

BASIC ECONOMETRICS



Damodar N. Gujarati

Third Edition



McGRAW-HILL INTERNATIONAL EDITIONS
Economic Series

BASIC ECONOMETRICS

Third Edition

Damodar N. Gujarati

United States Military Academy, West Point

McGraw-Hill, Inc.

New York St. Louis San Francisco Auckland Bogotá Caracas
Lisbon London Madrid Mexico City Milan Montreal New Delhi
San Juan Singapore Sydney Tokyo Toronto

BASIC ECONOMETRICS
International Editions 1995

Exclusive rights by McGraw-Hill Book Co. – Singapore for manufacture and export. This book cannot be re-exported from the country to which it is consigned by McGraw-Hill.

Copyright © 1995, 1988, 1978 by McGraw-Hill, Inc. All rights reserved. Except as permitted under the Copyright Act of 1976, no part of this publication may be reproduced or distributed in any form or by any means, or stored in a data base or retrieval system, without the prior written permission of the publisher.

4 5 6 7 8 9 0 CMO PMP 9 8 7 6

This book was set in New Aster by Publication Services, Inc.
The editors were Scott D. Stratford and Lucille H. Sutton;
the production supervisor was Friederich W. Schulte.
The cover was designed by Tana Kamine.
Project supervision was done by Publication Services, Inc.

Library of Congress Cataloging-in-Publication Data

Gujarati, Damodar N.

Basic econometrics / Damodar N. Gujarati. – 3rd ed.
p. cm.

Includes bibliographical references (p.) and index.

ISBN 0-07-025214-9

1. Econometrics. I. Title.

HB139.G84 1995

330'.01'5118–dc20

94-35295

When ordering this title, use ISBN 0-07-113964-8

Printed in Singapore

ABOUT THE AUTHOR

After teaching for more than 28 years at the City University of New York, Damodar N. Gujarati is currently a professor of economics in the Department of Social Sciences at the U.S. Military Academy at West Point, New York. 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*, the *Journal of Business*, the *American Statistician*, and the *Journal of Industrial and Labor Relations*. Dr. Gujarati is an editorial referee to several journals and book publishers and is 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 *Essentials of Econometrics* (McGraw-Hill, 1992). 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). As a regular participant in USIA's lectureship program abroad, Dr. Gujarati has lectured extensively on micro- and macroeconomic topics in countries such as Australia, Bangladesh, Germany, India, Israel, Mauritius, and the Republic of South Korea. Dr. Gujarati has also given seminars and lectures in Canada and Mexico.

To the Memory of
“Akka” (Shalini)
and
“Suru” (Suryakant)

PREFACE

As in the previous two editions, the primary objective of the third edition of *Basic Econometrics* is to provide an elementary but comprehensive introduction to econometrics without resorting to matrix algebra, calculus, or statistics beyond the elementary level.

In this edition I have attempted to incorporate some of the developments in the theory and practice of econometrics that have taken place since the publication of the second edition in 1988. In addition, this revision has given me the opportunity to simplify the discussion of some topics included in the previous editions and to add some new material on these topics. The major changes in this edition are as follows:

1. In Chapter 1, I have expanded the discussion of the nature and sources of data available for econometric analysis. In view of the increasing use of time series data in economic analysis, I have introduced very early the concept of a **stationary time series**, a concept that is crucial for analyzing data involving economic time series.
2. In Chapter 3, I present a more extended discussion of the assumptions of the classical linear regression model (CLRM). The CLRM is the foundation of econometrics. In this chapter I also discuss the **Monte Carlo** simulation experiments.
3. In Chapter 5, on hypothesis testing, I have introduced the concept of the ***p* value**, or the exact level of significance, of a test statistic. In this chapter, I also discuss the **Jarque-Bera test of normality**.
4. In Chapter 8, on hypothesis testing in the context of multiple regression models, I have streamlined the discussion. This chapter also includes a discussion of the choice between linear and log-linear regression models. In the appendix to this chapter, I discuss, at an elementary level, the **likelihood ratio (LR) test of hypothesis**.
5. In Chapter 10, on multicollinearity, I now give equal billing to **micro-numerosity** (smallness of sample size), a concept due to Arthur Gold-

- berger. I also introduce **tolerance** and **inflation-variance** tools for detecting multicollinearity.
6. In Chapter 11, on heteroscedasticity, I have now included the **Breusch-Pagan-Godfrey test** and **White's test** of heteroscedasticity. I also discuss White's heteroscedasticity-consistent variances and standard errors of OLS estimators.
 7. In Chapter 12, on autocorrelation, I have included these tests: asymptotic test of autocorrelation, Breusch-Godfrey test of higher-order autocorrelation, and Berenblut-Webb test. Included in this chapter is also the **ARCH model**, which has been increasingly used in financial economics.
 8. Chapter 13, on model building, discusses nominal versus true level of significance in the presence of data mining, and the **Lagrange multiplier (LM) test** for adding variables to a regression model.
 9. In Chapter 14, which is new, I discuss alternatives to the traditional econometric methodology. In particular, I discuss Leamer's and Hendry's approaches to econometrics. Also included in this chapter are tests of non-nested hypotheses, in particular the **Davidson-MacKinnon J test**.
 10. Chapter 15, on dummy variables, now includes a discussion of dummies in combining time series and cross-sectional data. I also show how the dummies can be used in the presence of autocorrelation and heteroscedasticity. An exercise in this chapter discusses Zellner's **seemingly unrelated regression (SURE)** technique.
 11. Chapter 16, on dummy dependent variable regression models, now includes a discussion of the **Tobit model**.
 12. Chapter 17, on dynamic regression models, now includes a discussion of both the **Granger test** and **Sims's test** of causality.
 13. Chapters 18, 19, and 20, on simultaneous-equation models, now contain tests of simultaneity and exogeneity. These chapters also discuss the relationship between causality and exogeneity.
 14. In recognition of the growing importance of time series data in economic analysis, I have included two new chapters on time series econometrics. In Chapter 21, I introduce the key concepts of time series analysis, such as **stationarity**, **random walk**, **unit root**, **Dickey-Fuller** and **augmented Dickey-Fuller tests of stationarity**, **deterministic and stochastic trends**, **trend-stationary** and **difference-stationary stochastic processes**, **cointegration**, **Engle-Granger tests of cointegration**, **error correction mechanism**, and **spurious regression**. In Chapter 22 I discuss the **Box-Jenkins**, or **ARIMA**, and **vector autoregression (VAR)** approaches to economic forecasting. These are alternatives to the traditional single- and simultaneous-equation approaches to forecasting.

I have added several new exercises. The exercises given at the end of chapters are now divided into two groups: questions and problems. The latter are data-based exercises (I am a firm believer in learning by doing).

All these changes have considerably expanded the scope of this book. I hope this gives the instructor substantial flexibility in choosing topics that are appropriate to the intended audience. Here are some suggestions about how this book may be used: **One-semester course for the nonspecialist:** Appendix A; Chaps. 1 through 8; an overview of Chaps. 10, 11, and 12 (omitting all the proofs); and Chap. 15. The theoretical exercises can be omitted. **One-semester course for economics majors:** Appendix A; Chaps. 1 through 8; and Chaps. 10 through 15. If matrix algebra is used, include Appendix B and Chap. 9. Some of the theoretical exercises can be omitted. **Two-semester course for economics majors:** Appendices A and B, and Chaps. 1 through 22. Mathematical proofs given in the various appendices can be covered on a selective basis. Additionally, the instructor may want to cover the topic of nonlinear (in parameters) regression models.

This revision would not have been possible without the constructive comments, suggestions, and encouragement that I have received from several people who have read the various drafts. In particular, I would like to acknowledge my debt to the following professors, without, of course, holding them responsible for any deficiencies that remain in the book: Ted Amato, University of North Carolina; Dale Belman, University of Wisconsin–Milwaukee; Tom Daula, U.S. Military Academy; Mary Deily, Lehigh University; Frank Diebold, University of Pennsylvania; David Garman, Tufts University; Sushila Gidwani-Bushchi, Manhattan College; William Greene, New York University; Dennis Jansen, Texas A&M University; Jane Lillydahl, University of Colorado; Dagmar Rajagopal, Ryerson Polytechnic University; Bo Ruck, U.S. Military Academy; John Spitzer, State University of New York, Brockport; and H. D. Vinod, Fordham University.

I owe a great debt of gratitude to Professor Kenneth J. White and Steven A. Theobald for checking numerical calculations and for preparing the handbook: *Basic Econometrics: A Computer Handbook Using SHAZAM*, 3d ed., McGraw-Hill, New York, 1995. My wife Pushpa and my daughters Joan and Diane have been a constant source of inspiration and encouragement. Beyond a simple thank-you, I want them to know that I love them all very dearly.

On a personal note, after some 28 years of teaching at the City University of New York (CUNY), I have now joined the Department of Social Sciences at the U.S. Military Academy at West Point, New York. I am grateful to CUNY for providing me with my first job and to the Academy for offering me new challenges and opportunities.

Damodar N. Gujarati

BASIC ECONOMETRICS

CONTENTS

Preface	xxi
Introduction	1

Part 1 Single-Equation Regression Models

1	The Nature of Regression Analysis	15
1.1	Historical Origin of the Term "Regression"	15
1.2	The Modern Interpretation of Regression	16
	Examples	16
1.3	Statistical vs. Deterministic Relationships	19
1.4	Regression vs. Causation	20
1.5	Regression vs. Correlation	21
1.6	Terminology and Notation	22
1.7	The Nature and Sources of Data for Econometric	
	Analysis	23
	Types of Data	23
	The Sources of Data	24
	The Accuracy of Data	26
1.8	Summary and Conclusions	27
	Exercises	28
	Appendix 1A	29
	1A.1 Sources of Economic Data	29
	1A.2 Sources of Financial Data	31
2	Two-Variable Regression Analysis:	
	Some Basic Ideas	32
2.1	A Hypothetical Example	32
2.2	The Concept of Population Regression Function (PRF)	36
2.3	The Meaning of the Term "Linear"	36
	Linearity in the Variables	37
	Linearity in the Parameters	37
2.4	Stochastic Specification of PRF	38

2.5	The Significance of the Stochastic Disturbance Term	39
2.6	The Sample Regression Function (SRF)	41
2.7	Summary and Conclusions	45
	Exercises	45
3	Two-Variable Regression Model:	
	The Problem of Estimation	52
3.1	The Method of Ordinary Least Squares	52
3.2	The Classical Linear Regression Model: The Assumptions	
	Underlying the Method of Least Squares	59
	How Realistic Are These Assumptions?	68
3.3	Precision or Standard Errors of Least-Squares	
	Estimates	69
3.4	Properties of Least-Squares Estimators:	
	The Gauss-Markov Theorem	72
3.5	The Coefficient of Determination r^2 : A Measure	
	of "Goodness of Fit"	74
3.6	A Numerical Example	80
3.7	Illustrative Examples	83
	Coffee Consumption in the United States, 1970–1980	83
	Keynesian Consumption Function for the United	
	States, 1980–1991	84
3.8	Computer Output for the Coffee Demand Function	85
3.9	A Note on Monte Carlo Experiments	85
3.10	Summary and Conclusions	86
	Exercises	87
	Questions	87
	Problems	89
	Appendix 3A	94
	3A.1 Derivation of Least-Squares Estimates	94
	3A.2 Linearity and Unbiasedness Properties	
	of Least-Squares Estimators	94
	3A.3 Variances and Standard Errors of Least-Squares	
	Estimators	95
	3A.4 Covariance between $\hat{\beta}_1$ and $\hat{\beta}_2$	96
	3A.5 The Least-Squares Estimator of σ^2	96
	3A.6 Minimum-Variance Property of Least-Squares	
	Estimators	97
	3A.7 SAS Output of the Coffee Demand Function	
	(3.7.1)	99
4	The Normality Assumption: Classical Normal	
	Linear Regression Model (CNLRM)	101
4.1	The Probability Distribution of Disturbances u_i	101
4.2	The Normality Assumption	102
4.3	Properties of OLS Estimators under the Normality	
	Assumption	104
4.4	The Method of Maximum Likelihood (ML)	107
4.5	Probability Distributions Related to the	
	Normal Distribution: The t , Chi-square (χ^2),	
	and F Distributions	107

4.6	Summary and Conclusions	109
	Appendix 4A	110
	Maximum Likelihood Estimation of Two-Variable Regression Model	110
	Maximum Likelihood Estimation of the Consumption-Income Example	113
	Appendix 4A Exercises	113
5	Two-Variable Regression: Interval Estimation and Hypothesis Testing	115
5.1	Statistical Prerequisites	115
5.2	Interval Estimation: Some Basic Ideas	116
5.3	Confidence Intervals for Regression Coefficients β_1 and β_2	117
	Confidence Interval for β_2	117
	Confidence Interval for β_1	119
	Confidence Interval for β_1 and β_2 Simultaneously	120
5.4	Confidence Interval for σ^2	120
5.5	Hypothesis Testing: General Comments	121
5.6	Hypothesis Testing: The Confidence-Interval Approach	122
	Two-Sided or Two-Tail Test	122
	One-Sided or One-Tail Test	124
5.7	Hypothesis Testing: The Test-of-Significance Approach	124
	Testing the Significance of Regression Coefficients: The t -Test	124
	Testing the Significance of σ^2 : The χ^2 Test	128
5.8	Hypothesis Testing: Some Practical Aspects	129
	The Meaning of "Accepting" or "Rejecting" a Hypothesis	129
	The "Zero" Null Hypothesis and the "2- t " Rule of Thumb	129
	Forming the Null and Alternative Hypotheses	130
	Choosing α , the Level of Significance	131
	The Exact Level of Significance: The p Value	132
	Statistical Significance versus Practical Significance	133
	The Choice between Confidence-Interval and Test-of-Significance Approaches to Hypothesis Testing	134
5.9	Regression Analysis and Analysis of Variance	134
5.10	Application of Regression Analysis: The Problem of Prediction	137
	Mean Prediction	137
	Individual Prediction	138
5.11	Reporting the Results of Regression Analysis	140
5.12	Evaluating the Results of Regression Analysis	140
	Normality Test	141
	Other Tests of Model Adequacy	144
5.13	Summary and Conclusions	144
	Exercises	145
	Questions	145
	Problems	147

Appendix 5A	152
5A.1 Derivation of Equation (5.3.2)	152
5A.2 Derivation of Equation (5.9.1)	152
5A.3 Derivations of Equations (5.10.2) and (5.10.6)	153
Variance of Mean Prediction	153
Variance of Individual Prediction	153
6 Extensions of the Two-Variable Linear Regression Model	155
6.1 Regression through the Origin	155
r^2 for Regression-through-Origin Model	159
An Illustrative Example: The Characteristic Line of Portfolio Theory	159
6.2 Scaling and Units of Measurement	161
A Numerical Example: The Relationship between GPD and GNP, United States, 1974–1983	163
A Word about Interpretation	164
6.3 Functional Forms of Regression Models	165
6.4 How to Measure Elasticity: The Log-Linear Model	165
An Illustrative Example: The Coffee Demand Function Revisited	167
6.5 Semilog Models: Log-Lin and Lin-Log Models	169
How to Measure the Growth Rate: The Log-Lin Model	169
The Lin-Log Model	172
6.6 Reciprocal Models	173
An Illustrative Example: The Phillips Curve for the United Kingdom, 1950–1966	176
6.7 Summary of Functional Forms	176
6.8 A Note on the Nature of the Stochastic Error Term: Additive versus Multiplicative Stochastic Error Term	178
6.9 Summary and Conclusions	179
Exercises	180
Questions	180
Problems	183
Appendix 6A	186
6A.1 Derivation of Least-Squares Estimators for Regression through the Origin	186
6A.2 SAS Output of the Characteristic Line (6.1.12)	189
6A.3 SAS Output of the United Kingdom Phillips Curve Regression (6.6.2)	190
7 Multiple Regression Analysis: The Problem of Estimation	191
7.1 The Three-Variable Model: Notation and Assumptions	192
7.2 Interpretation of Multiple Regression Equation	194
7.3 The Meaning of Partial Regression Coefficients	195
7.4 OLS and ML Estimation of the Partial Regression Coefficients	197
OLS Estimators	197
Variances and Standard Errors of OLS Estimators	198

	Properties of OLS Estimators	199
	Maximum Likelihood Estimators	201
7.5	The Multiple Coefficient of Determination R^2 and the Multiple Coefficient of Correlation R	201
7.6	Example 7.1: The Expectations-Augmented Phillips Curve for the United States, 1970–1982	203
7.7	Simple Regression in the Context of Multiple Regression: Introduction to Specification Bias	204
7.8	R^2 and the Adjusted R^2	207
	Comparing Two R^2 Values	209
	Example 7.2: Coffee Demand Function Revisited	210
	The “Game” of Maximizing \hat{R}^2	211
7.9	Partial Correlation Coefficients	211
	Explanation of Simple and Partial Correlation Coefficients	211
	Interpretation of Simple and Partial Correlation Coefficients	213
7.10	Example 7.3: The Cobb-Douglas Production Function: More on Functional Form	214
7.11	Polynomial Regression Models	217
	Example 7.4: Estimating the Total Cost Function	218
	Empirical Results	220
7.12	Summary and Conclusions	221
	Exercises	221
	Questions	221
	Problems	224
	Appendix 7A	231
7A.1	Derivation of OLS Estimators Given in Equations (7.4.3) and (7.4.5)	231
7A.2	Equality between a_1 of (7.3.5) and β_2 of (7.4.7)	232
7A.3	Derivation of Equation (7.4.19)	232
7A.4	Maximum Likelihood Estimation of the Multiple Regression Model	233
7A.5	The Proof that $E(b_{12}) = \beta_2 + \beta_3 b_{32}$ (Equation 7.7.4)	234
7A.6	SAS Output of the Expectations-Augmented Phillips Curve (7.6.2)	236
7A.7	SAS Output of the Cobb-Douglas Production Function (7.10.4)	237
8	Multiple Regression Analysis: The Problem of Inference	238
8.1	The Normality Assumption Once Again	238
8.2	Example 8.1: U.S. Personal Consumption and Personal Disposal Income Relation, 1956–1970	239
8.3	Hypothesis Testing in Multiple Regression: General Comments	242
8.4	Hypothesis Testing about Individual Partial Regression Coefficients	242
8.5	Testing the Overall Significance of the Sample Regression	244

	The Analysis of Variance Approach to Testing the Overall Significance of an Observed Multiple Regression: The F Test	245
	An Important Relationship between R^2 and F	248
	The “Incremental,” or “Marginal,” Contribution of an Explanatory Variable	250
8.6	Testing the Equality of Two Regression Coefficients	254
	Example 8.2: The Cubic Cost Function Revisited	255
8.7	Restricted Least Squares: Testing Linear Equality Restrictions	256
	The t Test Approach	256
	The F Test Approach: Restricted Least Squares	257
	Example 8.3: The Cobb-Douglas Production Function for Taiwanese Agricultural Sector, 1958–1972	259
	General F Testing	260
8.8	Comparing Two Regressions: Testing for Structural Stability of Regression Models	262
8.9	Testing the Functional Form of Regression: Choosing between Linear and Log-Linear Regression Models	265
	Example 8.5: The Demand for Roses	266
8.10	Prediction with Multiple Regression	267
8.11	The Troika of Hypothesis Tests: The Likelihood Ratio (LR), Wald (W), and Lagrange Multiplier (LM) Tests	268
8.12	Summary and Conclusions	269
	The Road Ahead	269
	Exercises	270
	Questions	270
	Problems	273
	Appendix 8A	280
	Likelihood Ratio (LR) Test	280

9 The Matrix Approach to Linear Regression Model 282

9.1	The k -Variable Linear Regression Model	282
9.2	Assumptions of the Classical Linear Regression Model in Matrix Notation	284
9.3	OLS Estimation	287
	An Illustration	289
	Variance-Covariance Matrix of $\hat{\beta}$	290
	Properties of OLS Vector $\hat{\beta}$	291
9.4	The Coefficient of Determination R^2 in Matrix Notation	292
9.5	The Correlation Matrix	292
9.6	Hypothesis Testing about Individual Regression Coefficients in Matrix Notation	293
9.7	Testing the Overall Significance of Regression: Analysis of Variance in Matrix Notation	294
9.8	Testing Linear Restrictions: General F Testing Using Matrix Notation	295
9.9	Prediction Using Multiple Regression: Matrix Formulation	296
	Mean Prediction	296