

POLI 803: Longitudinal Data Analysis

Autumn 2005

Professor Christopher Zorn
Department of Political Science
University of South Carolina
zorn at sc.edu

Description

This course is designed as an introduction to statistics for longitudinal data – that is, data with repeated measurements on the same units over time. Such data have become increasingly widespread in political science, particularly in comparative politics and international relations. Longitudinal data offer both challenges and opportunities for the applied researcher. The current course is divided more or less in half. In the first half of the class, we’ll discuss and apply methods for data which varies both across units and over time; these include models for “panel” data, “time-series cross-sectional” data, and the like. Topics will include fixed- and random-effects models, GLS-based approaches to panel data, GEE models, random coefficient models and dynamic models with lagged dependent variables. Along the way, we’ll try, to the extent possible, to include methods for “ugly” (i.e., discrete) dependent variables. The second half of the course will focus on survival models (also known as event-history or duration models), which are used in analyzing data on time(s) to event(s).

Much of the material in this course is fairly technical. While I have tried to chose readings that present the models as clearly and with as little jargon as possible, most of the material will still require several readings to fully comprehend. While not a prerequisite, experience in statistics at the level of a linear algebra-based treatment of least-squares regression is highly advisable for those enrolled in this class. Students are also expected to have a nodding acquaintance with basic differential and integral calculus and distribution theory. Additionally, it is impossible to learn statistics by reading books or articles and attending lectures. Because of this incontrovertible fact, enrollees will be required to complete five lab exercises over the course of the semester, typically receiving the exercise in one week and turning it in the following week at the beginning of class. These exercises will be computer-based; the two preferred software packages for the class are **Stata 9.0** (<http://www.stata.com>) and **R 2.1.1** (<http://cran.r-project.org>), both of which will be supported by the instructor, but you are welcome to use any software you care to. The same goes for the final project, the details of which will be announced in class.

Required Text

Box-Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. New York: Cambridge University Press.

Grading

Grading will be based on several lab exercises and a final project, as follows:

- Lab exercises: Five worth 10 percent each.
- Final Project: 50 percent.

Part One: Panel and TSCS Data Analysis

Reference Works

- Arellano, Manuel. 2002. *Panel Data Econometrics*. Oxford: Oxford University Press.
- Baltagi, Badi. 2005. *Econometric Analysis of Panel Data*, 3rd Ed. New York: Wiley.
- Diggle, P., P. Heagerty, K-Y Liang, and S. Zeger. 2002. *Analysis of Longitudinal Data*, 2nd Ed. Oxford: Oxford University Press.
- Finkel, Stephen E. 1995. *Causal Analysis With Panel Data*. Thousand Oaks, CA: Sage Publications.
- Frees, Edward W. 2004. *Longitudinal and Panel Data: Analysis and Applications in the Social Sciences*. New York: Cambridge University Press.
- Hand, David and Martin Crowder. 1996. *Practical Longitudinal Data Analysis*. London: Chapman and Hall.
- Hsiao, Cheng. 2003. *The Analysis of Panel Data*, 2nd Ed. New York: Cambridge University Press.
- Lee, Myoung-Jae. 2002. *Panel Data Econometrics: Methods-of-Moments and Limited Dependent Variables*. Academic Press.
- Mátyás, László and Patrick Sevestre, eds. 1996. *The Econometrics of Panel Data: A Handbook of the Theory with Applications*. 2nd Revised Ed. Dordrecht: Kluwer Academic Publishers.
- Singer, Judith D. and John B. Willett. 2003. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. New York: Oxford University Press (also discusses survival analysis).
- Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*. Cambridge: MIT Press.

Some Useful Resources

- Neal Beck’s “Longitudinal Data Analysis” course page (<http://www.nyu.edu/classes/nbeck/longdata/longdata.html>).
- Oxford University’s ESRC Spring School on Panel Data Analysis (http://springschool.politics.ox.ac.uk/springschool/courses_instructors_2005.asp).
- Thomas Plümper’s webpage at University of Konstanz (<http://www.uni-konstanz.de/FuF/Verwiss/GSchneider/pluemper/>).
- Robert Yaffee’s “A Primer for Panel Data Analysis” (http://www.nyu.edu/its/pubs/connect/fall03/yaffee_primer.html).

Class Schedule

August 22: Class Introduction and Overview

- No readings assigned.

August 29: Introduction to Panel/TSCS Models – Data, Pooling, Software, and so forth.

Required readings:

- Bartels, Larry M. 1996. “Pooling Disparate Observations.” *American Journal of Political Science* 40(August):905–42.
- Hsiao, Cheng. 2003. *Analysis of Panel Data*, Chapter 1.
- Nuamah, Nicholas N. N. 1986. “Pooling Cross Section and Time Series Data.” *The Statistician* 35:345–51.

Recommended readings:

- Bartels, Larry. 2000. “Panel Effects in the American National Election Studies.” *Political Analysis* 8(Winter):1-20.
- Gurland, John, and Jayaram Sethuraman. 1995. “How Pooling Failure Data May Reverse Increasing Failure Rates.” *Journal of the American Statistical Association* 90(December):1416–23.

September 5: No Class.

September 12: Fixed- and Random-Effects Models.

Required readings:

- Hsiao, Cheng. 2003. *The Analysis of Panel Data*, Chapter 3.
- Stimson, James. 1985. "Regression in Space and Time: A Statistical Essay." *American Journal of Political Science* 29:914–47.
- Zorn, Christopher. 2001. "Estimating Between- and Within-Cluster Covariate Effects, with an Application to Models of International Disputes." *International Interactions* 27(4):433–45.

Recommended readings:

- Finkel, Steven E., and Edward N. Muller. 1998. "Rational Choice and the Dynamics of Political Action: Evaluating Alternative Models with Panel Data." *American Political Science Review* 92(March):37–50.
- Neuhaus, J. M., and J. D. Kalbfleisch. 1998. "Between- and Within-Cluster Covariate Effects in the Analysis of Clustered Data." *Biometrics* 54:638–45.

September 19: GLS–ARMA Models.

Required readings:

- Beck, Nathaniel, and Jonathan N. Katz. 1995. "What To Do (And Not To Do) With Time-Series Cross-Section Data." *American Political Science Review* 89(September):634–47.
- Beck, Nathaniel, and Jonathan N. Katz. 1996. "Nuisance vs. Substance: Specifying and Estimating Time-Series Cross-Section Models." *Political Analysis*. 6:1–36.
- Wilson, Sven E., and Daniel M. Butler. 2004. "A Lot More to Do: The Promise and Peril of Panel Data in Political Science." Working paper: Brigham Young University.

Recommended readings / Applications:

- Beck, Nathaniel. 2001. "Time-Series Cross-Section Data: What Have We Learned in the Past Few Years?" *Annual Review of Political Science*. 4:271–93.
- Blais, André, Donald Blake, and Stephane Dion. 1996. "Do Parties Make a Difference: A Reappraisal." *American Journal of Political Science* 40:514–20.
- Burkhart, Ross E. and Michael S. Lewis-Beck. 1994. "Comparative Democracy: The Economic Development Thesis." *American Political Science Review* 88:903–10.

- Radcliff, Benjamin, and Patricia Davis. 2000. “Labor Organization and Electoral Participation in Industrial Democracies.” *American Journal of Political Science* 44(1): 132–141.
- Smith, Mark A. 2001. “The Contingent Effects of Ballot Initiatives and Candidate Races on Turnout.” *American Journal of Political Science* 45(3): 700–706.
- Wawro, Gregory, and Ida Pagter Kristensen. 2005. “Lagging the Dog?: The Robustness of Panel Corrected Standard Errors in the Presence of Serial Correlation and Observation Specific Effects.” Working paper: Columbia University. Contact Dr. Wawro (gjw10@columbia.edu) if you’re interested in this paper.

September 26: Models for Dynamic Panel Data.

Required readings:

- Beck, Nathaniel, and Jonathan Katz. 2004. “Time-Series Cross-Section Issues: Dynamics, 2004.” Working paper: New York University.
- Wawro, Gregory. 2002. “Estimating Dynamic Panel Data Models in Political Science.” *Political Analysis* 10(Winter):25–48.

Recommended readings:

- Achen, Christopher. 2000. “Why Lagged Dependent Variables Can Suppress the Explanatory Power of Other Independent Variables.” Presented at the Annual Meeting of the Society for Political Methodology, UCLA.
- Anderson, T.W., and C. Hsiao. 1982. “Formulation and Estimation of Dynamic Models Using Panel Data.” *Journal of Econometrics* 18:47–82.
- Beck, Nathaniel. 1991. “Comparing Dynamic Specifications: The Case of Presidential Approval.” *Political Analysis* 3:51–87.

October 3: Random Coefficient and Other Flexible Models.

Required readings:

- Beck, Nathaniel and Jonathan Katz. 2004. “Random Coefficient Models for Time-Series Cross-Section Data.” Working paper: California Institute of Technology.
- Western, Bruce. 1998. “Causal Heterogeneity in Comparative Research: A Bayesian Hierarchical Modelling Approach.” *American Journal of Political Science* 42:1233–59.

Recommended readings:

- Jackson, John E. 1991. “Estimation of Models with Variable Coefficients.” *Political Analysis* 3:27-49.
- Wood, B. Dan. 2000. “Weak Theories and Parameter Instability: Using Flexible Least Squares to Take Time-Varying Relationships Seriously.” *American Journal of Political Science*. 44:603-18.

October 10: Models for Ugly Response Variables, Part I: Binary Data.

Required readings:

- Beck, Nathaniel, Jonathan N. Katz, and Richard Tucker. 1998. “Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable.” *American Journal of Political Science* 42(October):1260–88.
- Hsiao, Cheng. 2003. *The Analysis of Panel Data*, Chapter 7, §7.1–7.3.

Recommended readings:

- Green, Donald P., Soo Yeon Kim, and David Yoon. 2001. “Dirty Pool.” *International Organization*, 55:441–68 (and commentary by Beck & Katz, Oneal & Russett, and King).
- Katz, Ethan. 2001. “Bias in Conditional and Unconditional Fixed Effects Logit Estimation.” *Political Analysis* 9(Autumn):379–84 (and also see Coupé, Tom (2005) “Bias in Conditional and Unconditional Fixed Effects Logit Estimation: A Correction,” *Political Analysis* 13(Summer):292–95).

October 17: Models for Ugly Response Variables, Part II and Wrap-Up.

Required readings:

- Cameron, A. Colin, and Pravin K. Trivedi. 1998. *Regression Analysis of Count Data*. New York: Cambridge University Press. Chapter 9.
- Neuhaus, J. M., J. D. Kalbfleisch, and W. W. Hauck. 1991. “A Comparison of Cluster-Specific and Population-Averaged Approaches for Analyzing Correlated Binary Data.” *International Statistical Review* 59(1):25–35.
- Zorn, Christopher. 2001. “Generalized Estimating Equation Models for Correlated Data: A Review with Applications.” *American Journal of Political Science* 45(April):470–90.

Recommended readings / Applications:

- Ballinger, Gary A. 2004. “Using Generalized Estimating Equations for Longitudinal Data Analysis.” *Organizational Research Methods* 7:127–50.

- Caldeira, Gregory A., John R. Wright, and Christopher Zorn. 1999. “Strategic Voting and Gatekeeping in the Supreme Court.” *Journal of Law, Economics and Organization* 15(3):549–72.
- Leeds, Brett Ashley, and David R. Davis. 1997. “Domestic Political Vulnerability and International Disputes.” *Journal of Conflict Resolution* 41(December):814–34.
- Razaghian, Rose. 2004. “Financing the Civil War: The Confederacy’s Financial Strategy.” Yale ICF Working Paper No. 04–45.

Part Two: Survival Analysis

Reference Works

- Blossfeld, Hans-Peter, and Götz Rohwer. 2001. *Techniques of Event History Modeling: New Approaches to Casual Analysis*, 2nd Ed. Mahwah, NJ: Lawrence Erlbaum.
- Box–Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*. New York: Cambridge University Press.
- Collett, Dave. 2003. *Modeling Survival Data in Medical Research*, 2nd Ed. London: Chapman & Hall.
- Elandt–Johnson, Regina, and Norman Johnson. 1999. *Survival Models and Data Analysis*. New York: Wiley.
- Fleming, Thomas R., and David P. Harrington. 1991. *Counting Processes and Survival Analysis*. New York: Wiley.
- Hosmer, David, and Stanley Lemeshow. 1999. *Applied Survival Analysis: Regression Modeling of Time to Event Data*. New York: Wiley.
- Kalbfleisch, J. D., and R. L. Prentice. 2002. *The Statistical Analysis of Failure Time Data*, 2nd Ed. New York: Wiley.
- Klein, John P., and Melvin L. Moeschberger. 1997. *Survival Analysis: Techniques for Censored and Truncated Data*. New York: Springer–Verlag.
- Lancaster, Tony. 1990. *The Econometric Analysis of Transition Data*. New York: Cambridge University Press.
- Le, Chap T. 1997. *Applied Survival Analysis*. New York: Wiley.
- Lawless, J. F. 2002. *Statistical Methods and Models for Lifetime Data*, 2nd Ed. New York: Wiley.
- Vermunt, Jeroen K. 1997. *Log–Linear Models for Event Histories*. Thousand Oaks, CA: Sage Publications.

Some Useful Resources

- Essex/ISER Course, “Survival Analysis with Stata”
(<http://www.iser.essex.ac.uk/teaching/degree/stephenj/ec968/>).
- Jan Box–Steffensmeier’s “Event History” course page
(<http://psweb.sbs.ohio-state.edu/faculty/jbox/Courses/ps786eh/ps786ehhome.html>).
- Stata and R examples from Hosmer and Lemeshow (1999)
(<http://www.ats.ucla.edu/stat/r/examples/asa/default.htm>).
- Survival analysis in R page
(<http://spider.stat.umn.edu/R/library/survival/html/00Index.html>).
- UCLA “Survival Analysis with Stata” page
(http://www.ats.ucla.edu/stat/stata/seminars/stata_survival/default.htm).

Course Schedule

October 24: Data, Descriptive Statistics, and Bivariate Analyses.

Required readings:

- Hosmer, David W., and Stanley Lemeshow. 1999. *Applied Survival Analysis: Regression Modeling of Time to Event Data*, pp. 27-84 and Appendix 1.
- Box–Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*, Chapters 1–2.

Recommended readings:

- Cioffi–Revilla, Claudio. 1984. “The Political Reliability of Italian Governments: An Exponential Survival Model.” *American Political Science Review* 78(2):318-37.
- Zelditch, Morris Jr. and Joan Butler Ford. 1994. “Uncertainty, Potential Power, and Nondecisions.” *Social Psychology Quarterly* 57(1):64-73.

October 31: Parametric Survival Models.

Required readings:

- Box–Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*, Chapter 3.
- Alt, James, and Gary King. 1994. “Transfers of Governmental Power: The Meaning of Time Dependence.” *Comparative Political Studies* 27(2):190–210.

Recommended readings / Applications:

- Bennett, D. Scott and Allan C. Stam III. 1996. "The Duration of Interstate Wars." *American Political Science Review* 90(June):239–57.
- Bueno de Mesquita, Bruce, and Randolph M. Siverson. 1995. "War and the Survival of Political Leaders: A Comparative Study of Regime Types and Political Accountability." *American Political Science Review* 89(June):841–55.
- Hosmer, David W., and Stanley Lemeshow. 1999. *Applied Survival Analysis: Regression Modeling of Time to Event Data*, pp. 271-305.
- Martinek, Wendy L., Mark Kemper, and Steven R. Van Winkle. 2002. "To Advise and Consent: The Senate and Lower Federal Court Nominations, 1977-1998." *Journal of Politics* 337–61.
- Teachman, Jay D., and Mark D. Hayward. 1993. "Interpreting Hazard Rate Models." *Sociological Methods and Research* 21(February):340–71.

November 7: Cox’s Proportional Hazards Model and Discrete–Time Models.

Required readings:

- Box–Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*, Chapters 4–5.
- Laird, Nan, and Donald Oliver. 1981. "Covariance Analysis of Censored Survival Data Using Log-Linear Analysis Techniques." *Journal of the American Statistical Association* 96(June):231–40.
- Review Beck, Katz, and Tucker (1998 *AJPS*).

Recommended readings / Applications:

- Alt, James E., Gary King and Curtis S. Signorino. 2001. "Aggregation Among Binary, Count and Duration Models: Estimating the Same Quantities from Different Levels of Data." *Political Analysis* 9(Winter):21–44.
- Box-Steffensmeier, Janet M., Laura W. Arnold, and Christopher Zorn. 1997. "The Strategic Timing of Position Taking in Congress: A Study of the North American Free Trade Agreement." *American Political Science Review* 91(June):324–38.
- Hegre, Havard, Tanja Ellingsen, Scott Gates and Nils Petter Gleditsch. 2001. "Toward a Democratic Civil Peace? Democracy, Political Change, and Civil War, 1816-1992." *American Political Science Review* 95(March):33–48.
- Lindsey, J.K. 1998. "Counts and Times to Events." *Statistics in Medicine* 17:1745–51.

- Royston, Patrick, and M. K. B. Pumar. 2002. “Flexible Parametric Models for Censored Survival Data, with Applications to Prognostic Modeling and Estimation of Treatment Effects.” *Statistics in Medicine* 21(13):1275–97.
- Singer, Judith D., and John B. Willett. 1993. “Its About Time: Using Discrete–Time Survival Analysis to Study Duration and the Timing of Events.” *Journal of Educational Statistics* 18(Summer):155–95.

November 14: Proportionality and Duration Dependence.

Required readings:

- Box–Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*, Chapter 8.
- Box–Steffensmeier, Janet M., and Christopher Zorn. 2001. “Duration Models and Proportional Hazards in Political Science.” *American Journal of Political Science* 45(October):951–67.
- Heckman, James J. 1991. “Identifying the Hand of the Past: Distinguishing State Dependence from Heterogeneity.” *American Economic Review* 81(May):75–79.
- Zorn, Christopher. 2000. “Modeling Duration Dependence.” *Political Analysis* 8(Autumn):367–80.

Recommended readings:

- Bennett, D. Scott. 1999. “Parametric Duration Models, Duration Dependence, and Time–Varying Data Revisited.” *American Journal of Political Science* 43(January):256–70.
- Ng’andu, N. H. 1997. “An Empirical Comparison of Statistical Tests for Assessing the Proportional Hazards Assumption of Cox’s Model.” *Statistics in Medicine* 16:611–26.

November 21: Heterogeneity.

Required readings:

- Box–Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*, Chapter 9.
- Zorn, Christopher, and Janet M. Box–Steffensmeier. 2005. “Cure Models for Political Science Research.” Working paper: Ohio State University.

Recommended readings / Applications:

- Carpenter, Daniel. 2002. "Groups, the Media, Agency Waiting Costs and FDA Drug Approval." *American Journal of Political Science* 46(July):490–505.
- Chiozza, Giacomo, and Hein E. Goemans. 2004. "International Conflict and the Tenure of Leaders: Is War Still *Ex Post* Inefficient?" *American Journal of Political Science* 48(July):604–18.
- Hettinger, Virginia, and Christopher Zorn. 2005. "Explaining the Incidence and Timing of Congressional Responses to the U.S. Supreme Court." *Legislative Studies Quarterly* 30(February):5–28.
- Omori, Yasuhiro, and Richard A. Johnson. 1993. "The Influence of Random Effects on the Unconditional Hazard Rate and Survival Functions." *Biometrika* 80(4):910–14.
- Schmidt, Peter, and Anne D. Witte. 1989. "Predicting Recidivism Using 'Split-Population' Survival Time Models." *Journal of Econometrics* 40(1):141–59.

November 28: Multiple and Competing Events.

Required readings:

- Box–Steffensmeier, Janet M., and Bradford S. Jones. 2004. *Event History Modeling: A Guide for Social Scientists*, Chapter 10.
- Box-Steffensmeier, Janet M., and Christopher Zorn. 2002. "Duration Models for Repeated Events." *Journal of Politics* 64(November):1069–94.

Recommended readings / Applications:

- Cleves, Mario. 1999. "Analysis of Multiple Failure–Time Data with Stata." *Stata Technical Bulletin* 49:30–39.
- Diermeier, Daniel, and Randy T. Stevenson. 1999. "Cabinet Survival and Competing Risks." *American Journal of Political Science* 43(October):1051–68.
- Gordon, Sanford C. 2002. "Stochastic Dependence in Competing Risks." *American Journal of Political Science* 46(January):200–17.
- Kelly, Patrick J., and Lynette L–Y. Lim. 2000. "Survival Analysis for Recurrent Event Data." *Statistics in Medicine* 19:12–33.
- Wei, L. J., and David V. Glidden. 1997. "An Overview of Statistical Methods for Multiple Failure Time Data in Clinical Trials." *Statistics in Medicine* 16:833–39.
- Zorn, Christopher, and Steven R. Van Winkle. 2000. "A Competing Risks Model of U.S. Supreme Court Vacancies, 1789–1992." *Political Behavior* 22(June):145–66.