Pooled Time Series and Cross-Sectional Analysis
Political Science 597C
Spring 2004

Professor De Boef
Office Hours:
Mondays, 10:00-11:30
Fridays, 1:30-2:45

sdeboef@psu.edu
N160 Burrowes
863-9402

This course is designed to teach students to model relationships involving data that has both a time serial and cross sectional dimension. Examples include modeling voting outcomes across states and over time, attitudes across individuals and over time, budgets across agencies and over time, etc. These types of data allow us to conduct dynamic comparison, but they also present unique difficulties for the analyst. This class will teach students how to conduct dynamic comparison. Class lectures will be a mixture of theory and application; homework will be applied and will require Stata. Readings should be read prior to the class period for which they are listed. Students should come to class with a dataset suitable for analysis. In addition to weekly homework assignments, students will be required to produce and present a poster and take a final exam. Finally, I will introduce students to latex and expect all homework assignments will be done in latex.

This course meets for the first half of the semester only and is 1.5 credits. Students may choose to enroll in 597B, Limited Dependent Variable Analysis, the courses will run consecutively. While there are no formal prerequisites for the course, I expect students to have had the equivalent of PLSC 502 and PLSC 503.

The Plan.
Weekly reading assignments should be read before (and again after!) the class period in which they are listed. Typically the readings are quite short, but you will find that a solid understanding of the material will require you to read and reread the material. I did not order books for the course as we will be reading selections from multiple texts. If you choose to purchase a book on time-series cross-section and panel data analysis for your reference, I recommend Cheng Hsiao's 2003 text, *Analysis of Panel Data*.

Each week homework assignments will require you to apply the techniques covered in class and readings to your own datasets and to interpret your results. To that end and because the course meets for only 8 weeks, it is imperative that you select and prepare a dataset to use for weekly assignments and for your poster before January 23rd. If you do not have a dataset to work with, I recommend talking with faculty or your peers for ideas and sources ASAP. The only restriction on the dataset used for the course is that it contains a continuous dependent variable. We will work with binary dependent variables, as well, but most of the course will focus on linear models. Finally, you will find that life is simpler for you if you have time-series cross-section (rather than panel) data that contains 8 or more units and as many or more time periods.

In addition, most weeks you will also write a one page summary of one of the reading assignments. Brief versions of the assignments are given on the syllabus, but more detailed instructions will be handed out each class period.

Given the brevity of the course, rather than write a full paper for this class, everyone will present a poster as if you were presenting a poster at the APSA, for example. We will hold a special class session and students and faculty will be invited to attend. We'll talk more about the posters in class.

Grading:

Weekly homework assignments:	50%
Final Exam (last class period):	25%
Poster:	25%

A word on attendance. Given the few numbers of course meetings it is particularly important that you do not miss class unless absolutely necessary. When possible, please let me know in advance if you are ill or cannot make class.

Tentative Schedule

Friday January 16. Data, Stata, and Latex.

Required Readings:

- De Boef, lecture notes Part 1. "Introduction."
- De Boef, notes on "Stata Commands for Pooled Time-series and Cross-sections."
- De Boef, "Getting Started with WinEdt and Latex."
- De Boef, "Introduction to Latex".

I will email each of these readings to you. The last is a tex file called ltx_slides.tex. See "Getting Started with WinEdt and Latex" for instructions for opening and viewing this file.

- Nagler, Jonathan. 1995. "Coding Style and Good Computing Practices." *The Political Methodologist*. 6(2):2-9. This is available in pdf format online at http://polmeth.wustl.edu/. Select *The Political Methodologist* and the correct volume.
- Stata Manual section on –xt- commands. Read the manual itself, rather than the help menu in Stata.

Recommended Readings:

- The last 2 volumes of *The Political Methodologist* contain a "Latex Corner", which provides helps and hints for getting started using Latex. Check these out at http://polmeth.wustl.edu/. Additional help is easily found through google.

Homework Assignment 1 (due Friday January 23): Produce a dataset and read it into Stata. Describe the dataset (both basic xtsum statistics and substantive description) using latex. See handout.

Friday January 23. Introduction to Panel and Pooled Time-Series Cross-Section. OLS and Pooled Times-Series Cross-Section Data.

Required Readings:

- De Boef, lecture notes Parts 2&3. "Advantages of Times-Series Cross-Section Analysis" and "OLS and Pooled Designs."
- Hsiao, Cheng. 2003. *Analysis of Panel Data*, Chapter 1 and Chapter 2.
- Nuamah, Nicholas N.N.N. 1986. "Pooling Cross Section and Time Series Data." *The Statistician* 35:345-351. (Available on JSTOR)
- Greene, William H. 2003. *Econometric Analysis*, 5th edition, sections 13.1-13.2.

Homework Assignment 2 (due Friday January 30): Estimate and Interpret an OLS model using your data. What concerns do you have about this model? Why? Give some basic descriptive information on model residuals and discuss model fit across the units. See handout.

Friday January 30. Dealing with Heterogeneity: Fixed Effects (LSDV) and Random Effects (GLS-Error components).

Required Readings:

- De Boef, lecture notes Part 4. "Dealing with Heterogeneity."
- Hsiao, Cheng. 2003. *Analysis of Panel Data*, Chapter 3.
- Greene, William H. 2003. *Econometric Analysis*, 5th edition, sections 13.3-13.4.
- Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data*, chapter 10.
- Stimson, James A. 1985. "Regression in Space and Time: A Statistical Essay." *American Journal of Political Science* 29:914-947.

Read one of the following applications:

- Zahariadis, Nikolaos. 1997. "Why State Subsidies? Evidence from European Community Countries 1981-1986." *International Studies Quarterly*, 41(2):341-354.
- 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.
- Wood, B. Dan. 1998. "The Dynamics of Senatorial Representation, 1952-1991." *The Journal of Politics*, 60(3):705-736.
- Zorn, Christopher J.W. 2001. "Estimating Between- and Within-Cluster Covariate Effects, with an Application to Models of International Disputes." *International Interactions*, 27(4):433-445.

Homework Assignment 3 (due Friday February 6):Estimate and Interpret fixed effects and random effect models using your data. Estimate both a LSDV and standardize your data and estimate fixed effects models. Estimate a random effects model. Explain which model is a better estimator for your research question and why. Write a one page summary of one of the applications.

Friday February 6. Dynamic Models for pooled time-series cross-section data.
Required Readings:

- De Boef, lecture notes Part 5. "Dealing with Time Dependence".
- Hsiao, Cheng. 2003. *Analysis of Panel Data*, Chapter 4, sections 1 through 5.
- Greene, William H. 2003. *Econometric Analysis*, 5th edition, sections 13.6-13.7.
- Wawro, Gregory. 2002. "Estimating Dynamic Panel Data Models in Political Science." *Political Analysis* 10(Winter):25-48.

Read one of the following applications:

- Abramson, Paul R., Susan Ellis and Ronald Inglehart. 1997. "The Economics of Politics in Comparative Perspective Revisited." *Political Behavior* 19(2): 41-59.
- Green, Donald P., David H. Yoon. 2002. "Reconciling Individual and Aggregate Evidence Concerning Partisan Stability: Applying Time-Series models to Panel Survey Data. *Political Analysis* 10:1-24.

Homework Assignment 4 (due Friday February 13):Estimate and Dynamic models using your data. Write a one page summary of one of the applications.

Friday February 13: Additional Issues in Estimation.
GLS Extensions:
Heteroscedasticity & Unit (Panel) Contemporaneous Correlations.
Panel Corrected Standard Errors.
Random Coefficient Models and more.
Required Readings:

- De Boef, lecture notes Parts 6&7. "Dealing with Both Unit and Time Effects" and "Unit or Panel Contemporaneous Correlations".
- Beck, Nathaniel and Jonathan Katz. 1995. "What to Do (And Not to Do) With Time-Series Cross-Section Data." *American Political Science Review* 89(September):634-647.
- Beck, Nathaniel and Jonathan Katz. 1996. "Nuisance vs. Substance: Specifying and Estimating Time-Series Cross-Section Models." *Political Analysis* 6:1-36.
- Greene, William H. 2003. *Econometric Analysis*, 5th edition, sections 13.8-13.9.
- Zorn, Christopher J.W. 2001. "Generalized Estimating Equation Models for Correlated Data: A Review with Applications." *American Journal of Political Science* 45(2):470-490.

Read one of the following applications:

- Burkhart, Ross E. and Michael S. Lewis-Beck.1994. "Comparative Democracy: The Economic Development Thesis." *The American Political Science Review*. 88(4):903-910.

- Fording, Richard. C. 1997. "The Conditional Effect of Violence as a Political Tactic: Mass Insurgency, Welfare Generosity, and Electoral Context in the American States." *American Journal of Political Science* 41(1): 1-29.

Homework Assignment 5 (due Friday February 20): TBA

Friday February 20: Final Wrap Up on Linear Models for Pooled Time-Series and Cross-Sections.

Required Readings:

- De Boef, lecture notes Part 8. "A Strategy for Pooled Time-series and Cross-Section Analysis".

Homework Assignment 6 (due Friday February 27): TBA

Friday February 27: Binary Dependent Variables (and a brief word on counts and duration models).

Required Readings:

- De Boef, lecture notes Part 9. "Binary Dependent Variables".
- Hsiao, Cheng. 2003. *Analysis of Panel Data*, Chapter 7.
- Beck, Nathaniel and Jonathan Katz. 1998. "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable." *American Journal of Political Science* 42(October):1260-1288.

Recommended Readings:

- Cameron, Colin and Pravin K. Trivedi. 1998. *Regression Analysis of Count Data*. New York: Cambridge University Press. Chapter 9.
- Green, Donald P., Soo Yeon Kim, and David Yoon. 2001. "Dirty Pool." *International Organization* 55:441-468.
- 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 J.W. 2001. "Generalized Estimating Equation Models for Correlated Data: A Review with Applications." *American Journal of Political Science* 45(April):470-90 (and erratum).
- Wawro, Gregory. 2001. "A Panel Probit Analysis of Campaign Contributions and Roll Call Votes." *American Journal of Political Science* 45(July):563-579.

Homework Assignment 7 (due Friday March 2): Estimate and Interpret a binary pooled time-series cross-section model. If you do not have a binary dependent variable in your data set, find a class mate who does or see me for data.

Friday March 2: Exam