SC706: LONGITUDINAL DATA ANALYSIS

Instructor: Natasha Sarkisian
Email: natasha@sarkisian.net
Office phone: (617) 552-0495
Home phone: (617) 755-3178
Mailbox: McGuinn 426
Office: McGuinn 417
Office hours: By appointment
Class time: Tuesdays 3:30-6:00 PM
Class location: O'Neill 245
Webpage: http://www.sarkisian.net/sc706/

COURSE DESCRIPTION*

This applied course is designed for graduate students with a prior background in statistics at the level of SC703: Multivariate Statistics (or its equivalent). This means that students should have considerable experience with ordinary-least-squares (OLS) regression: I assume you have an understanding of multiple OLS regression and an ability to conduct such analyses using some statistical software (e.g., SPSS, SAS, Stata, etc.). The major topics of the course will include event history analysis (also known as survival analysis), time series analysis (ARIMA models), and panel data analysis (fixed effects, random effects, and mixed models).

The goals of the course are to develop the skills necessary to identify an appropriate technique, estimate models, and interpret results for independent research and to critically evaluate contemporary social research using advanced quantitative methods. The course will be applied in the sense that we will focus on estimating models and interpreting the results, rather than on understanding in detail the mathematics behind the techniques. I hope that the course will provide you with a solid foundation in longitudinal data analysis, which is a type of advanced quantitative skill that is in high demand in many fields, both in and out of academia. For those of you in the Sociology Department, the course can also provide a foundation for the "Advanced Quantitative Methods" area examination.

We will be using Stata for all the analyses throughout the course. No previous Stata experience is necessary: I will provide an introduction to Stata in the beginning of the course and guide you throughout the course.

COURSE POLICIES

For each topic in the course, I will give a lecture focusing on the reasoning behind the technique, and provide a review of the syntax used to do analyses and the output generated by Stata. Throughout that process, you will get a chance to practice conducting the analyses and interpreting the results. We will also discuss and critically evaluate published research based on the various techniques. Make sure that you carefully read these examples of published research before class and be prepared to discuss them. The course is based on an interactive relationship between the instructor and students, as well as on collaboration among the students. You are strongly encouraged to ask questions and discuss the material in class. I also encourage collaboration among the students. Please feel free to help each other when running analyses for assignments. However, everyone must turn in their own report and statistical output.

* This syllabus draws upon ideas presented in syllabi by a number of people, including Robert Kunovich, John Williamson, Joya Misra, and Doug Anderton.

I also would like to stress that you are always welcome to come and see me with any additional questions. Email is the best way to get in touch with me – I check my email very often. Email is the best way to get a quick question answered or to set up an appointment to discuss something at length. You are also welcome to call me either in my office or at home (any time between 9 AM and 10 PM); however, be prepared to leave your name and number if I am not available to pick up the phone. Also, please check our course website regularly: various course materials (assignments, handouts, etc.) will be posted there regularly. And make sure to check your email, too – from time to time I may send some announcements.

Finally, a note on feedback. I would like to know how I could make this course experience as useful and interesting as possible. Therefore, every class in the end of class I will ask you to submit a sheet of paper with the date and at least one sentence of reaction to that class meeting, indicating what you learned, or something you liked or did not like, found interesting or controversial, found clear or too simplistic, or found confusing and in need of further (or better) explanation. You may also submit comments on the course in general.

REQUIRED MATERIALS

Books:
The following books should be available for purchase at the BC bookstore. They are also placed on reserve at the library.
1. Taris, Toon W. 2000. *A Primer in Longitudinal Data Analysis.* Thousand Oaks, CA: Sage Publications.
2. Cleves, Mario Alberti, William Gould, and Roberto G. Gutierrez. 2004. *An Introduction to Survival Analysis Using Stata. Revised Edition.* College Station, TX: Stata Corporation.

Other readings:
Other required readings (listed below in the course outline) will be available on electronic reserve in the library: see http://www.bc.edu/libraries/services/reserves/

COURSE REQUIREMENTS AND GRADING

There will be two assignments in this course, each worth 50% of your grade. These assignments will involve selecting a research question and variables, running analyses, and writing up the results like you would for a journal publication (including introduction, data and methods, and results sections). I will provide data, although you can use your own data if they are appropriate for the technique (see me in advance). My intention is that these assignments will assist in the completion of the advanced quantitative methods area exam in sociology and/or will facilitate your own independent research projects.

Each assignment will consist of two drafts, to be submitted electronically (by email or using MyFiles). If you turn in the first draft by the due date, I will comment on it, assign a temporary grade, and return it to you. At that point, we will also discuss the common problems and mistakes. You will then get a chance to submit a revised draft. If you are satisfied with your temporary grade, you do not need to revise the assignment – just let me know. This system will

allow you to push yourself beyond your comfort level without worrying how it will affect your grade. For example, you might try to interpret some of the results not required for the assignment or you might decide to present the results in a more meaningful and perhaps less conventional way.

The letter grades for the final drafts of your assignments will be determined as follows:

93-100	A
90-92	A-
87-89	B+
83-86	B
80-82	B-
60-79	C
0-59	F

COURSE OUTLINE

January 16. Introduction to Longitudinal Data Analysis.

Taris, Toon W. 2000. Chapter 1 from: *A Primer in Longitudinal Data Analysis.* Thousand Oaks, CA: Sage Publications.

January 23. Event History Analysis: Discrete Time Models.

Willett, John B., and Judith D. Singer. 2004. Discrete-Time Survival Analysis. Chapter 11 (pp. 199-214) from: David Kaplan (Ed.), *The Sage Handbook of Quantitative Methodology for the Social Sciences.* Thousand Oaks, CA: Sage Publications. RESERVE

Gupta, Sanjiv, Pamela J. Smock, and Wendy D. Manning. 2004. Moving Out: Transition to Nonresidence among Resident Fathers in the United States, 1968-1997. *Journal of Marriage and Family*, 66, August, 627-638. RESERVE

January 30. Event History Analysis: Nonparametric and Semi-parametric Models.

Taris, Toon W. 2000. Chapter 6 from: *A Primer in Longitudinal Data Analysis.* Thousand Oaks, CA: Sage Publications.

Cleves, Mario Alberti, William Gould, and Roberto G. Gutierrez. 2004. Chapters 8-11 from: *An Introduction to Survival Analysis Using Stata.* College Station, TX: Stata Corporation.

Hill, Twyla J. 2000. Legally Extending the Family: An Event History Analysis of Grandparent Visitation Rights Laws. *Journal of Family Issues*, 21, 246-261. RESERVE

February 6. Event History Analysis: Parametric models.

Cleves, Mario Alberti, William Gould, and Roberto G. Gutierrez. 2004. Chapters 12-15 from: *An Introduction to Survival Analysis Using Stata.* College Station, TX: Stata Corporation.

Teachman, Jay D., and Mark D. Hayward. 1993. Interpreting Hazard Rate Models. *Sociological Methods and Research*, 21, 3, 340-371. RESERVE

Palloni, Alberto, and Elizabeth Arias. 2004. Paradox Lost: Explaining the Hispanic Adult Mortality Advantage. *Demography*, 41, 385-415. RESERVE

February 13. Event History: Time-Varying Predictors and Repeated Events.

Cleves, Mario Alberti, William Gould, and Roberto G. Gutierrez. 2004. Chapters 4-7 from: *An Introduction to Survival Analysis Using Stata.* College Station, TX: Stata Corporation.

Allison, Paul D. 1985. Chapter 6 from: E*vent History Analysis: Regression for Longitudinal Event Data.* Beverly Hills, CA: Sage Publications. RESERVE

February 20. Event History: Competing Risk Models.

Allison, Paul D. 1985. Chapter 5 from: E*vent History Analysis: Regression for Longitudinal Event Data.* Beverly Hills, CA: Sage Publications. RESERVE

Metraux, Stephen, and Dennia P. Culhane. 1999. Recurring Homelessness among Women. *Journal of Family Issues*, 20(3): 371-396. RESERVE

De Graaf, Paul M. and Matthijs Kalmijn. 2003. Alternative Routes in the Remarriage Market: Competing-Risk Analyses of Union Formation after Divorce. *Social Forces*, 81(4):1459-1498. RESERVE

February 27. Sequence Analysis.

Taris, Toon W. 2000. Chapter 7 from: *A Primer in Longitudinal Data Analysis.* Thousand Oaks, CA: Sage Publications.

Macindoe, Heather, and Andrew Abbott. 2004. Sequence Analysis and Optimal Matching Techniques for Social Science Data. Pp. 387-406 (Chapter 17) from: Melissa Hardy and Alan Bryman (Eds.), *Handbook of Data Analysis*. Thousand Oaks, CA: Sage Publications. RESERVE

March 6. No class, Spring Break

March 13. Introduction to Panel Data Analysis: Fixed and Random Effects Models.
******Assignment 1 draft due******

Taris, Toon W. 2000. Chapters 3, 4 and 5 in: *A Primer in Longitudinal Data Analysis.* Thousand Oaks, CA: Sage Publications.

Baum, Christopher. 2006. Chapter 9, pp.219-232, from: An Introduction to Modern Econometrics Using Stata. College Station, TX: Stata Press. RESERVE

Rabe-Hesketh, Sophia, and Anders Skrondal. 2005. Chapter 2 from: Multilevel and Longitudinal Modeling Using Stata. College Station, TX: Stata Press. RESERVE

March 20. Panel Data Analysis: Error Covariance Structure.

Singer, Judith D., & John B. Willett. 2003. Chapter 7 from: Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence. New York: Oxford University Press. RESERVE

Pampel, Fred C., and John Williamson. 1988. "Welfare Spending in Advanced Industrial Democracies: 1950-1980." *American Journal of Sociology, 93*: 1424-1456. RESERVE

March 27. Panel Data Analysis: Mixed Effects Models.

Rabe-Hesketh, Sophia, and Anders Skrondal. 2005. Chapter 3 from: Multilevel and Longitudinal Modeling Using Stata. College Station, TX: Stata Press. RESERVE

Farkas, George, and Kurt Beron. 2004. "The Detailed Age Trajectory of Oral Vocabulary Knowledge: Differences By Class And Race." *Social Science Research*, 33, 464-497. RESERVE

April 3. Panel Data Analysis: Dynamic Models.

Halaby, Charles N. 2004. Panel Models in Sociological Research: Theory into Practice. *Annual Review of Sociology*, 30, 507-544. RESERVE

Baum, Christopher. 2006. Chapter 9, pp.232-236, from: An Introduction to Modern Econometrics Using Stata. College Station, TX: Stata Press. RESERVE

April 10. Introduction to Time Series Analysis.

******Assignment 2 draft due******

Wonnacott, Thomas H., and Ronald J. Wonnacott. 1981. Chapter 6, pp.208-247, from: *Regression: A Second Course in Statistics*. Malabar, FL: Krieger Press. RESERVE.

April 17. Time Series Analysis: Univariate ARIMA Models.

Wonnacott, Thomas H., and Ronald J. Wonnacott. 1981. Chapter 6, pp.248-264, from: *Regression: A Second Course in Statistics*. Malabar, FL: Krieger Press. RESERVE.

McCleary, Richard. 1980. Chapter 3 from *Applied Time Series Analysis for Social Sciences*. Beverly Hills, CA: Sage. RESERVE

April 24. Time Series – multivariate ARIMA Models.

McCleary, Richard. 1980. Chapter 5 from *Applied Time Series Analysis for Social Sciences*. Beverly Hills, CA: Sage. RESERVE

Morris, Theresa. 2003. "Unionization Matters: An Analysis of Post-World War II Strikes." *Sociological Inquiry*, *73*: 245-264. RESERVE.

May 1. Missing Data in Longitudinal Research.

******Assignments 1 and 2 final drafts due May 8******

Taris, Toon W. 2000. Chapter 2 from: *A Primer in Longitudinal Data Analysis.* Thousand Oaks, CA: Sage Publications.