SOC6708H: Advanced Data Analysis: Analysis of Longitudinal/Panel Data

Fall 2018
Lecture: Fridays, 10 am – 12 pm
Lecture Location: Room 240
Lab: Fridays, 12 pm – 1 pm
Lab Location: Room 36

Instructor: Andrew Miles
Office location: 364
Email address: andrew.miles@utoronto.ca
Office hours: by appointment

Teaching Assistant: TBD

Course Description

Panel data are any data that are collected on (or from) the same people (or other units) over time. This course teaches fundamental methods in the analysis of panel data, focusing on ways of using panel data to strengthen causal claims, and on modeling change over time. Topics include random effects models, fixed effects models, growth curve models, dynamic models (i.e., models with lagged dependent variables), and special issues confronted in the analysis of panel data. Students will make use of both multilevel modeling and structural equation modeling frameworks. No prior experience with either is required, though students should be familiar with regression.

Textbooks and Other Materials

The textbooks for the course are:

Allison, Paul D. 2009. *Fixed Effects Regression Models*. Los Angeles, C.A.: Sage.

Grimm, Kevin J., Nilam Ram, and Ryne Estabrook. 2017. *Growth Modeling: Structural Equation and Multilevel Modeling Approaches*. New York: Guilford Press.

Both texts will be available in the bookstore.

Statistical Software

All analyses will be performed using Stata. Stata is available in the sociology computer lab (room 36), at locations listed here (https://mdl.library.utoronto.ca/technology/statistical-software), and possibly at other locations as well. Stata can also be purchased from U of T's software office (http://sites.utoronto.ca/ic/software/alphalist.html). If you decide to buy Stata, be

sure to select the student pricing, and purchase either Stata/IC or Stata/SE. You need Stata 13 or above (current version is 15).

If you are not familiar with Stata, I strongly encourage you to become familiar with its basic operations prior to the first day of class. If you are familiar with other statistical software packages you should be able to learn the basics quickly. Good references include:

http://www.ats.ucla.edu/stat/stata/

Longest, Kyle C. 2012, 2015. *Using Stata for Quantitative Analysis* (1st or 2nd editions). [2012 edition available as a digital holding in U of T Libraries]

Class Format

Lecture time will be devoted to presenting statistical concepts. You should be prepared to take notes by hand, as laptops are not permitted during class (see section on classroom etiquette below). To help you focus on learning the material rather than scrambling to take notes, simplified versions of lecture slides will generally be posted prior to lecture. It is recommended that you print a copy of these slides and bring them with you to lecture to take notes on. We will take a 10-minute break about halfway through each lecture period. Lectures are designed to lay down foundational concepts, with practical instruction in data analysis being given during the lab tutorials.

Lab Tutorials

The goal of lab tutorials is to review material from lecture as needed, and learn to apply statistical concepts using software designed for statistical analysis. Labs are led by the course TA, and will typically include an in-lab exercise to help you practice your data analytic skills.

Weekly Reading

Reading assigned chapters from the textbook and completing assignments provide useful repetition and practice of concepts and skills presented during lecture. This is especially important when dealing with advanced topics like panel data analysis.

Lectures allow for dialogue and clarification of key concepts. For this reason, I believe lectures are generally better as the first introduction to material. Consequently, in this course reading is intended as review, and should occur *after* lecture, but before the following lecture (though I will not insist on this if students wish to read before lecture). In this way, lectures can provide a framework that will help students better understand and retain the material they read.

Evaluation and Grading

Coursework is weighted as follows in calculating the final grade.

Weekly assignments 45%

Research briefs	30%
Final paper	25%

Final grades will be assigned using the grading scale below, taken from the University Assessment and Grading Practices Policy.

Percentage	Grade
90-100	A+
85-89	A
80-84	A-
77-79	B+
73-76	B
70-72	B-
0-69	Fail

Description of Evaluation Components

Weekly Assignments

There will be an assignment posted on Blackboard each week that will allow you to review key concepts and practice data analysis. Assignments will generally include a computing component using Stata, the how-to's of which will be taught in lab sessions. Assignments can be found under *Course Materials → Assignments*.

Assignments will be graded as pass/fail. Failing grades will be given for incomplete work, or for work that exhibits little evidence of a good faith effort at engagement with the material.

Research Briefs

You will complete three research briefs during the course. The goal of each research brief is to give you a chance to apply the skills you've been learning to a question that is of interest to you.

<u>What to Do</u>

1. Find a topic that interests you, and come up with some aspect of that topic that you'd like to learn more about. Keep in mind that you will need to be able to answer this question using techniques you've learned in class.
2. Find some data that will allow you to answer your question. NOTE: If you are having trouble finding data, the easiest thing to do is probably to pick a new question, one that can be answered using the data you have access to.
3. Figure out which variables you will need to answer your question.
4. Determine which statistical technique you can use to answer your question.
5. Run the analysis in Stata.
6. Interpret the results. That is, what is the answer to your question? How do you know/what evidence do you have?

7. Write up the results in the Research Brief (see guidelines below)
8. Turn the Research Brief in
9. Celebrate your statistical prowess in a manner of your choosing

Specific Requirements

Research Brief #1 will need to use one or more techniques from weeks 1-4 of the class.

Research Brief #2 will need to use one or more techniques from weeks 5-8.

Research Brief #3 will need to use one or more techniques from weeks 9-11.

A Research Brief should be no longer than 2 pages, with normal (12 point) sized font. Often a single page will be sufficient.

Include at least one table, graph, or figure to help display your results.

Use the layout described below for your Research Brief (i.e., all the same headings in the same order).

Layout

Your Research Brief should include

- **Question:** a clearly worded question
- **Data:** a description of the data you are using, including a description of all variables and how they are coded
- **Plan of Analysis:** a description of your analysis (i.e., how are you going to use the data to answer the question?)
- **Results:** a presentation of the results, including a table/figure/graph
- **Discussion:** Provide an interpretation of the results – that is, given the results, what is the answer to the question? Are there any other possible interpretations? What limitations might there be in your data or analysis that affect how well we can answer the question?

Note that you are welcome to use the same general question and data for multiple research briefs, as well as the final paper (described below).

Links to submit Research Briefs can be found on Blackboard under *Course Materials* → *Research Briefs*.

Grading

Research Briefs will be graded as follows:

Score	Meaning
3	Excellent
2	Adequate

1	Needs attention

For the purposes of final grade calculation, scores on Research Briefs will be averaged and assigned the following values:

Score range	Grade	Numeric value
0	Fail	0
0-0.09	Fail	50
1-1.3	B-	71
1.4-1.6	B	75
1.6-1.7	B+	78
1.8-1.9	A-	82
2-2.4	A	87
2.5-3	A+	95

Final Paper

The final paper gives you an opportunity to apply what you have learned to question of your choosing. It must use one (or more) of the methods discussed in class. It should be written in research paper format with two exceptions.

1. The front end (introduction and theory section) should be condensed considerably. Essentially, just give the introduction part of a normal research paper (where you state why the paper is important and what you will be doing), and skip the theory section where you substantiate all the claims you made in the introduction. A rough guideline might be 2 double-spaced pages, or 1 single-spaced page for this section.
2. The discussion should focus on points relevant to the analyses of the paper. That is, what did you find? How strong was the evidence? What limitations exist in your research design/data and how might they be overcome in future research? You can restate why your question is important, but don't spend time discussing all of the theoretical implications of your work, how it relates to past work, etc. – i.e., all the big picture implications of your work.

In writing the methods and results sections, pay particular attention to:

1) developing an appropriate plan of analysis for your question
2) performing the analysis correctly
3) interpreting the results correctly

Alternately, you may choose to write a paper on a methodological topic related to panel data analysis. This might take the form of a review of existing literature on the topic, or an original work designed to contribute to the methodological literature in some way (e.g., a simulation study). This option requires instructor approval, and consultation with the instructor to choose an appropriate topic.

The final paper will be graded using the same scale as the research briefs.

Course Schedule

All readings are from the textbook unless otherwise noted.

Week	Date	Topic	Reading	Due this week (by the start of lecture)
1	Sept 14	Introduction to Longitudinal/Panel Data	Singer, Judith D. and John B. Willett. 2003. *Applied Longitudinal Data Analysis: Modeling Change and Event Occurrence*. Oxford University Press. Pages 16-33. Grimm et al (2017): 29-37 Stata manual entry for *simulate*	
2	Sept 21	Directed Acyclic Graphs and Structural Equation Modeling	Berkout, Olga V., Alan M. Gross, and John Young. 2014. "Why So Many Arrows? Introduction to Structural Equation Modeling for the Novitiate User." *Clinical Child and Family Psychology Review* 17(3):217–29. Kline (2016) chapter 7:198-206; 217-219 Kline (2016) chapter 8 Bonus: intro4 from Stata SEM manual	assignment 1
3	Sept 28	Cross-Lagged Panel Models	Selig, James P. and Todd D. Little. 2012. "Autoregressive and Cross-Lagged Panel Analysis for Longitudinal Data." Pages 265-278 in *Handbook of Developmental Research Methods*, edited by B. Laursen, T. D. Little, and N. A. Card. New York: Guilford Press.	assignment 2
4	Oct 5	Random effects models	Hox, Joop J. 2010. *Multilevel Analysis: Techniques and Applications*. 2nd ed. New York: Routledge. Chapter 5.	assignment 3
5	Oct 12	Fixed Effect Models	Allison (2009) chapters 1 and 2	assignment 4 research brief #1
6	Oct 19	Fixed Effect Logit and Count Models	Allison (2009) chapters 3 and 4	assignment 5
7	Oct 26	FE Estimation with Dynamic Panel Models	TBD	assignment 6
8	Nov 2	Fixed Effects models using SEM	Bollen, Kenneth A. and Jennie E. Brand. 2010. "A General Panel	assignment 7

			Model with Random and Fixed Effects: A Structural Equations Approach." *Social Forces* 89(1):1–34. Allison, Paul D., Richard Williams, and Enrique Moral-Benito. 2017. "Maximum Likelihood for Cross-Lagged Panel Models with Fixed Effects." *Socius* 3:1-17.	
9	Nov 9	Growth Models	from Grimm et al (2017)	assignment 8 research brief #2
10	Nov 16	Growth Models using SEM	from Grimm et al (2017)	assignment 9
11	Nov 23	Advanced Growth Modeling	from Grimm et al (2017)	assignment 10
12	Nov 30	Panel Conditioning	Halpern-Manners, Andrew, John Robert Warren, and Florencia Torche. 2017. "Panel Conditioning in the General Social Survey." *Sociological Methods & Research* 46(1):103–24.	assignment 11 research brief #3
	Dec 12	Final paper due by midnight		

Every attempt will be made to follow this schedule, but it is subject to change at the discretion of the instructor.

Procedures and Rules

Late Work

All assignments are due by the beginning of lecture on the date listed in the syllabus. Late work will not be accepted.

Late penalties can be waived for a legitimate reason with proper documentation (e.g., illness, family emergency, religious observance, but NOT family vacations, weddings, I want a long weekend, etc.). Where possible, these arrangements must be made in advance of the missed work.

Grade Appeals

If you believe that a mistake was made in grading your work, you may appeal the grade by submitting a written explanation of why you think your mark should be altered to the instructor. The instructor will then re-grade your work with the additional information in mind. Although in most cases re-grading results in a higher mark, this is not guaranteed, and your mark might go down.

Academic Integrity

You are expected to abide by the University's standards of academic integrity, which can be found in the "Code of Behaviour on Academic Matters" (http://www.governingcouncil.utoronto.ca/policies/behaveac.htm). Plagiarism or other violations will be addressed in accordance with University guidelines. Please be cautious in this matter, as the penalties for academic misconduct can be quite severe.

Working With Other Students

Working with other students is often a useful way to learn statistics. You are therefore encouraged (but not required) to work with other class members in completing assignments (including research briefs). However, each student must complete and submit his/her own work, written in his/her own words. Students who work together on class work should also indicate whom they worked with on each assignment (if anyone). These steps guard against situations where a student's academic integrity might be called into question (see section on Academic Integrity).

Classroom Etiquette

Students are expected to arrive at class on time. If you need to leave during lecture, please do so in a way that will minimize disruption to the class.

Laptop computer are not to be used during class time. *This means that you should be prepared to take notes by hand*. Notes may be taken on tablet devices, but these should not be used for gaming, checking email, or any of the many other things that might distract from classroom engagement. Cellphones should not be used during class time, and should be turned off or set to silent until class is over.

Attendance

Data analysis is a skill, and like any skill mastering requires time on task. Attendance is therefore mandatory at all lectures and lab tutorials. Any absences should be cleared with the instructor. More than two absences at either lecture or lab may result in a reduction in a student's final grade, usually one half grade per absence beyond the two (e.g., from A to A-). These grade reductions will be at the instructor's discretion.