SOC 570: MULTILEVEL AND LONGITUDINAL DATA ANALYSIS
Spring 2020
Mondays 1:00 – 3:50 pm
Department of Sociology
Davison Hall, Room 129 (Seminar Room)

Professor:	Lei Lei, Ph.D.
Office:	039 Davison Hall
Office Hours:	Mondays 4:00 – 5:00 pm or by appointment
E-mail:	llei@sociology.rutgers.edu

Course overview

This course is about applied multilevel and longitudinal modeling in social sciences. Multilevel models have many other names, including hierarchical models, random-effect models, or mixed-effect models. Longitudinal data are also referred to as panel data, repeated measures, or cross-sectional time series. The focus of this course will be regression modeling when data are clustered in some way. In cross-sectional settings, students may be nested in schools, people in neighborhoods, employees in firms, and twins in twin pairs. Longitudinal data are clustered because multiple observations over time are nested within units, which are often individuals.

This course will cover topics including 1) management of multilevel and longitudinal data; 2) two-level models including random-intercept and random-coefficient models; 3) dynamics models and fixed-effect models for causal inference; 4) models for longitudinal data, including marginal models and growth curve models; 5) models with nested and crossed random effects; 6) multilevel models for categorical/binary responses; 7) multilevel models for counts data.

A central goal of the course will be to provide you with a conceptual and practical understanding of multilevel models, so that you can apply the methods to **your own data** and **research interests**. You will learn through lectures, lab exercises, assignments, and a final research project in which you apply the methods to your own research. *Stata* will be the main software used in this course, but you can complete the exercises and assignments using a software that you prefer.

Textbook

Sophia Rabe-Hesketh and Anders Skrondal. (2012). *Multilevel and Longitudinal Modeling Using Stata, 3rd Edition (Volume I: Continuous Responses; Volume II: Categorical Responses, Counts, and Survival)*. College Station, TX: Stata Press.

Davide Garson. (2020). *Multilevel Modeling: Applications in Stata, IMB SPSS, SAS, R & HLM*. Thousand Oaks, CA: Sage.

Requirements

1. <u>Assignments</u>: Eight homework assignments will be made throughout the semester to give you practical experience working with the topic covered in the prior class. You are encouraged to work together on these exercises…but everyone must hand in his or her own completed assignment. **Each assignment should be turned in electronically via Sakai by noon of the following Monday.** The assignments are worth 50% of your course grade.

2. <u>Empirical paper discussion</u>: Two students will be assigned as discussion leaders who will present to the class an empirical research article in certain weeks. The presentation should focus on introducing the research questions of the paper, explaining the data structure, reviewing the methods used to answer the research questions, and interpreting the results. It will be helpful for us to see how multilevel/longitudinal data analysis methods and results are presented in research articles published in top social sciences journals. Leading the discussion will account for 5% of your course grade.

3. <u>Research paper</u>: A central goal of the class is for students to gain the ability to apply the methods to analyze their own data and answer their research questions. You will be asked to develop a four-page research proposal by the middle of the semester, complete a research paper by the end of the semester (around 20 pages, excluding tables, figures, and references), and present the paper in class. Although the research paper can be brief in the theory section, you will need to clearly state the hypotheses, describe the data and methods, properly present and interpret the results, and draw meaningful conclusions. The research proposal (5%), the presentation (5%), and the final research paper (35%) together are worth 45% of your course grade.

4. <u>Attendance policy</u>: The material covered in this class is incremental and cumulative. That is, each week the presentation of new material will build on what we've covered previously. Therefore, regular attendance is essential and you must keep on track with your readings and assignments (meticulous notes and clear organization will help, too). **It is imperative that you do not fall behind when learning this material**. During the semester you will likely have some competing obligations, both professional and personal. If you know that you are going to miss a class, please let me know so that we can plan ahead.

Useful Websites

1. All datasets used in the textbook are available from http://www.stata-press.com/data/mlmus3.html You can download the dataset into a local directory on your computer.

   Alternatively, individual datasets can be loaded directly into net-aware Stata by specifying the complete URL. For example,

   > . use http://www.stata-press.com/data/mlmus3/pefr

2. Solutions to exercises that are available to readers can be downloaded from http://www.stata-press.com/books/mlmus3-answers.html.

3. Additional readings, assignment questions, handouts, and other course-related materials will be posted on **Sakai.**

Diversity Statement

The Rutgers Sociology Department strives to create an environment that supports and affirms diversity in all manifestations, including race, ethnicity, gender, sexual orientation, religion, age, social class, disability status, region/country of origin, and political orientation. We also celebrate diversity of theoretical and methodological perspectives among our faculty and

students and seek to create an atmosphere of respect and mutual dialogue. We have zero tolerance for violations of these principles and have instituted clear and respectful procedures for responding to such grievances

Academic Misconduct

Academic honesty is a subject that I take very seriously. I encourage all students to familiarize themselves with the Rutgers University policies and procedures on academic honesty, available at http://academicintegrity.rutgers.edu/academic-integrity-at-rutgers. All violations of academic integrity, for example, copying or plagiarizing others' work for your written assignments, will be referred to the appropriate authorities and sanctioned accordingly.

Students with Disabilities

In accordance with University policy, if you have a documented disability and require accommodations to obtain equal access in this course, please contact me at the beginning of the semester. Students with disabilities must be registered with Office of Student Disability Services and must provide verification of their eligibility for such accommodations.

Course Outline

The dates indicated are tentative and may change based upon how this class proceeds. Some topics may take a bit more time and some may take less time than indicated below.

Week 1. January 27
Introduction to the course
Brief review of regression models and Stata programing
    Readings:
    Rabe-Hesketh and Skrondal, Chapter 1
    Garson, Chapter 1

Week 2. February 3
Data preparation and management
Variance-components models and estimation using Stata
    Readings:
    Rabe-Hesketh and Skrondal, Chapter 2.1-2.10
    Garson, Chapter 2 (Overview, model specification, sample size, balanced and unbalanced designs, linearity and nonlinearity) and Chapter 3

Week 3. February 10
Variance-components model continued – more examples
Hypothesis test and confidence intervals and parameter estimation
    Readings:
    Garson, Chapters 4 and 5
    (Assignment 1 due)

Week 4. February 17
Random-intercept models with covariates and estimation using Stata
    Readings:
    Rabe-Hesketh and Skrondal, Chapter 3.1-3.9
    Garson, Chapter 6
    (Empirical paper discussion 1)
    (Assignment 2 due)

Week 5. February 24
Random-coefficient models
    Readings:
    Rabe-Hesketh and Skrondal, Chapter 4
    Garson, Chapter 7
    (Empirical paper discussion 2)
    (Assignment 3 due)

Week 6. March 2
Models for longitudinal and panel data
Random-intercept model and lagged-response models
    Readings:
    Rabe-Hesketh and Skrondal, Chapter 5

(Empirical paper discussion 3)
(Assignment 4 due)

Week 7. March 9
Mean structure and covariance structures
Generalized Estimating Equation (GEE)
    Readings:
    Rabe-Hesketh and Skrondal, Chapter 6
    (Assignment 5 due)

Week 8. March 23
Growth curve models
    Readings: Rabe-Hesketh and Skrondal, Chapter 7
    (Empirical paper discussion 4)
    (Research Proposal due)

Week 9. March 30
Three-level models
    Readings: Rabe-Hesketh and Skrondal, Chapter 8
    Garson, Chapter 8
    (Empirical paper discussion 5)
    (Assignment 6 due)

Week 10. April 6
Crossed random effects
    Readings:
    Rabe-Hesketh and Skrondal, Chapter 9
    Garson, Chapter 11
    (Empirical paper discussion 6)

Week 11. April 13
Multilevel logistic regression models – random-effect and fixed-effect
    Readings:
    Rabe-Hesketh and Skrondal, Chapter 10
    (Empirical paper discussion 7)
    (Assignment 7 due)

Week 12. April 20
Multilevel models for ordinal and count outcomes
    Readings:
    Rabe-Hesketh and Skrondal, Volume 2, Chapters 11, 12, 13
    (Empirical paper discussion 8)
    (Assignment 8 due)

Week 13. April 27
    Student presentations of research paper

Paper due on Friday May 8th, 11:59pm

Empirical Papers for Discussion

Papers	Student 1	Student 2
1. Variance component model and two-level model De Clercq, B., Vyncke, V., Hublet, A., Elgar, F. J., Ravens-Sieberer, U., Currie, C., ... & Maes, L. (2012). Social capital and social inequality in adolescents' health in 601 Flemish communities: A multilevel analysis. *Social Science & Medicine*, *74*(2), 202-210.		
2. Random coefficients and cross-level interactions Huffman, M. L., & Cohen, P. N. (2004). Racial wage inequality: Job segregation and devaluation across US labor markets. *American Journal of Sociology*, *109*(4), 902-936.		
3. Lagged dependent variable model Shandra, J. M., Shandra, C. L., & London, B. (2010). Do non-governmental organizations impact health? A cross-national analysis of infant mortality. *International Journal of Comparative Sociology*, *51*(1-2), 137-164.		
4. Growth curve model Chen, F., Yang, Y., & Liu, G. (2010). Social change and socioeconomic disparities in health over the life course in China: A cohort analysis. *American sociological review*, *75*(1), 126-150. Cheng, S. (2014). A life course trajectory framework for understanding the intracohort pattern of wage inequality. *American Journal of Sociology*, *120*(3), 633-700.		
5. Fixed-effect model Lu, Y. (2012). Household migration, social support, and psychosocial health: The perspective from migrant-sending areas. *Social Science & Medicine*, *74*(2), 135-142. Cheng, S. (2016). The accumulation of (dis) advantage: The intersection of gender and race in the long-term wage effect of marriage. *American Sociological Review*, *81*(1), 29-56.		
6. Cross-classified model Sykes, B., & Musterd, S. (2011). Examining neighbourhood and school effects simultaneously: what does the Dutch evidence show? *Urban Studies*, *48*(7), 1307-1331.		
7. Multilevel logistic regression Flaherty, J., & Brown, R. B. (2010). A multilevel systemic model of community attachment: Assessing the relative importance of the community and individual levels. *American Journal of Sociology*, *116*(2), 503-542.		
8. Multilevel Poisson regression Huebner, B. M. (2003). Administrative determinants of inmate violence: A multilevel analysis. *Journal of Criminal Justice*, *31*(2), 107-117.		