

SOCY7706: LONGITUDINAL DATA ANALYSIS

Instructor: Natasha Sarkisian
Preferred pronouns: she/her/hers
Email: natasha@sarkisian.net
Phone: (617) 755-3178
Office: McGuinn 417
Zoom: <https://bccte.zoom.us/j/95791630926>

Class time: Wednesdays 9-11:30am
Class location: O'Neill 245
Office hours: Mondays 11:30-12:30 and by appointment
Mailbox: McGuinn 426
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 SOCY7703: 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 handling and describing longitudinal data, two-wave panel data models, fixed, random, and mixed effects models, panel models for categorical data, and missing data issues in longitudinal analyses.

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://bcapps.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

In-Class Activities: For each topic covered in the course, I will give a lecture involving a demonstration in Stata, and you will get a chance to follow along and do all the necessary steps in Stata on your laptop. We will also discuss article examples for each major topic.

Communication: 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 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 if additional materials are posted on the website and send other announcements from time to time.

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 send me an email with at least one sentence of reaction to that class meeting, indicating what you learned, or something you liked or did not like, 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. Please be honest in your comments – if something is unclear or doesn't work for you, I really do want to know about that and will not penalize you in any way! Make sure to submit your feedback every week – this is a part of the class participation component of your grade.

Coursework: Throughout the course, you are expected to do all your coursework on time. Ordinarily, no late assignments will be accepted, unless you have a valid (and documented) excuse.

Academic Integrity: It is your obligation to be fully aware of the Boston College policies on academic honesty. ANY violation may subject the offender to severe penalty, including course failure. If you are not familiar with the Boston College policy on academic honesty, see: <http://www.bc.edu/offices/stserv/academic/integrity.html>

Disability Accommodation: If you have a disability and will be requesting accommodations for this course, please register with either Dr. Kathy Duggan (dugganka@bc.edu), Associate Director, Connors Family Learning Center (learning disabilities or AHD) or Dean Rory Stein (rory.stein@bc.edu), Assistant Dean for students with disabilities (all other disabilities, including temporary disabilities). Advance notice and appropriate documentation are required for accommodations.

Health and Wellness: If you are feeling stressed, having challenges managing your time, sleep, or making choices around alcohol and food, the Office of Health Promotion (OHP) offers Individual and Group Health Coaching appointments with a trained Health Coach. Please reach out by going to OHP website or walk over to Gasson 025 and talk with a staff member. Be Well.

Diversity: It is my intent that students from all diverse backgrounds and perspectives be well-served by this course, that students' learning needs be addressed both in and out of class, and that the diversity that the students bring to this class be viewed as a resource, strength, and benefit. It is my intent to present materials and activities that are respectful of diversity: gender identity, sexuality, disability, age, socioeconomic status, ethnicity, race, nationality, religion, and culture. Your suggestions are encouraged and appreciated. Please let me know ways to improve the effectiveness of the course for you personally, or for other students or student groups.

COURSE REQUIREMENTS AND GRADING

All the required readings will be available on electronic reserve on Canvas.

There will be five major assignments for this course, each 15% of your grade. These assignments will involve utilizing the techniques we are learning. For each technique, you will run all the necessary analyses, conduct diagnostics and apply remedies, and interpret the results, using additional interpretation tools (standardized coefficients, graphs, etc.) as needed to assist in interpreting your findings. You will submit an annotated log that will contain the output (with your comments) for all of the tasks that you will perform for each 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. I will provide the datasets for all the major assignments. If, however, you have a research project that involves longitudinal data analysis, you can obtain permission from me to use that dataset for one or more of your assignments instead. Please contact me in the beginning of the course if you intend to use a different dataset and send me the information about the dataset and your project.

In addition to that, there will be four mini-assignments (5% of your grade each) that will involve answering questions on a published article that utilizes some of the same methods that will be discussed in class or doing an interpretation exercise; questions will be posted on the course website one week prior to the assignment due date. These assignments will be graded pass/fail, with the grades reflecting demonstrated effort rather than correct answers (correct answers will, however, be provided during in-class discussion).

All assignments will be submitted electronically. Small files can be sent by email; any large files should be submitted using Google Drive, Dropbox or another file sharing website.

Finally, the remaining 5% of your grade will be based on your in-class participation – asking and answering questions, participating in discussions, submitting weekly feedback, etc.

COURSE OUTLINE

August 31. Introduction to Longitudinal Data Analysis and to Stata.

September 7. Longitudinal Data Management.

Ch. 7, 9 from Mitchell, Michael N. 2020. *Data Management Using Stata: A Practical Handbook*. 2 ed. Stata Press.

September 14. Describing Longitudinal Data.

Andreß, Hans-Jürgen, Katrin Golsch, and Alexander W. Schmidt. 2013. Chapter 3 from: *Applied Panel Data Analysis for Economic and Social Surveys*. Springer.

September 21. Two-Wave Panel Data Analysis.

****Assignment 1 (Data Management and Description) due by 11:59pm****

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

Johnson, David. 2005. “Two-Wave Panel Analysis: Comparing Statistical Methods for Studying the Effects of Transitions.” *Journal of Marriage and Family* 67(4):1061-75.

September 28. Fixed Effects Models.

****Mini-assignment 1: Answers to Laurence article questions due by 9am****

Worrall, John L. 2008. An Introduction to Pooling Cross-Sectional and Time Series Data. Chapter 15 from *Handbook of Longitudinal Research: Design, Measurement, and Analysis* edited by Scott Menard. Academic Press.

Laurence, James. 2015. (Dis)placing Trust: The Long-term Effects of Job Displacement on Generalised Trust over the Adult Lifecourse. *Social Science Research* 50: 46-59.

October 5. Fixed Effects Models: Diagnostics.

****Assignment 2 (Two-Wave Panel Data) due by 11:59 pm* ***

Allison, Paul D. 2009. Chapter 2 and Appendix 1 (portion related to Ch. 2), from: *Fixed Effects Regression Models*. Sage Publications.

Baum, Christopher. 2006. Chapter 9, pp.219-226, from: *An Introduction to Modern Econometrics Using Stata*. College Station, TX: Stata Press.

October 12. Random Effects Models.

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

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

October 19. Introduction to Mixed Effects Models.

****Mini-assignment 2: Answers to Jorgenson and Clark article questions due by 9am****

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

Hedecker, Donald. 2004. An Introduction to Growth Modeling. Chapter 12 from: David Kaplan (Ed.), *The Sage Handbook of Quantitative Methodology for the Social Sciences*. Thousand Oaks, CA: Sage Publications.

Jorgenson, Andrew and Brett Clark. 2009. The Economy, Military, and Ecologically Unequal Exchange Relationships in Comparative Perspective: A Panel Study of the Ecological Footprints of Nations, 1975–2000. *Social Problems* 56 (4): 621–646.

October 26. Model Building Strategies for Mixed Effects Models.

***Assignment 3 (Fixed and Random Effects Models) due by 11:59 pm ***

Hox, Joop. 2010. Chapters 3 and 4 from *Multilevel Analysis: Techniques and Applications*. 2nd edition. Routledge.

November 2. Diagnostics for Mixed Effects Models.

Snijders, Tom A. B., and Roel J. Bosker. 2012. Chapter 10 from *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Thousand Oaks, CA: Sage.

November 9. Interpreting Mixed Effects Models.

Mini-assignment 3: Answers to Farkas article questions due by 9am

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.

November 16. Missing Data in Longitudinal Research.

Mini-assignment 4: Coefficient interpretation exercise due by 9am

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

Acock, Alan C. 2005. Working with Missing Values. *Journal of Marriage and Family* 67: 1012–1028.

November 23. No class.

November 30. Missing Data in Longitudinal Research (continued).

Twisk, Jos, and Wieke de Vente. 2002. Attrition in Longitudinal Studies: How to Deal with Missing Data. *Journal of Clinical Epidemiology* 55: 329–37.

December 7. Panel Data Models for Categorical Data.

Assignment 4 (Mixed Effects Models) due by 11:59 pm *

Snijders, Tom A. B., and Roel J. Bosker. 1999. Chapter 14 from *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. Thousand Oaks, CA: Sage.

December 14. Sample Selection and Endogeneity Biases

Assignment 5 (Missing Data) due by 11:59 pm

Fu, Vincent Kang, Christopher Winship, and Robert D. Mare. 2009. Sample Selection Bias Models.

Chapter 18 from *Handbook of Data Analysis*, edited by Melissa A. Hardy and Alan Bryman. Sage.

Greenland, S. 2000. An Introduction to Instrumental Variables for Epidemiologists. *International Journal of Epidemiologists*, 29, 722–729.