SOC 590: Longitudinal Data Analysis
Fall 2007, Thursdays 4-7 (first half of semester), 165 Wallace Hall

Professor: Scott M. Lynch
Office: 159 Wallace Hall (Hours: by appointment)
Phone: 258-7255
Email: slynch@princeton.edu

Objectives

The days of using exclusively one-wave cross-sectional data to answer social science research questions are mostly over. Instead, many of the questions researchers ask today involve life course concepts and theories that require the use of repeated cross-sectional or panel (i.e., longitudinal) data to answer. Such data provide more leverage than one-wave cross-sectional data in answering research questions, but they also pose difficulties that necessitate more sophisticated methods of analysis than one-wave cross-sectional data require. This course is intended to be a broad and highly-applied exposition of methods appropriate for the analysis of longitudinal data. In an effort to cover many of the most common models, this course provides greater breadth than depth of the material and can be viewed as a survey of numerous methods. I expect that students will investigate these methods in greater depth as needed in their own research outside the class, but we will discuss required data structures, software syntax, and interpretation of results for all methods covered in order to provide the necessary foundation for undertaking one's own analyses.

Longitudinal data analysis involves the modeling of two types of outcomes: repeated measures (like wages, health, or political affiliation at different points in time) and events (like death, marriage/divorce, and Ph.D. completion). We will cover methods for both types of outcomes in the course. Methods for analyzing repeated measures have become increasingly commmon in sociology over the last two decades, and we will discuss the most common methods, including fixed and random effects models, growth models, and classic econometric time series models. Methods for analyzing events have always been common in demography and are increasingly used in sociology. Thus, we will cover a variety of traditionally "demographic" methods also, including discrete and continuous time hazard models and life table methods. In the process of covering all of these methods, we will construct a flowchart to help direct one in determining the appropriate method for answering research questions based on (1) the type of question asked and (2) the type of data available for analysis. We will also consider how to handle missing data and how to assess (and possibly correct for) the influence of selective mortality in longitudinal analyses.

Prerequisites

It is expected that students have had both a basic statistics course and a course covering multiple regression (e.g., SOC 504). Familiarity with matrix algebra, probability theory, and maximum likelihood estimation is expected, but we will review both in the first week.

Format and Requirements

The course consists of 6 weekly three-hour sessions. Weekly homework will be given; most will involve the use of Stata, MPlus (structural equation modeling software), and possibly R and WinBugs. These latter two software packages are free and downloadable; the first two are available on the department's cluster computers.

There are a number of required readings, each of which is available through amazon.com or via JSTOR or related electronic sources. The required readings are listed in the course schedule below. Please have the material read BEFORE class on the day the readings are listed, as we will discuss them on that day.

Course Schedule and Readings*

Week 1; September 20

- Overview of course
- Review of matrix algebra, probability theory, maximum likelihood estimation, and mathematical notation
- The life course perspective, age, period, cohort effects and the identity problem
- Data types and structures; problems longitudinal data present
- Causality

Readings

NONE

Week 2; September 25

- Individual change vs. cohort turnover
- Age, Period, Cohort (APC) models for repeated cross-sectional data (linear and generalized linear regression)

Readings

Firebaugh, G. (1997). "Analyzing Repeated Surveys." Sage University Paper Series on Quantitative Applications in the Social Sciences, no. 07-115. Thousand Oaks, CA.

Glenn, N. (2005). "Cohort Analysis." Sage University Paper Series on Quantitative Applications in the Social Sciences, no. 07-005, 2nd ed. Thousand Oaks, CA.

Ryder, NB. (1965). "The Cohort as a Concept in the Study of Social Change." *American Sociological Review* 30:843-861.

Week 3; October 2

- Autocorrelation and Time Series models (ARiMA) (Stata's `arima` and related procedures)
- Fixed and Random effects models (Stata's `xtreg` procedure)
- Change score models for two-wave panel data

Readings

Menard, S. (2002). "Longitudinal Research." Sage University Paper Series on Quantitative Applications in the Social Sciences, no. 07-076, 2nd ed. Thousand Oaks, CA.

Ostrom, CW. (1990). "Time Series Analysis: Regression Techniques." Sage University Paper Series on Quantitative Applications in the Social Sciences, no. 07-009, 2nd ed. Thousand Oaks, CA.

Sayrs, LW. (1989). "Pooled Time Series Analysis." Sage University Paper Series on Quantitative Applications in the Social Sciences, no. 07-070, 2nd ed. Thousand Oaks, CA.

Week 4; October 9

- "Growth" models (SEM approach using MPlus, Stata's `xtmixed` procedure, and other approaches)

Readings

George, LK and SM Lynch. (2003). "Race Difference in Depressive Symptoms: A Dynamic Perspective on Stress Exposure and Vulnerability." *Journal of Health and Social Behavior* 44(March);353-369.

Lynch, SM. (2003). "Cohort and Life Course Patterns in the Relationship Between Education and Health: A Hierarchical Approach." *Demography* 40(2):309-331.

Lynch, SM and LK George. (2002). "Interlocking Trajectories of Loss-Related Events and Depressive Symptoms Among Elders." *Journals of Gerontology: Social Sciences* 57B(2):S117-S125.

Meredith, W and J Tisak. (1990). "Latent Curve Analysis." *Psychometrika* 55:107-122.

Willett, JB and AG Sayer. (1994). "Using Covariance Structure Analysis to Detect Correlates and Predictors of Individual Change Over Time." *Psychological Bulletin* 116:363-381.

Week 5; October 16

- continuous and discrete time hazard models (discrete time regression models)
- life table methods (single decrement, multiple decrement, and multistate tables)
- Sullivan's method; cross-sectional survival models

Readings

Allison, PD. (1984). *Event History Analysis: Regression for Longitudinal Event Data.* Sage University Paper series on Quantitative Applications in the Social Sciences, 07-046. Beverly Hills, CA: Sage.

Hill, ME. (1999). "Multivariate Survivorship Analysis Using Two Cross-Sectional Samples." *Demography* 36(4):497-503.

Lauderdale, DS. (2001). "Education and Survival: Birth Cohort, Period, and Age Effects." *Demography* 38(4):551-562.

Lynch, SM and JS Brown. (2005). "A New Approach to Estimating Life Tables with Covariates and Constructing Interval Estimates of Life Table Quantities". *Sociological Methodology* 35:177-225.

Molla, MT, DK Wagener, and JH Madans. (2001). "Summary Measures of Population Health: Methods for Calculating Healthy Life Expectancy." *Healthy People Statistical Notes, no. 21,* Hyattsville, MD: National Center for Health Statistics.

Palloni, A. (2000). "Increment-Decrement Life Tables," Pp. 256-272 in SH Preston, P. Heuveline, and M. Guillot *Demography: Measuring and Modeling Social Processes.* Oxford, England: Blackwell.

Sullivan, DF. (1971). "A Single Index of Mortality and Morbidity." *HMSHA Health Reports* 86:347-354.

Week 6; October 23

- Handling missing data in longitudinal analyses
- Selective mortality

Readings

Allison, PD. (2002). *Missing Data.* Sage University Paper Series on Quantitative Applications in the Social Sciences, 07-136. Oaks, CA: Sage.

Beckett, M. (2000). "Converging Health Inequalities in Later Life—An Artifact of Mortality Selection?" *Journal of Health and Social Behavior* 41:106-119.

Lynch, SM, JS Brown, and KG Harmsen. (2003). "Black-White Differences in Mortality Deceleration and Compression and the Mortality Crossover Reconsidered." *Research on Aging* 25(5):456-483.

Vaupel, JW and AI Yashin. (1985). "Heterogeneity's Ruses: Some Surprising Effects of Selection on Population Dynamics." *The American Statistician* 39(3):176-185.

* NOTE: I expect to provide additional readings each week as needed; these are the minimum.