

Microeconometrics Using Stata

Volume II: Nonlinear Models and Causal Inference Methods

Second Edition

A. COLIN CAMERON

Department of Economics

University of California, Davis, CA

and

School of Economics

University of Sydney, Sydney, Australia

PRAVIN K. TRIVEDI

School of Economics

University of Queensland, Brisbane, Australia

and

Department of Economics

Indiana University, Bloomington, IN



STATA Press

A Stata Press Publication

StataCorp LLC

College Station, Texas

Contents

	List of tables	xiii
	List of figures	xv
16	Nonlinear optimization methods	819
16.1	Introduction	819
16.2	Newton–Raphson method	819
16.3	Gradient methods	824
16.4	Overview of ml, moptimize(), and optimize()	829
16.5	The ml command: lf method	831
16.6	Checking the program	837
16.7	The ml command: lf0–lf2, d0–d2, and gf0 methods	844
16.8	Nonlinear instrumental-variables (GMM) example	851
16.9	Additional resources	854
16.10	Exercises	854
17	Binary outcome models	857
17.1	Introduction	857
17.2	Some parametric models	858
17.3	Estimation	860
17.4	Example	862
17.5	Goodness of fit and prediction	869
17.6	Marginal effects	877
17.7	Clustered data	880
17.8	Additional models	881
17.9	Endogenous regressors	887
17.10	Grouped and fractional data	895

17.11	Additional resources	898
17.12	Exercises	898
18	Multinomial models	901
18.1	Introduction	901
18.2	Multinomial models overview	901
18.3	Multinomial example: Choice of fishing mode	905
18.4	Multinomial logit model	908
18.5	Alternative-specific conditional logit model	914
18.6	Nested logit model	922
18.7	Multinomial probit model	929
18.8	Alternative-specific random-parameters logit	934
18.9	Ordered outcome models	938
18.10	Clustered data	942
18.11	Multivariate outcomes	943
18.12	Additional resources	946
18.13	Exercises	946
19	Tobit and selection models	949
19.1	Introduction	949
19.2	Tobit model	950
19.3	Tobit model example	953
19.4	Tobit for lognormal data	961
19.5	Two-part model in logs	970
19.6	Selection models	974
19.7	Nonnormal models of selection	982
19.8	Prediction from models with outcome in logs	986
19.9	Endogenous regressors	989
19.10	Missing data	991
19.11	Panel attrition	995
19.12	Additional resources	1019
19.13	Exercises	1019

20	Count-data models	1021
20.1	Introduction	1021
20.2	Modeling strategies for count data	1022
20.3	Poisson and negative binomial models	1026
20.4	Hurdle model	1044
20.5	Finite-mixture models	1050
20.6	Zero-inflated models	1069
20.7	Endogenous regressors	1079
20.8	Clustered data	1089
20.9	Quantile regression for count data	1090
20.10	Additional resources	1096
20.11	Exercises	1096
21	Survival analysis for duration data	1099
21.1	Introduction	1099
21.2	Data and data summary	1100
21.3	Survivor and hazard functions	1104
21.4	Semiparametric regression model	1109
21.5	Fully parametric regression models	1118
21.6	Multiple-records data	1129
21.7	Discrete-time hazards logit model	1132
21.8	Time-varying regressors	1135
21.9	Clustered data	1136
21.10	Additional resources	1137
21.11	Exercises	1137
22	Nonlinear panel models	1139
22.1	Introduction	1139
22.2	Nonlinear panel-data overview	1139
22.3	Nonlinear panel-data example	1145
22.4	Binary outcome and ordered outcome models	1148
22.5	Tobit and interval-data models	1167

22.6	Count-data models	1172
22.7	Panel quantile regression	1184
22.8	Endogenous regressors in nonlinear panel models	1187
22.9	Additional resources	1188
22.10	Exercises	1188
23	Parametric models for heterogeneity and endogeneity	1191
23.1	Introduction	1191
23.2	Finite mixtures and unobserved heterogeneity	1192
23.3	Empirical examples of FMMs	1195
23.4	Nonlinear mixed-effects models	1224
23.5	Linear structural equation models	1231
23.6	Generalized structural equation models	1251
23.7	ERM commands for endogeneity and selection	1261
23.8	Additional resources	1266
23.9	Exercises	1266
24	Randomized control trials and exogenous treatment effects	1269
24.1	Introduction	1269
24.2	Potential outcomes	1271
24.3	Randomized control trials	1272
24.4	Regression in an RCT	1282
24.5	Treatment evaluation with exogenous treatment	1290
24.6	Treatment evaluation methods and estimators	1292
24.7	Stata commands for treatment evaluation	1302
24.8	Oregon Health Insurance Experiment example	1305
24.9	Treatment-effect estimates using the OHIE data	1312
24.10	Multilevel treatment effects	1323
24.11	Conditional quantile TEs	1332
24.12	Additional resources	1334
24.13	Exercises	1335

25 Endogenous treatment effects 1337

- 25.1 Introduction 1337
- 25.2 Parametric methods for endogenous treatment 1338
- 25.3 ERM commands for endogenous treatment 1341
- 25.4 ET commands for binary endogenous treatment 1348
- 25.5 The LATE estimator for heterogeneous effects 1356
- 25.6 Difference-in-differences and synthetic control 1363
- 25.7 Regression discontinuity design 1369
- 25.8 Conditional quantile regression with endogenous regressors 1388
- 25.9 Unconditional quantiles 1394
- 25.10 Additional resources 1401
- 25.11 Exercises 1402

26 Spatial regression 1405

- 26.1 Introduction 1405
- 26.2 Overview of spatial regression models 1406
- 26.3 Geospatial data 1407
- 26.4 The spatial weighting matrix 1411
- 26.5 OLS regression and test for spatial correlation 1413
- 26.6 Spatial dependence in the error 1414
- 26.7 Spatial autocorrelation regression models 1417
- 26.8 Spatial instrumental variables 1427
- 26.9 Spatial panel-data models 1428
- 26.10 Additional resources 1429
- 26.11 Exercises 1430

27 Semiparametric regression 1433

- 27.1 Introduction 1433
- 27.2 Kernel regression 1434
- 27.3 Series regression 1438
- 27.4 Nonparametric single regressor example 1440
- 27.5 Nonparametric multiple regressor example 1450

27.6	Partial linear model	1453
27.7	Single-index model	1456
27.8	Generalized additive models	1458
27.9	Additional resources	1461
27.10	Exercises	1462
28	Machine learning for prediction and inference	1465
28.1	Introduction	1465
28.2	Measuring the predictive ability of a model	1466
28.3	Shrinkage estimators	1477
28.4	Prediction using lasso, ridge, and elasticnet	1482
28.5	Dimension reduction	1493
28.6	Machine learning methods for prediction	1496
28.7	Prediction application	1501
28.8	Machine learning for inference in partial linear model	1507
28.9	Machine learning for inference in other models	1516
28.10	Additional resources	1523
28.11	Exercises	1524
29	Bayesian methods: Basics	1527
29.1	Introduction	1527
29.2	Bayesian introductory example	1528
29.3	Bayesian methods overview	1532
29.4	An i.i.d. example	1538
29.5	Linear regression	1549
29.6	A linear regression example	1552
29.7	Modifying the MH algorithm	1560
29.8	RE model	1562
29.9	Bayesian model selection	1567
29.10	Bayesian prediction	1569
29.11	Probit example	1572

29.12	Additional resources	1576
29.13	Exercises	1576
30	Bayesian methods: Markov chain Monte Carlo algorithms	1579
30.1	Introduction	1579
30.2	User-provided log likelihood	1579
30.3	MH algorithm in Mata	1584
30.4	Data augmentation and the Gibbs sampler in Mata	1589
30.5	Multiple imputation	1595
30.6	Multiple-imputation example	1599
30.7	Additional resources	1608
30.8	Exercises	1608
	Glossary of abbreviations	1611
	References	1617
	Author index	1635
	Subject index	1641