

Econ 276: Econometrics of Limited Dependent Variables

Stanford University
Spring 2009

Lecture: MW 9:15-10:50, Landau 206

Professor: Matt Harding

Office: Landau 238

Office Hours: M 3:30-5:30pm

Email: mch@stanford.edu

Assistant: Patricia Luna, luna@stanford.edu

Course Objectives

This course aims to provide an overview of a variety of topics in the econometrics of limited dependent variables with a particular focus on estimation and practical implementation. First, the course will look at a series of general econometric approaches relating to panel data, instrumental variables and quantile regression. Second, the course will review classical approaches dealing with binary variables and censoring/truncation. Third, the course will provide a brief overview of more advanced topics such as set valued estimation, estimation of social networks, Bayesian methods and computationally intensive methods.

The focus of the course is on providing a toolbox of estimation strategies. Less emphasis is being placed on rigorous proofs and derivations while emphasizing applications and actual implementation on a computer. Additionally, it aims to prepare students for independent research in applied econometrics and encourage students to take the first steps towards initiating an original applied project.

Pre-requisites

Basic knowledge of regression analysis and first year graduate level econometrics or statistics is required. Familiarity with a statistical package such as STATA or R and some basic programming knowledge in a modern language such as Matlab will be helpful.

Course Requirements

Course credit is given for class presentations (20%) and a final project (80%).

Students are required to work and develop an applied project during the course. There will be two project based presentations (one at mid-term and one at the end of the term).

The final project can consist of either of the following:

- A short applied research note on a topic in applied economics similar to a note published in *Economics Letters*
- A replication of some of the main results in an existing paper in applied economics. The student is responsible for identifying the paper, obtaining the data and code and replicating some of the main tables in the paper

The final project will consist of a short written report of no more than 5-10 pages but it must be empirical in nature and show a careful analysis of an economic dataset.

Course Outline

Week	Dates	Description
1	Apr 1	Panel Data: Individual Effects, Hausman-Taylor, Robust Std Errors
2	Apr 6,8	Instrumental Variables; MLE, GMM, MD, Quantile Regression
3	Apr 13, 15	Logit, Probit, Tobit, Censoring/Truncation
4	Apr 20, 22	Duration Models; <i>Initial Project Presentations</i>
5	Apr 27, 29	Index Models; Bootstrap; Non-linear Panel Data Models
6	May 4,6	Discrete Choices
7	May 11, 3	Simulation Based Estimation
8	May 18, 20	Bayesian Modeling, MCMC
9	May 27	Factor Analysis/Random Matrix Theory; Estimation of Social Networks
10	June 1	Final Project Presentations
11	June 5	<i>Final Project Due</i>

Reading List

Textbooks

Angrist, J. and Pischke J-S (2009) *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton.

Cameron, C. and Trivedi, P. (2009) *Microeconometrics Using Stata*, Stata Press.

Gourieroux, C. (2000) *Econometrics of Qualitative Dependent Variables*, Cambridge

Train, K. (2003) *Discrete Choice Methods with Simulation*. Cambridge University Press.

Wooldridge, J. (2002) *Econometric Analysis of Cross-Section and Panel Data*, MIT.

More readings will be added by topic and the papers will be available on the course website as the term progresses.

1. Panel Data (Fixed Effects vs Random Effects, Hausman Test, Hausman-Taylor Method, Robust Standard Errors)

Chamberlain, G. (1984) "Panel Data" in Griliches, Z. and Intriligator, MD (Eds) *Handbook of Econometrics*, Vol. 2, Chapter 22, Elsevier.

Hausman, J. (1978) "Specification Tests in Econometrics", *Econometrica*, 46(6), 1273-1291.

Hausman, J. and Taylor, W. (1981) "Panel Data and Unobservable Individual Effects", *Econometrica* 49:6, 1377-1398.

Stock, J. and Watson, M. (2008) Heteroskedasticity-Robust Standard Errors for Fixed Effects Panel Data Regressions", *Econometrica* 76:1, 155-174.

2. Instrumental Variables (Structural vs Reduced Form; IV Estimators; Finite Sample Bias, k-class Estimators, Quantile IV)

Hahn, J. and Hausman, J. (2002) "A New Specification Test for the Validity of Instrumental Variables", *Econometrica* 70:1, 163-189.

Hahn, J. and Hausman, J. (2002) "Notes on Bias in Estimators for Simultaneous Equation Models", *Economics Letters* 75, 237-241.

Hahn, J. and Hausman, J. (2003) "Weak Instruments: Diagnosis and Cures in Empirical Economics", *American Economic Review*, 93:2, 118-125.

Hausman, J. (1983) "Specification and Estimation of Simultaneous Equation Models", " in Griliches, Z. and Intriligator, MD (Eds) *Handbook of Econometrics*, Vol. 1, Chapter 7, Elsevier.

Phillips, P. (2006) "A Remark on Bimodality and Weak Instrumentation in Structural Equation Estimation", *Econometric Theory* 22, 947-960.

3. Classical Estimation Techniques (MLE, GMM, MD); Quantile Regression (Robustness, Computational Implementation)

Amemiya, T. (1985) *Advanced Econometrics*. Cambridge, MA: Harvard University Press.

Chernozhukov, V. and Hansen, C. (2008) "Instrumental Variable Quantile Regression: A Robust Inference Approach", *Journal of Econometrics*, 142(1), 379-398.

Hansen, L. (1982) "Large Sample Properties of Generalized Method of Moments Estimators", *Econometrica* 50, 1029-1054.

Hansen, L. and Singleton, K. (1982) "Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models", *Econometrica* 50, 1269-1286.

Harding, M. and Lamarche, C. (2008) "A Quantile Regression Approach for Estimating Panel Data Models using Instrumental Variables", mimeo.

Koenker, R. (2006) "Quantile Regression in R: A Vignette", mimeo.

Koenker, R. and Bassett, G. (1978) "Regression Quantiles", *Econometrica*, 46:1, 33-50.

Newey, W. (1985) "Generalized Method of Moments Specification Testing", *Journal of Econometrics* 29, 229-256.

Newey, W. and McFadden, D. (1999) "Large Sample Estimation and Hypothesis Testing." In McFadden, D. and Engle, R. *Handbook of Econometrics*. Vol. 4., Chapter 26, Elsevier.

Powell, J. (1986) "Censored Regression Quantiles", *Journal of Econometrics*, 32, 143-155.

Reid, N. (2003) "Asymptotics and the Theory of Inference", *Annals of Statistics* 31:6

Wooldridge (2002) ch. 12-14

4. Binary Variables (Logit, Probit, Thick Tails and Rare Events, Count Data, Semiparametric Methods)

Cosslett, S. (1981) "Maximum Likelihood Estimator for Choice-Based Samples", *Econometrica* 49(5), 1289-1316.

Imbens, G. (1992) "An Efficient Method of Moments Estimator for Discrete Choice Models with Choice-Based Sampling", *Econometrica* 60(5), 1187-1214.

Hausman, J., Hall, B. and Griliches, Z. (1984) "Econometric Models for Count Data with an Application to Poisson Models", *Econometrica* 52(3), 701-720.

Koenker, R. and Yoon, J. (2007) "Parametric Links for Binary Choice Models: A Fisherian-Bayesian Colloquy", mimeo.

Manski, C. and Lerman, S. (1977) "The Estimation of Choice Probabilities from Choice Based Samples", *Econometrica* 45(8), 1977-1988.

Ruud, P. (1983) "Sufficient Conditions for the Consistency of Maximum Likelihood Estimation Despite Misspecification of Distribution in Multinomial Discrete Choice Models", *Econometrica* 51(1), 225-228.

Wooldridge, J. (2002) Ch. 15.1-15.7, 19

5. Truncation, Censoring, Sample Selection

Amemiya, T. (1973) "Regression Analysis When the Dependent Variable is Truncated Normal", *Econometrica* 41(6), 997-1016.

Heckman, J. (1979) "Sample Selection Bias as a Specification Error", *Econometrica* 47(1), 153-161.

Winkelmann, R. (1998) "Count Data Models with Selectivity", *Econometric Reviews* 17(4), 339-359.

Wooldridge, J. (2002) Ch 16, 17

6. Duration Models (Proportional Hazard, Unobserved Heterogeneity, Semiparametric Models)

Chen, S. (2002) "Rank Estimation of Transformation Models", *Econometrica* 70(4), 1683-1697.

Deng, Y., Quigley, J. and van Order, R (2000) "Mortgage Terminations, Heterogeneity and the Exercise of Mortgage Options", *Econometrica* 68(2), 275-307.

Elbers, C. and Ridder, G. (1982) "True and Spurious Duration Dependence: The Identifiability of the Proportional Hazard Model", *Review of Economic Studies* 49(3), 403-409.

Han, A. and Hausman, J. (1990) "Flexible Parametric Estimation of Duration and Competing Risk Models", *Journal of Applied Econometrics* 5(1), 1-28.

Horowitz, J. (1999) "Semiparametric Estimation of a Proportional Hazard Model with Unobserved Heterogeneity", *Econometrica* 67(5), 1001-1028.

Meyer, B. (1990) "Unemployment Insurance and Unemployment Spells", *Econometrica* 58(4), 757-782.

Van den Berg, G. (2000) "Duration Models: Specification, Identification, and Multiple Durations", in Heckman, J. and Leamer, E. eds *Handbook of Econometrics Vol. V*, Elsevier.

Wooldridge, J. (2002) Ch 20

7. Bootstrap

Horowitz, J. (2001) "The Bootstrap" in Heckman, J. and Leamer, E. eds. *Handbook of Econometrics Vol. V.*, chapter 52, Elsevier, 3159-3228.

8. Nonlinear Panel Data

Arellano, M. and Bonhomme, S. (2009) "Robust Priors in Nonlinear Panel Data Models", *Econometrica* 77(2), 489-536.

Arellano, M. and Honoré, B. (2001) "Panel Data Models: Some Recent Developments in Heckman, J. and Leamer, E. eds. *Handbook of Econometrics Vol. V.*, chapter 53, Elsevier, 3229-3296.

Fernandez-Val, I. (2009) "Fixed Effects Estimation of Structural Parameters and Marginal Effects in Panel Probit Models", *Journal of Econometrics*.

Hahn, J. (2001): "The Information Bound of a Dynamic Panel Logit Model with Fixed Effects", *Econometric Theory*, 17, 913-932.

Hahn, J. and Newey, W. "Jackknife and Analytical Bias Reduction for Nonlinear Panel Data Models." *Econometrica* 72, 1295-1319.

Honoré, B. and Kyriazidou, E. (2000): "Panel data discrete choice models with lagged dependent variables", *Econometrica* 68, 839-874.

Lancaster, T. (2000) "The Incidental Parameter Problem Since 1948", *Journal of Econometrics*, 391-413.

Wooldridge, J. (2002) Ch 13

9. Stated Choice Models; Simulation Methods

Barndorff-Nielsen, O. and D. R. Cox (1979) "Edgeworth and Saddle-point Approximations with Statistical Applications," *Journal of the Royal Statistical Society B*, 41, 179-312.

Cardell, S. and F. Dunbar (1980) "Measuring the Societal Impacts of Automobile Downsizing," *Transportation Research A*, 14, 423-434.

Harding, M. C., and J. A. Hausman (2007) "Using a Laplace Approximation to Estimate the Random Coefficients Logit Model by Nonlinear Least Squares," *International Economic Review*, 48(4), 1311-1328.

Hausman, J. A., and D. A. Wise (1978): "A Conditional Probit Model for Qualitative Discrete-Choice Decisions Recognizing Interdependence and Heterogeneous Preferences," *Econometrica*, 46(2), 403-426.

Hensher, D. and W. Greene (2003) "The Mixed Logit Model: The State of Practice," *Transportation*, 133-176.

Knittel, C. R., and K. Metaxoglou (2008) "Estimation of random coefficient demand models: challenges, difficulties and warnings," working paper series, NBER.

Lerman, S. and C. Manski (1981) "On the Use of Simulated Frequencies to Approximate Choice Probabilities," in *Structural Analysis of Discrete Data with Econometric*, ed. by C. Manski and D. McFadden, Cambridge: MIT Press, 305-320.

Manski, C. and D. McFadden, eds. (1981) *Structural Analysis of Discrete Data with Econometric Applications*, Cambridge: MIT Press.

McFadden, D. (1974) "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers in Econometrics*, ed. by A. Karlqvist, L. Lundqvist, F. Snickars, and J. Weibull, Amsterdam: Academic Press, 75-96.

McFadden, D. (1989) "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration," *Econometrica*, 995-1026.

McFadden, D. and K. Train (2000) "Mixed MNL Models for Discrete Response," *Journal of Applied Econometrics*, 447-470.

Pakes, A. and D. Pollard (1989) "Simulation and the Asymptotics of Optimization Estimators," *Econometrica*, 1027-1057.

Revelt, D. and K. Train (1998) "Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level," *Review of Economics and Statistics*, 647-657.

Small, K. A., C. Winston, and J. Yan (2005) "Uncovering the Distribution of Motorists' Preferences for Travel Time and Reliability," *Econometrica*, 1367-1382.

10. Bayesian Methods; MCMC; Dirichlet Processes

Antoniak, C. E. (1974) "Mixtures of Dirichlet Processes with Applications to Bayesian Nonparametric Problems," *The Annals of Statistics*, 1, 1152–1174.

Athey, S., and G. Imbens (2007) "Discrete Choice Models with Multiple Unobserved Choice Characteristics," working paper, Harvard University.

Burda, M., M. C. Harding, and J. A. Hausman (2008) "A Bayesian Mixed Logit-Probit Model for Multinomial Choice," *Journal of Econometrics*, 147(2), 232–246.

Burda, M. and M. C. Harding (2009) "Dynamic Panel Probit with Flexible Correlated Effects," mimeo.

Blackwell, D., and J. B. MacQueen (1973) "Ferguson Distribution via Polya Urn Schemes," *The Annals of Statistics*, 1, 353–355.

Chib, S., and B. Hamilton (2002) "Semiparametric bayes analysis of longitudinal data treatment models," *Journal of Econometrics*, 110, 67–89.

Dahl, D. B. (2005) "Sequentially-Allocated Merge-Split Sampler for Conjugate and Nonconjugate Dirichlet Process Mixture Models," under revision, *Journal of Computational and Graphical Statistics*.

Ferguson, T. S. (1973) "A Bayesian Analysis of some Nonparametric Problems," *The Annals of Statistics*, 1, 209–230.

Hirano, K. (2002) "Semiparametric Bayesian Inference in Autoregressive Panel Data Models," *Econometrica*, 70, 781–799.

Jain, S., and R. M. Neal (2007) "Splitting and merging components of a nonconjugate Dirichlet process mixture model," *Bayesian Analysis*, 2(3), 445–472.

Muller, P., and F. A. Quintana (2004) "Nonparametric Bayesian Data Analysis," *Statistical Science*, 19(1), 95–110.

Neal, R. (2000) "Markov Chain Sampling Methods for Dirichlet Process Mixture Models," *Journal of Computational and Graphical Statistics*, 9(2), 249–265.

Rossi, P. E., G. M. Allenby, and R. McCulloch (2005) *Bayesian Statistics and Marketing*. Wiley series in Probability and Statistics.

11. Factor Analysis; Random Matrix Theory

- Bai, J. (2003) "Inferential Theory for Factor Models of Large Dimensions," *Econometrica*, 71, 135–171.
- Bai, J., and S. Ng (2002) "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70(1), 191–221.
- Bai, Z. D., and J. W. Silverstein (1998) "No Eigenvalues Outside the Support of the Limiting Spectral Distribution of Large Dimensional Sample Covariance Matrices," *Annals of Probability*, 26(1), 316–345.
- Bai, Z. D., and J. W. Silverstein (2004) "CLT for Linear Spectral Statistics of Large-Dimensional Sample Covariance Matrices," *Annals of Probability*, 32(1A), 553–605.
- Bai, Z. D., and J. W. Silverstein (2005) "Exact Separation of Eigenvalues of Large Dimensional Covariance Matrices," mimeo.
- Chamberlain, G., and M. Rothschild (1983) "Arbitrage, Factor Structure and Mean-Variance Analysis in Large Asset Markets," *Econometrica*, 51(5), 1305–324.
- Edelman, A., and N. R. Rao (2005): "Random Matrix Theory," *Acta Numerica*, pp. 233–297.
- Harding (2008) "Explaining the Single Factor Bias of Arbitrage Pricing Models in Finite Samples," *Economics Letters*, 99(1).
- Harding (2009) "Structural Estimation of High-Dimensional Factor Models", mimeo.
- Jolliffe, I. T. (2002) *Principal Component Analysis*. Springer, New York.
- Stock, J., and M. Watson (2003) "Forecasting Output and Inflation: The Role of Asset Prices," *Journal of Economic Literature*, 41, 788–829.
- Stock, J., and M. Watson (2005) "Implications of Dynamic Factor Models for VAR Analysis," mimeo.
- Stock, J., and M. Watson (2006) "Forecasting with Many Predictors," in *Handbook of Economic Forecasting*, ed. by G. Elliott, C. J. Granger, and A. Timmermann. Elsevier, Amsterdam.

12. Estimation of Random Networks; p^* Models

- Anderson, C., S. Wasserman and B. Crouch. (1999) "A p^* Primer: Logit Models for Social Networks", *Social Networks* 21, 37-66.
- Handcock, M., Hunter, D. and S. Goodreau (2008) "Goodness of Fit of Social Network Models", *Journal of the American Statistical Association* 103:1, pp. 248-258.

Hunter, D. (2007) "Curved Exponential Family Models for Social Networks", *Social Networks* 29, 216-230.

Robins, G. (2007) "Recent Developments in Exponential Random Graph Models for Social Networks", *Social Networks* 29, 192-215.