

Sociology 362: Statistics for Sociologists III
University of Wisconsin-Madison
Spring 2020

[Last Updated: January 20, 2020]

Who and Where?

Lecture Time: Tuesday & Thursday 2:30-3:45PM

Lecture Location: Van Vleck B115

Instructor: Prof. Felix Elwert

Office: 4426 Social Science Building

Office Hours: By electronic appointment via:
<https://felix-elwert.youcanbook.me>
Tuesday & Wednesday 9:30-10:30am

Email: elwert@wisc.edu

Teaching Assistant: Jia Wang

Office: 7110 Social Science

Office Hours: Monday & Wednesday 12-1pm or by appointment via email.

Email: jia.wang@wisc.edu

Course Website: <https://canvas.wisc.edu/courses/186152>

Administration Information

Credits: 3

Credit hours are met by 150 minutes of lecture plus 90 minutes of lab per week, plus a minimum of 7 hours of out-of-class student work per week.

Course Designations:

Level – Advanced; L&S Credit - Counts as Liberal Arts and Science credit in L&S; Grad 50% - Counts toward 50% graduate coursework requirement; Honors - Accelerated Honors.

Prerequisites: Junior standing and Soc 361 or equivalent, or consent of instructor.

Instructional Model: All Face-to-face.

This is the third course in the statistics sequence for sociologists. It is required for CAR and sociology Ph.D. students; it is elective for others.

Course Description

This course aims to give students a foundation for understanding the purposes and methods of modern quantitative research in sociology. We will engage with two overarching topics.

1. Learn the key strategies for drawing causal inferences from experimental and observational data (“**identification**”).
2. Develop a deeper technical understanding of the main statistical techniques for empirical research in sociology (“**estimation**”).

We will move back and forth between identification and estimation throughout the course, since different identification strategies may require different estimation approaches. For example, when analysts believe that they can identify causal effects by *adjusting* or controlling for observed covariates, then they will resort to *regression* and *matching* estimators. When adjustment for observed covariates is not sufficient (perhaps because of unobserved confounding factors), they may defend specific *functional form assumptions* and resort to *difference-in-difference* or *fixed-effects estimators*. Alternatively, when they can pinpoint the source of *random variation* in treatment assignment, they may use *regression-discontinuity* or *instrumental variables estimators*. In sum, different identification strategies imply different statistical estimators. Therefore, Soc 362 teaches both identification and estimation.

Throughout the course, *linear regression* will serve as the traditional point of reference for all estimation topics. In contrast to Soc 361, however, we move beyond OLS by considering important extensions for non-metric outcomes (*general linear models*), nested data structures (*hierarchical or mixed models*), and censored data (*tobit*).

Inevitably, we will cover different topics at different levels of technical depth. We will emphasize concepts and intuitions where these are sufficient, and we will additionally develop more formal notation where it is useful for linking to more advanced courses.

Since quantitative research is an interdisciplinary enterprise, we will liberally draw on methodological contributions from econometrics, statistics, biostatistics, political methodology, and computer science.

Learning Outcomes

By the end of this course, students will be able to recognize what statistical techniques are appropriate to extract answers for common sociological questions from data in a variety of realistic scenarios.

Students who will primarily *consume*, but not produce, quantitative empirical research will acquire the necessary tools to become critical readers of the majority of quantitative research in sociology.

Students who want to *produce* quantitative empirical research, will learn the tools to execute competent work and be able to decide which of the more specialized methodological courses they should take next in order to participate at the research frontier.

Grading

Exams (65%): Students will complete 2 midterm exams and one cumulative final exam. Exams cover all material presented in lecture and strongly build on the problem sets in content and form. To respect that everybody can have a bad day, we will give greater weight to your better midterm (20%) and less weight to your weaker midterm (15%). The final exam counts 30%.

There is no possibility of make-up midterms. If you have to miss a midterm, please provide a compelling reason *before the beginning of the exam*, in which case your other midterm will count more (35%). If both midterms are missed, the final exam will constitute the entire exam grade (65%).

For each exam, you may bring *one single-sided sheet* of your own handwritten notes (for formulas etc.).

Problem sets (30%): Students will complete nine computational and conceptual problem sets, some of which contain data-analysis exercises. Problem sets are typically due one week after they have been distributed. Late assignments receive a score of zero. You can drop one assignment without penalty. Assignments will be graded on effort for partial credit and for correctness to achieve full credit.

Attendance (3%): Lecture and section attendance are required and accounts for 3% of the grade. Students may miss up to a total of 2 lecture and 2 section meetings and still receive full attendance credit. Attendance credit drops to 0% with the fifth absence.

Feedback (2%): We want your feedback! After every lecture (i.e. before the next lecture), students will submit quick feedback on their comprehension of the readings and lecture. The feedback should consist of a concise list comprising the two things they understood best, and the two things they understood least. Please submit this feedback on the course website under "Discussions." Students who submit feedback in at least 10 weeks will receive full credit (2%).

Extra Credit (3%): Students may submit well-organized and reasonably comprehensive glossaries of key terms and acronyms covered in lecture and the readings for extra credit. The glossary should be typed or neatly hand-written. Students who hand in one glossary before each of the three exams will receive 3% extra credit.

Readings

We post required and recommended readings on the course website. Students must complete all required readings before lecture. Students should revisit their readings, and the lecture script again after lecture. I *cannot stress this enough*: in order to learn statistics, you must read the material *repeatedly*.

This course does not follow any one textbook. Instead, we are assembling the readings from a journal articles, published textbooks, and the *de facto* draft of a textbook (Chris Winship's lecture scripts) to cover the material you need at an appropriate technical level.

Students are encouraged to buy at least one of two excellent and highly accessible books on the logic of empirical research, which focus on foundational insights and guiding intuition that will be useful for both producers and consumers of statistics, whatever their level of technical sophistication.

1. Rosenbaum, Paul. 2017. *Observation and Experiment. An Introduction to Causal Inference*. Harvard University Press: Cambridge, MA.
2. Angrist, Joshua D., and Jörn-Steffen Pischke. 2015. *Mastering Metrics: The Path from Cause to Effect*. Princeton University Press: Princeton, NJ.

Additionally, graduate and CAR students who plan to produce quantitative research should obtain Angrist and Pischke's justly famous *Mostly Harmless Econometrics* (the graduate version of #2 above). Some of the material in this book is presented at a higher technical level than Soc 362, but it is the minimal level to which you will need to rise eventually to produce state-of-the-art quantitative research on your own.

3. Angrist, Joshua D., and Jörn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press: Princeton, NJ.

Several good texts on social science statistics in Stata exist. I recommend:

4. Baum, Christopher F. 2006. *An Introduction to Modern Econometrics Using Stata*. Stata Press: College Station, TX.

How much should you read? My advice is to *read more than required, especially if you struggle with the material*. Reading the same text twice is good; but reading the same material twice from different authors may be even better.

Do not be tricked by apparently low page counts (e.g. 30 pages per week). Reading statistics is different from reading prose. It is important to read slowly and carefully. The devil is in the details. Take note of symbols, formulas, and definitions. Highlight and annotate your text. Excerpt your readings by hand and on paper. It is perfectly normal to spend 2-4 or hours on 10 pages of technical reading.

Academic Integrity

By enrolling in this course, each student assumes the responsibilities of an active participant in UW-Madison's community of scholars in which everyone's academic work and behavior are held to the highest academic integrity standards. Academic misconduct compromises the integrity of the university. Cheating, fabrication, plagiarism, unauthorized collaboration, and helping others commit these acts are examples of academic misconduct, which can result in disciplinary action. This includes but is not limited to failure on the assignment/course, disciplinary probation, or suspension. Substantial or repeated cases of misconduct will be forwarded to the Office of Student Conduct & Community Standards for additional review. For more information, refer to studentconduct.wiscweb.wisc.edu/academic-integrity/.

Accommodations for Students with Disabilities

The University of Wisconsin-Madison supports the right of all enrolled students to a full and equal educational opportunity. The Americans with Disabilities Act (ADA), Wisconsin State Statute (36.12), and UW-Madison policy (Faculty Document 1071) require that students with disabilities be reasonably accommodated in instruction and campus life. Reasonable accommodations for students with disabilities is a shared faculty and student responsibility. Students are expected to inform me of their need for instructional accommodations by the end of the third week of the semester, or as soon as possible after a disability has been incurred or recognized. I will work either directly with you or in coordination with the McBurney Center to identify and provide reasonable instructional accommodations. Disability information, including instructional accommodations as part of a student's educational record, is confidential and protected under FERPA.

<http://mcburney.wisc.edu/facstaffother/faculty/syllabus.php>

Diversity and Inclusion

Diversity is a source of strength, creativity, and innovation for UW-Madison. We value the contributions of each person and respect the profound ways their identity, culture, background, experience, status, abilities, and opinion enrich the university community. We commit ourselves to the pursuit of excellence in teaching, research, outreach, and diversity as inextricably linked goals.

<https://diversity.wisc.edu/>

Parenting

If you are a parent, you are invited to bring your baby or child to class if you find yourself in a rare childcare crisis and would miss a lecture or section meeting otherwise. Make sure to bring food, drink, and something to keep your child entertained quietly. Position yourself near the door so you can step out with your child if need be.

Course Outline

Subject to change. Please monitor course announcements.

Readings and assignments are posted on the course website.

MHE: *Mostly Harmless Econometrics*

MM: *Measuring Metrics*

OE: *Observation and Experiment*

All readings are required, unless marked * for optional and ** for technically demanding (i.e. allow more time).

January	21	L1: Introduction & Overview * Lalonde (1986)
	23	L2: Potential Outcomes Causality & Randomized Experiments MHE: Chapter 2 * OE: Chapters 1-2
	28	L3: Matching I OE: Chapter 5 *Gerber and Green (2012), Chapter 4
	30	L4: Matching II Stuart (2010) Read one of the following two: Iacus, King, Porro (2011) (read for ideas, math optional) Blackwell et al. (2009) (Stata manual for Iacus et al.) * Imai, Stuart, King (2008) pp481-493
February	4	L5: Path Models I Pearl 2013 *Elwert 2014: 245-51, 256-8.
	6	L6: Path Models II (Same readings as L5)
	10	<i>Problem Set 1 due in lab section</i>
	11	L7: Technical Tools 1: Properties of Estimators Winship, Lecture Script R1
	13	L8: Technical Tools 2: Relationships Between Distributions Winship Lecture Script R2

- 17 *Problem Set 2 due in lab section*
- 18 L9: Technical Tools 3: Matrix Algebra
Winship Lecture Script L4
- 20 L10: Regression Assumptions & MLE Estimation
Winship Lecture Script L5 (skip p.9-11)
*Long 1997:25-33
- 24 *Problem Set 3 due in lab section*
- 25 L11: Identification and Misspecification
Winship Lecture Script L6
** MHE, CH3 pp. 68-77
- 27 Review
- March 2 *Problem Set 4 due in lab section*
- 3 Midterm I
- 5 L12: Robust Inference and GLS
Winship Lecture Script 12
*Baum 2006, Chapter 6
** Abadie et al. 2017: pages 1-7.
- 10 L13: HLM & Random Effects
Note: Raudenbush & Bryk and Winship overlap in substance but differ greatly in presentation.
Raudenbush and Bryk 2002, Ch. 2
Winship Lecture Script 13
Raudenbush and Bryk 2002, Ch. 3
- 12 L14: Tobit and Probit
Winship Lecture Script 10
* Long, pp34-54
- 17 Spring break (no class)
- 19 Spring break (no class)
- 23 *Problem Set 5 due in lab section*

- 24 L15: Logistic Regression I
Long, pp 61-79, *79-83
**MHE, Section 3.4.2
- 26 L16: Logistic regression II
Read at least one of these three:
Mize 2019
Mood 2010
Breen, Holm, Karlson 2018
- 30 *Problem Set 6 due in lab section*
- 31 L17: Instrumental Variables I
Winship L5 pp.9-11
MM Chapter 3
*MHE Sections 4.1 & 4.2
- April 2 L18: Instrumental Variables II
Same readings as L17
- 6 *Problem Set 7 due in lab section*
- 7 Review
- 9 Midterm II
- 14 L19: Regression Discontinuity Design
Dunning 2012 Ch. 3
MM Chapter 4
- 16 L20: Panel Models: DID and Fixed Effects I
MM Chapter 5
- 21 L21: Panel Models: DID and Fixed Effects II
TBD
- 23 L22: Panel Models: Mixed Models
Same readings as L13
- 27 *Problem Set 8 due in lab section*

28 L23: Mediation Analysis
VanderWeele 2016
* VanderWeele 2015, Chs. 1 & 2

30 Review

May 6 Final Exam: 2:45-4:45 (Place: TBD)

Problem Set 9 due before final exam