limit Theorem (CLT)	472
C.2	THE *t* DISTRIBUTION	473
	Properties of the *t* Distribution	474
C.3	THE CHI-SQUARE (χ^2) PROBABILITY DISTRIBUTION	477
	Properties of the Chi-square Distribution	478
C.4	THE *F* DISTRIBUTION	480
	Properties of the *F* Distribution	481
C.5	SUMMARY	483
	KEY TERMS AND CONCEPTS	483
	QUESTIONS	484
	PROBLEMS	484
Appendix D: Statistical Inference: Estimation and Hypothesis Testing		**487**
D.1	THE MEANING OF STATISTICAL INFERENCE	487
D.2	ESTIMATION AND HYPOTHESIS TESTING: TWIN BRANCHES OF STATISTICAL INFERENCE	489
D.3	ESTIMATION OF PARAMETERS	490
D.4	PROPERTIES OF POINT ESTIMATORS	493
	Linearity	494
	Unbiasedness	494
	Minimum Variance	495
	Efficiency	496
	Best Linear Unbiased Estimator (BLUE)	497
	Consistency	497
D.5	STATISTICAL INFERENCE: HYPOTHESIS TESTING	498
	The Confidence Interval Approach to Hypothesis Testing	499
	Type I and Type II Errors: A Digression	500
	The Test of Significance Approach to Hypothesis Testing	503

A Word on Choosing the Level of Significance, α, and the *p* Value 506
The χ^2 and *F* Tests of Significance 507
D.6 SUMMARY 510
KEY TERMS AND CONCEPTS 510
QUESTIONS 511
PROBLEMS 512

Appendix E: Statistical Tables 515

Appendix F: Computer Output of EViews, MINITAB, Excel, and STATA 534

SELECTED BIBLIOGRAPHY 541

INDEXES 545
Name Index 545
Subject Index 547