
CONTENTS

About the Authors	iii
Preface	xv
1 Relationships between Two Variables	1
1.1 Examples of Bivariate Relationships	1
1.1.1 Bivariate Frequency Distributions	5
1.2 The Correlation Coefficient	6
1.2.1 The Correlation Coefficient for a Bivariate Frequency Distribution	8
1.2.2 The Limits of r	9
1.2.3 Nonsense Correlations and Other Matters	9
1.2.4 A Case Study	11
1.3 Probability Models for Two Variables	12
1.3.1 Discrete Bivariate Probability Distribution	12
1.3.2 The Bivariate Normal Distribution	14
1.4 The Two-Variable Linear Regression Model	15
1.4.1 A Conditional Model	16
1.4.2 Estimates and Estimators	18
1.4.3 Least-Squares Estimators	19
1.4.4 Decomposition of the Sum of Squares	21
1.4.5 A Numerical Example	22
1.5 Inference in the Two-Variable, Least-Squares Model	23
1.5.1 Properties of LS Estimators	23
1.5.2 Gauss–Markov Theorem	25
1.5.3 Inference Procedures	25
1.5.4 Numerical Example (Continued from Section 1.4.5)	28
1.6 Analysis of Variance in the Two-Variable Regression Model	29
1.7 Prediction in the Two-Variable Regression Model	31
1.8 Gasoline Consumption: A Preliminary Analysis	33
Appendix	
1.1 To prove $\text{var}(b) = \sigma^2 / \sum x^2$	35
1.2 To derive the mean and variance of the sampling distribution of a	35

1.3	To derive $\text{cov}(a, b)$	36
1.4	Gauss–Markov theorem	36
1.5	To derive $\text{var}(e_0)$	37
	Problems	37
2	Further Aspects of Two-Variable Relationships	41
2.1	Time as a Regressor	42
2.1.1	Constant Growth Curves	43
2.1.2	Numerical Example	43
2.2	Transformations of Variables	44
2.2.1	Log-Log Transformations	45
2.2.2	Semilog Transformations	46
2.2.3	Reciprocal Transformations	47
2.3	An Empirical Example of a Nonlinear Relation: U.S. Inflation and Unemployment	49
2.4	Lagged Dependent Variable as Regressor	52
2.4.1	An Introduction to Asymptotics	53
2.4.2	Convergence in Probability	54
2.4.3	Convergence in Distribution	55
2.4.4	The Autoregressive Equation	56
2.5	Stationary and Nonstationary Series	57
2.5.1	Unit Root	59
2.5.2	Numerical Illustration	59
2.6	Maximum Likelihood Estimation of the Autoregressive Equation	61
2.6.1	Maximum Likelihood Estimators	61
2.6.2	Properties of Maximum Likelihood Estimators	63
	Appendix	
2.1	Change of variables in density functions	64
2.2	Maximum likelihood estimators for the AR(1) model	65
	Problems	66
3	The k-Variable Linear Equation	69
3.1	Matrix Formulation of the k -Variable Model	70
3.1.1	The Algebra of Least Squares	70
3.1.2	Decomposition of the Sum of Squares	72
3.1.3	Equation in Deviation Form	73
3.2	Partial Correlation Coefficients	76
3.2.1	Sequential Buildup of the Explained Sum of Squares	78
3.2.2	Partial Correlation Coefficients and Multiple Regression Coefficients	81
3.2.3	General Treatment of Partial Correlation and Multiple Regression Coefficients	82
3.3	The Geometry of Least Squares	83
3.4	Inference in the k -Variable Equation	86
3.4.1	Assumptions	86
3.4.2	Mean and Variance of b	87

3.4.3	Estimation of σ^2	89
3.4.4	Gauss–Markov Theorem	89
3.4.5	Testing Linear Hypotheses about β	90
3.4.6	Restricted and Unrestricted Regressions	95
3.4.7	Fitting the Restricted Regression	96
3.5	Prediction	99
Appendix		
3.1	To prove $r_{12.3} = (r_{12} - r_{13}r_{23})/\sqrt{1 - r_{13}^2}\sqrt{1 - r_{23}^2}$	100
3.2	Solving for a single regression coefficient in a multiple regression	101
3.3	To show that minimizing $\mathbf{a}'\mathbf{a}$ subject to $\mathbf{X}'\mathbf{a} = \mathbf{c}$ gives $\mathbf{a} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{c}$	103
3.4	Derivation of the restricted estimator \mathbf{b}_*	103
Problems		104
4	Some Tests of the k-Variable Linear Equation for Specification Error	109
4.1	Specification Error	109
4.1.1	Possible Problems with u	110
4.1.2	Possible Problems with X	110
4.1.3	Possible Problems with β	111
4.2	Model Evaluation and Diagnostic Tests	112
4.3	Tests of Parameter Constancy	113
4.3.1	The Chow Forecast Test	113
4.3.2	The Hansen Test	116
4.3.3	Tests Based on Recursive Estimation	117
4.3.4	One-Step Ahead Prediction Errors	118
4.3.5	CUSUM and CUSUMSQ Tests	119
4.3.6	A More General Test of Specification Error: The Ramsey RESET Test	121
4.4	A Numerical Illustration	121
4.5	Tests of Structural Change	126
4.5.1	Test of One Structural Change	127
4.5.2	Tests of Slope Coefficients	128
4.5.3	Tests of Intercepts	128
4.5.4	Summary	129
4.5.5	A Numerical Example	130
4.5.6	Extensions	132
4.6	Dummy Variables	133
4.6.1	Introduction	133
4.6.2	Seasonal Dummies	134
4.6.3	Qualitative Variables	135
4.6.4	Two or More Sets of Dummy Variables	137
4.6.5	A Numerical Example	137
Appendix		
4.1	To show $\text{var}(d) = \sigma^2 [\mathbf{I}_{n_2} + \mathbf{X}_2(\mathbf{X}'_1\mathbf{X}_1)^{-1}\mathbf{X}'_2]$	138
Problems		139

5	Maximum Likelihood (ML), Generalized Least Squares (GLS), and Instrumental Variable (IV) Estimators	142
5.1	Maximum Likelihood Estimators	142
5.1.1	Properties of Maximum Likelihood Estimators	143
5.2	ML Estimation of the Linear Model	145
5.3	Likelihood Ratio, Wald, and Lagrange Multiplier Tests	147
5.3.1	Likelihood Ratio (LR) Tests	147
5.3.2	The Wald (W) Test	148
5.3.3	Lagrange Multiplier (LM) Test	149
5.4	ML Estimation of the Linear Model with Nonspherical Disturbances	151
5.4.1	Generalized Least Squares	152
5.5	Instrumental Variable (IV) Estimators	153
5.5.1	Special Case	156
5.5.2	Two-Stage Least Squares (2SLS)	157
5.5.3	Choice of Instruments	157
5.5.4	Tests of Linear Restrictions	158
	Appendix	
5.1	Change of variables in density functions	158
5.2	Centered and uncentered R^2	159
5.3	To show that $e_*'X(X'X)^{-1}X'e_* = e_*'e_* - e'e$	160
	Problems	161
6	Heteroscedasticity and Autocorrelation	162
6.1	Properties of OLS Estimators	163
6.2	Tests for Heteroscedasticity	166
6.2.1	The White Test	166
6.2.2	The Breusch-Pagan/Godfrey Test	167
6.2.3	The Goldfeld-Quandt Test	168
6.2.4	Extensions of the Goldfeld-Quandt Test	168
6.3	Estimation Under Heteroscedasticity	170
6.3.1	Estimation with Grouped Data	171
6.3.2	Estimation of the Heteroscedasticity Relation	171
6.4	Autocorrelated Disturbances	174
6.4.1	Forms of Autocorrelation: Autoregressive and Moving Average Schemes	175
6.4.2	Reasons for Autocorrelated Disturbances	176
6.5	OLS and Autocorrelated Disturbances	176
6.6	Testing for Autocorrelated Disturbances	178
6.6.1	Durbin-Watson Test	179
6.6.2	The Wallis Test for Fourth-Order Autocorrelation	182
6.6.3	Durbin Tests for a Regression Containing Lagged Values of the Dependent Variable	182
6.6.4	Breusch-Godfrey Test	185
6.6.5	Box-Pierce-Ljung Statistic	187
6.7	Estimation of Relationships with Autocorrelated Disturbances	188

6.8	Forecasting with Autocorrelated Disturbances	192
6.9	Autoregressive Conditional Heteroscedasticity (ARCH)	195
	Appendix	
6.1	LM test for multiplicative heteroscedasticity	198
6.2	LR test for groupwise homoscedasticity	200
6.3	Properties of the ARCH(1) process	201
	Problems	202
7	Univariate Time Series Modeling	204
7.1	A Rationale for Univariate Analysis	205
7.1.1	The Lag Operator	206
7.1.2	ARMA Modeling	207
7.2	Properties of AR, MA, and ARMA Processes	207
7.2.1	AR(1) Process	207
7.2.2	AR(2) Process	209
7.2.3	MA Processes	213
7.2.4	ARMA Processes	214
7.3	Testing for Stationarity	215
7.3.1	Graphical Inspection	216
7.3.2	Integrated Series	220
7.3.3	Trend Stationary (TS) and Difference Stationary (DS) Series	220
7.3.4	Unit Root Tests	223
7.3.5	Numerical Example	226
7.4	Identification, Estimation, and Testing of ARIMA Models	228
7.4.1	Identification	228
7.4.2	Estimation	229
7.4.3	Diagnostic Testing	230
7.5	Forecasting	231
7.5.1	MA(1) Process	232
7.5.2	ARMA(1,1) Process	233
7.5.3	ARIMA(1,1,0) Process	234
7.6	Seasonality	235
7.7	A Numerical Example: Monthly Housing Starts	236
	Problems	243
8	Autoregressive Distributed Lag Relationships	244
8.1	Autoregressive Distributed Lag Relations	244
8.1.1	A Constant Elasticity Relation	245
8.1.2	Reparameterization	245
8.1.3	Dynamic Equilibrium	246
8.1.4	Unit Elasticity	246
8.1.5	Generalizations	247
8.2	Specification and Testing	248
8.2.1	General to Simple and Vice Versa	248
8.2.2	Estimation and Testing	250
8.2.3	Exogeneity	253

8.2.4	Exogeneity Tests	256
8.2.5	The Wu-Hausman Test	257
8.3	Nonstationary Regressors	259
8.4	A Numerical Example	265
8.4.1	Stationarity	266
8.4.2	Cointegration	266
8.4.3	A Respecified Relationship	270
8.4.4	A General ADL Relation	271
8.4.5	A Reparameterization	275
8.5	Nonnested Models	280
Appendix		
8.1	Nonsingular linear transformations of the variables in an equation	282
8.2	To establish the equality of the test statistics in Eqs. (8.37) and (8.41)	284
Problems		285
9	Multiple Equation Models	287
9.1	Vector Autoregressions (VARs)	287
9.1.1	A Simple VAR	287
9.1.2	A Three-Variable VAR	292
9.1.3	Higher-Order Systems	294
9.2	Estimation of VARs	295
9.2.1	Testing the Order of the VAR	296
9.2.2	Testing for Granger Causality	296
9.2.3	Forecasting, Impulse Response Functions, and Variance Decomposition	297
9.2.4	Impulse Response Functions	298
9.2.5	Orthogonal Innovations	299
9.2.6	Variance Decomposition	301
9.3	Vector Error Correction Models	301
9.3.1	Testing for Cointegration Rank	302
9.3.2	Estimation of Cointegrating Vectors	303
9.3.3	Estimation of a Vector Error Correction Model	305
9.4	Simultaneous Structural Equation Models	305
9.5	Identification Conditions	309
9.6	Estimation of Structural Equations	314
9.6.1	Nonstationary Variables	317
9.6.2	System Methods of Estimation	317
Appendix		
9.1	Seemingly Unrelated Regressions (SUR)	318
9.2	Higher-order VARs	320
9.2.1	A VAR(1) Process	320
9.2.2	A VAR(2) Process	321
Problems		322

10	Generalized Method of Moments	327
10.1	The Method of Moments	328
10.2	OLS as a Moment Problem	329
10.3	Instrumental Variables as a Moment Problem	330
10.4	GMM and the Orthogonality Condition	333
10.5	Distribution of the GMM estimator	335
10.6	Applications	336
10.6.1	Two-Stage Least Squares, and Tests of Overidentifying Restrictions	336
10.6.2	Wu-Hausman Tests Revisited	338
10.6.3	Maximum Likelihood	342
10.6.4	Euler Equations	343
10.7	Readings	344
	Problems	345
11	A Smorgasbord of Computationally Intensive Methods	348
11.1	An Introduction to Monte Carlo Methods	348
11.1.1	Some Guidelines for Monte Carlo Experiments	349
11.1.2	An Example	350
11.1.3	Generating Pseudorandom Numbers	352
11.1.4	Presenting the Results	354
11.2	Monte Carlo Methods and Permutation Tests	359
11.3	The Bootstrap	362
11.3.1	The Standard Error of the Median	362
11.3.2	An Example	363
11.3.3	The Parametric Bootstrap	365
11.3.4	Residual Resampling: Time Series and Forecasting	366
11.3.5	Data Resampling: Cross-Section Data	369
11.3.6	Some Remarks on Econometric Applications of the Bootstrap	369
11.4	Nonparametric Density Estimation	370
11.4.1	Some General Remarks on Nonparametric Density Estimation	375
11.4.2	An Application: The Wage Effects of Unions	376
11.5	Nonparametric Regression	379
11.5.1	Extension: The Partially Linear Regression Model	383
11.6	References	385
	Problems	385
12	Panel Data	388
12.1	Sources and Types of Panel Data	389
12.2	The Simplest Case—The Pooled Estimator	390
12.3	Two Extensions to the Simple Model	390
12.4	The Random Effects Model	391
12.5	Random Effects as a Combination of Within and Between Estimators	392

12.6	The Fixed Effects Model in the Two-Period Case	395
12.7	The Fixed Effects Model with More Than Two Time Periods	397
12.8	The Perils of Fixed Effects Estimation	399
12.8.1	Example 1: Measurement Error in X	399
12.8.2	Example 2: Endogenous X	402
12.9	Fixed Effects or Random Effects?	403
12.10	A Wu-Hausman Test	403
12.11	Other Specification Tests and an Introduction to Chamberlain's Approach	404
12.11.1	Formalizing the Restrictions	406
12.11.2	Fixed Effects in the General Model	407
12.11.3	Testing the Restrictions	407
12.12	Readings	408
	Problems	409
13	Discrete and Limited Dependent Variable Models	412
13.1	Types of Discrete Choice Models	412
13.2	The Linear Probability Model	414
13.3	Example: A Simple Descriptive Model of Union Participation	415
13.4	Formulating a Probability Model	418
13.5	The Probit	419
13.6	The Logit	424
13.7	Misspecification in Binary Dependent Models	426
13.7.1	Heteroscedasticity	426
13.7.2	Misspecification in the Probit and Logit	427
13.7.3	Functional Form: What Is the Right Model to Use?	430
13.8	Extensions to the Basic Model: Grouped Data	432
13.8.1	Maximum Likelihood Methods	432
13.8.2	Minimum χ^2 Methods	433
13.9	Ordered Probit	434
13.10	Tobit Models	436
13.10.1	The Tobit as an Extension of the Probit	436
13.10.2	Why Not Ignore "The Problem"?	439
13.10.3	Heteroscedasticity and the Tobit	440
13.11	Two Possible Solutions	441
13.11.1	Symmetrically Trimmed Least Squares	442
13.11.2	Censored Least Absolute Deviations (CLAD) Estimator	444
13.12	Treatment Effects and Two-Step Methods	446
13.12.1	The Simple Heckman Correction	447
13.12.2	Some Cautionary Remarks about Selectivity Bias	449
13.12.3	The Tobit as a Special Case	450
13.13	Readings	452
	Problems	452

Appendix A

A.1	Vectors	455
A.1.1	Multiplication by a Scalar	456
A.1.2	Addition and Subtraction	456
A.1.3	Linear Combinations	456
A.1.4	Some Geometry	456
A.1.5	Vector Multiplication	458
A.1.6	Equality of Vectors	458
A.2	Matrices	459
A.2.1	Matrix Multiplication	459
A.2.2	The Transpose of a Product	460
A.2.3	Some Important Square Matrices	461
A.2.4	Partitioned Matrices	463
A.2.5	Matrix Differentiation	464
A.2.6	Solution of Equations	465
A.2.7	The Inverse Matrix	466
A.2.8	The Rank of a Matrix	467
A.2.9	Some Properties of Determinants	471
A.2.10	Properties of Inverse Matrices	472
A.2.11	More on Rank and the Solution of Equations	473
A.2.12	Eigenvalues and Eigenvectors	476
A.2.13	Properties of Eigenvalues and Eigenvectors	478
A.2.14	Quadratic Forms and Positive Definite Matrices	483

Appendix B

B.1	Random Variables and Probability Distributions	485
B.2	The Univariate Normal Probability Distribution	486
B.3	Bivariate Distributions	487
B.4	Relations between the Normal, χ^2 , t , and F Distributions	489
B.5	Expectations in Bivariate Distributions	490
B.6	Multivariate Densities	490
B.7	Multivariate Normal pdf	492
B.8	Distributions of Quadratic Forms	493
B.9	Independence of Quadratic Forms	495
B.10	Independence of a Quadratic Form and a Linear Function	496

Appendix C	497
------------	-----

Appendix D	499
------------	-----

Index	521
-------	-----