Instructor: Natasha Sarkisian
Email: natasha@sarkisian.net
Phone: (617) 755-3178
Office: McGuinn 417

Class time: Wednesdays 2:00-4:20 PM
Class location: O’Neill 245
Office hours: By appointment
Webpage: http://www.sarkisian.net/socy7706/

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 change models, fixed, random, and mixed effects models, panel models for categorical and count data, techniques to deal with sample selection issues and endogeneity bias, and discrete-time event history analysis.

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. For your assignment, you can use Stata on Citrix: see http://apps.bc.edu. Some additional helpful resources for learning Stata include http://www.ats.ucla.edu/stat/stata/ as well as Stata forum, Statalist: http://www.statalist.org/.

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. 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.

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 to get a quick question answered or to set up an appointment to discuss something at length. Also, please check your email regularly: I will let you know by email when course notes are posted on the website and send other announcements from time to time.

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 (or send an email) 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.
COURSE REQUIREMENTS AND GRADING
All the required readings will be available on electronic reserve in the library: see http://www.bc.edu/reserves

The main assignment for this course will be to write a solid draft of a journal article based on panel data analyses. In preparation for that, you will submit a proposal (5% of your grade), the first draft of your data management (5%), the first draft of your data analysis and article (10% of your grade), and the final drafts of your data analysis (30% of your grade) and article (30% of your grade). Four additional mini-assignments (5% of your grade each) will involve answering questions on a published article that utilizes some of the same methods that will be discussed in class; questions will be posted on the course website one week prior to the assignment due date.

Proposal. The proposal will involve identifying a research question and conducting preliminary literature review, selecting a dataset, and identifying relevant variables. You can select a panel dataset with either two waves or with three or more waves; these two options will lead you to estimate two distinct sets of models. You should also identify your target journal. Please consult me early on if you need help to locate appropriate data or advice on journals. I would recommend selecting a continuous rather than categorical variable as your outcome; please consult with me if you prefer working with a categorical outcome.

Data Management. For this part of your assignment, you will conduct all the data management steps to prepare your dataset for analysis; you will merge and reshape the data as necessary and recode and label all the variables needed for your analysis. You will also conduct multiple imputation to deal with missing data, if any. You will submit your do file or files (please make sure to include enough comments to explain what you are doing at each step) and a log file resulting from running your do file(s).

Data Analysis. For this part of your assignment, you will run all the necessary analyses, conduct diagnostics and apply remedies, and use additional interpretation tools (standardized coefficients, graphs, etc.) to assist in interpreting your findings. You will submit an annotated log that will contain the output (with your brief comments) for all of the tasks that you will perform for this assignment. Please paste your graphs into the log right after the command that was used to generate each graph. Make sure to divide your log into clear sections (with headings) and create a table of contents, to be included as the first page of your log. Use Courier New 9-point font for your log to ensure that the output is well aligned.

Article. For your article, you will select one or more models from your data analysis to present. Your article will include an introduction, literature review, data and methods, results, and conclusion, all written in journal format.

All components of this assignment will be submitted electronically (by email or using MyFiles). When you submit the first drafts of your analysis and article, I will provide you with detailed feedback that you will then use to prepare your final drafts. The letter grades will be determined as follows:

<table>
<thead>
<tr>
<th>Grade</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>93-100</td>
</tr>
<tr>
<td>A-</td>
<td>90-92</td>
</tr>
<tr>
<td>B+</td>
<td>87-89</td>
</tr>
<tr>
<td>B</td>
<td>83-86</td>
</tr>
<tr>
<td>B-</td>
<td>80-82</td>
</tr>
<tr>
<td>C</td>
<td>60-79</td>
</tr>
<tr>
<td>F</td>
<td>0-59</td>
</tr>
</tbody>
</table>
COURSE OUTLINE


***Proposal due***

February 11. Missing Data in Longitudinal Research.

February 18. Two-Wave Panel Data Analysis

February 25. Fixed Effects Models.
***Mini-assignment 1: Answers to article questions due 10 AM***
Baum, Christopher. 2006. Chapter 9, pp.219-226, from: An Introduction to Modern Econometrics Using Stata. College Station, TX: Stata Press.

March 4. No class: Spring Break

***First draft of data management due***
Baum, Christopher. 2006. Chapter 9, pp.226-232, from: An Introduction to Modern Econometrics Using Stata. College Station, TX: Stata Press.

***Mini-assignment 2: Answers to article questions due 10 AM***

**March 25. Model Building Strategies for Mixed Effects Models**

**April 1. Diagnostics for Mixed Effects Models**

**April 8. Panel Data Models for Categorical and Count Data**
***Mini-assignment 3: Answers to article questions due 10 AM***

**April 15. Sample Selection and Endogeneity Biases**
***First draft of data analysis and article due***

**April 22. Discrete-time Event History Analysis**

**April 29. Diagnostics for Discrete-time Event History Models**
***Mini-assignment 4: Answers to article questions due 10 AM***

***Final drafts of data analysis and article due May 13***