

Sociology 362: Statistics for Sociologists III
University of Wisconsin-Madison
Spring 2025

Who and Where?

Lecture Time: Tuesday & Thursday 2:30-3:45PM
Lecture Location: 1263 Computer Sciences
Lab Time: Monday 1:20-3:15PM/3:30-5:25PM
Lab Location: 3218 Sewell Social Sciences

Instructor: Prof. Felix Elwert
Office Hours: 4426 Social Science Building
Tuesday, 11-noon
Thursday, 1-2pm
Email: elwert@wisc.edu

Teaching Assistant: Ang Yu
Office Hours: 2458 Social Science Building
Monday, 10-11am
Thursday, 10-11 am
Email: ayu33@wisc.edu

Course Website <https://canvas.wisc.edu/courses/444896>

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 in-person.

Course Description

This is the third course of the Statistics-for-Sociologists sequence, designed to survey key methods and approaches of modern quantitative research in sociology. We will engage with two main topics.

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

We will move back and forth between identification and estimation throughout the course, since different identification strategies often employ different estimation tools. For example, when researchers are willing to assume that they can identify causal effects by *adjusting* for observed covariates, then they could resort to any number of *regression* or *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 a source of *random variation* in treatment assignment, they may use *regression-discontinuity* or *instrumental variables estimators*. In sum, different identification strategies cohabit with 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 most 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*), censored outcomes (*tobit*), and binary outcomes (*probit* and *logit*).

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.

Lectures

Lectures introduce the main course content based on, and sometimes going beyond, required course readings. Students will *complete all required readings before lecture, without exception*.

Labs

Labs serve three purposes: they review key material from readings and lecture; assist students with homework assignments by discussing related exercises; giving students the opportunity to ask questions. Sections also provide programming help for software exercises.

Learning Outcomes

By the end of this course, students will be able to recognize what statistical techniques (“estimators”) are appropriate to answer sociological questions from data in a variety of circumstances (“identification conditions”).

This course serves students with a variety of goals.

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

Students who will go on 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, or which books to read, next in order to participate at the research frontier.

Note: statistical methodology is a fast-moving field. This course cannot advance you to the research frontier. It will do its best, however, to get you started.

Grading Components (graduate/undergraduate)

Graduate students are graded on all components listed below. Undergraduates are graded on all components listed below, except the final project. Percentages X/Y refer to X=graduates and Y=undergraduates.

Exams (65/65%): Students will complete two non-cumulative 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.

Recognizing that everybody can have a bad day, I 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 must 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%). Midterms missed without a compelling reason receive a score of 0.

You will take all exams in person.

For each exam, you may use *one single-sided sheet* of your own *handwritten* notes (for formulas etc.). Preparing this sheet will also be great exam prep!

Problem sets (20/30%): 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. All late assignments receive a score of zero. We will drop your lowest-score assignment when computing final grades. Assignments will be graded on effort for partial credit

and for correctness to achieve full credit. We ask that you turn in a hard copy of your problem set answers; you may type up or handwrite your problem sets. Use of LaTeX is encouraged but not required. Your TA will collect your problem sets at the beginning of your lab on Mondays. You may also submit your assignments electronically by emailing them to your TA.

Attendance (3/3%): 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 (lecture and section combined), unless students have had a compelling reason to miss all three (or more) sessions.

Feedback (2/2%): We want your feedback! After every lecture (i.e., same day), 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 “Home>Feedback.” Students who submit feedback for at least 15 will receive full credit (2%), and 0% otherwise.

Folks, attendance and feedback together combine to a half letter grade of essentially “free” points.

Extra Credit (3/3%): 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 (= three glossaries in total) will receive 3% extra credit. Please submit your glossaries on the course website, under “Assignments.”

Final Project (10/0%)

The final project (for graduate students only!) consists of an additional, and more demanding, problem set that emphasizes the synthesis and application of new skills learned throughout the semester to evaluating distinct methodological opportunities and challenges in recent sociological research. This assignment will be distributed at the beginning of April.

Grades

Grades are assigned based on the weighted average of students’ scores on all grade components.

A:	≥92%
AB:	≥86%
B:	≥80%
BC	≥75%
C:	≥65%
D:	≥60%
F:	<60%

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: to learn statistics, you must read the material *repeatedly*. There's no other way.

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 like:

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. The internet is also full of good tutorials—but beware that the internet is, well, the internet; errors abound.

Do not be deceived 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 more hours on 10 pages of technical reading (yes, also for me!).

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 <https://conduct.students.wisc.edu/academic-misconduct/>.

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

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/>

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 ** is for required-and-technically demanding (allow more time).

January	21	L1: Introduction & Overview * Lalonde (1986) * Lundberg, Johnson, Stewart (2021) (instant classic, revisit often) * Petersen and van der Laan (2014) (easier than LJS)
	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 at the beginning of lab</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 (Read sections 1-4)
	17	<i>Problem Set 2 due at the beginning of lab</i>
	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); Read Section 5 in Script R2 *Long 1997:25-33
	24	<i>Problem Set 3 due at the beginning of lab</i>
	25	L11: Identification and Misspecification Winship Lecture Script L6 Cinelli, Forney, Pearl (2022) ** MHE, CH3 pp. 68-77
	27	Review
March	3	<i>Problem Set 4 due at the beginning of lab</i>
	4	Midterm I Regular lecture time
	6	L12: Robust Inference and GLS Winship Lecture Script 12 *Baum (2006), Chapter 6 ** Abadie et al. (2017): pages 1-7.
	11	L13: HLM & Random Effects Note: Raudenbush & Bryk and Winship overlap in substance but differ in presentation. Raudenbush and Bryk (2002), Ch. 2 Winship Lecture Script 13 Raudenbush and Bryk (2002), Ch. 3
	13	L14: Tobit and Probit Winship Lecture Script 10 * Long, pp34-54
	17	<i>Problem Set 5 due at the beginning of lab</i>

	18	L15: Logistic Regression I Long, pp 61-79, * 79-83 **MHE, Section 3.4.2
	20	L16: Logistic regression II Mize (2019)
	22-30	Spring Break
	31	<i>Problem Set 6 due at the beginning of lab</i>
April	1	L17: Instrumental Variables I Winship L5 pp.9-11 MM Chapter 3 *MHE Sections 4.1 & 4.2
	3	L18: Instrumental Variables II Same readings as L17
	7	<i>Problem Set 7 due at the beginning of lab</i>
	8	Review
	10	Midterm II Regular lecture time
	15	L19: Regression Discontinuity Design Dunning (2012) Ch. 3 MM Chapter 4
	17	L20: Panel Models: DID and Fixed Effects I MM Chapter 5
	22	L21: Panel Models: DID and Fixed Effects II TBD
	24	L22: Panel Models: Mixed Models Same readings as L13
	28	<i>Problem Set 8 due at the beginning of lab</i>

	29	L23: Mediation Analysis VanderWeele (2016) * VanderWeele (2015), Chs. 1 & 2
May	1	Review
	3	<i>Problem Set 9 due (all students)</i> <i>Final Project due (graduate students only)</i>
	6	Final exam, 12:25 PM - 2:25 PM