Panel Data Analysis using SAS and SPSS

2011 ICPSR SUMMER WORKSHOP AT THE UNIVERSITY OF MICHIGAN
JUNE 6-10, 2011

Professor Patricia McManus       Teaching Assistant
Department of Sociology       Michael F. Thompson
Indiana University       mfthomps@indiana.edu
pmcmanus@indiana.edu

Description
Panel data, also known as cross-sectional time-series data, contain many individual units that are observed at more than one point in time. Statistical techniques that exploit the within-group correlation structures of panel data offer powerful advantages over conventional regression analysis. This workshop uses an intensive mix of lecture and hands-on computer sessions to introduce students to the specification, estimation, and interpretation of models for the analysis of panel data. The workshop begins with the basic theory underlying the analysis of panel data along with examples of how to use - and how not to use - panel models. Essential terminology and notation are introduced using a review of the classical linear regression model. The second day covers practical applications of the linear error components model from assumptions to estimation and interpretation of the first difference model, fixed effects models, and random effects models. The third day is devoted to specification tests, autocorrelation and heteroskedastic errors, and a presentation of a hybrid approach that combines information from the random effects and fixed effects models. The fourth day covers linear dynamic panel models and instrumental variable estimation. The final day of the workshop covers fixed-effects and random effects models for binary and count models. The course assumes familiarity with the linear regression model.

Daily Schedule
Lecture  9-12:30
Break     12:30-1:30
Lab/Lecture 1:30-5:30

Texts
McManus, Patricia A. 2011. Lecture Notes for Panel Data Using SAS and SPSS.

The notes contain lecture material, overheads, lab notes and exercises for the course. Bring these materials to class, and to get the most out of the class, review the days lecture notes in advance.
Recommended Texts


These texts will prove useful after the workshop. If you are already familiar with and comfortable with the level of presentation in Wooldridge (2010), you will find the pace of this course too slow. Allison (2005) is a slim and very accessible introduction to fixed effects models in SAS for applied data analysts, and includes several chapters on nonlinear fixed effects models. Unfortunately, it does not include coverage of SAS/ETS and PROC PANEL.

Workshop Outline
Each topic below will be included in the workshop, but the time spent on each topic will vary based on student background and interests.

1. Overview: introduction to panel data, advantages, disadvantages, examples, common errors (*Day 1*)

2. Review of general linear model: assumptions of single-equation linear regression model and estimation in the presence of heteroskedastic or correlated error structures using OLS, GLS, FGLS (*Day 1*)

3. Pooled Cross-Sections & Difference-in-Difference Analysis (*Day 1*)

4. Error components models: first differences model, fixed effects, random effects (*Day 2*)

5. Panel data workflow and decision tree: Specification tests, serial correlation and heteroskedastic errors in panel models (*Days 2-3*)

6. The Mundlak transformation hybrid panel model (*Day 3*)

7. Dynamic panel models: instrumental variables & GMM approaches (*Days 4-5*)

8. Panel models for dichotomous outcomes: FE (conditional) logit and RE logit and probit, fixed effects models for count outcomes (*Day 5*)

9. Introduction to generalized estimation equations (GEE) for population-averaged effects (*Day 5*)