ARE 212 Syllabus
Multiple Equation Estimation

Department of Agricultural and Resource Economics
University of California, Berkeley
Spring 2025

Lectures: Mon/Wed 10:00-12:00, 241 Giannini
Section (as needed): Fri 10:00-11:00, 141 Giannini

Course Website: https://bcourses.berkeley.edu/

Instructor: Michael Anderson
Contact: mlanderson@berkeley.edu, https://are.berkeley.edu/~mlanderson
Office Hours: Tu 1:30-2:30 pm, Fr 9:30-10:30 am
Office hours can be attending in person (Fridays only) or remote (Zoom).
Sign up online at https://are.berkeley.edu/~mlanderson
Zoom link https://berkeley.zoom.us/j/94250028292?pwd=L60l8xhFaAsjhN5wlW2ZOMNOPBxJbS.1

GSI: N/A

Course Description

This is the second course in the graduate econometrics sequence, following ARE 210 (probability and statistics). The primary goal of this course is to provide you with an in-depth understanding of multiple linear regression and how it behaves under different sets of assumptions. In addition it covers causality, experimental design, nonparametric methods, and selection-on-observables designs. Lectures will focus on technical material, supplemented with illustrations of applications. Assignments will deal with the application of techniques covered in class to (real-world) data sets as well as some proofs and derivations. Two in-person exams will test your understanding of the material.

Prerequisites

You are expected to have a working knowledge of matrix algebra and familiarity with the basic fundamentals of probability and statistics. Courses that satisfy the prerequisite are ARE 210, two upper division or graduate mathematical probability and statistics courses (e.g. STAT 134/135 or STAT 201A/201B), or equivalent.

Assignments and Grading

We will assign 4 to 6 problem sets during the course of the semester. You must work cooperatively on the problem sets in groups of 2 to 5. They will be graded on a "completed/not completed" basis. There will also be midterm and final examinations. Grades will be based on

completion of problem sets (10%), midterm (30%), final exam (55%), and class participation (5%). Late problem sets will not be accepted, but you are allowed to drop one problem set.

Information regarding the schedule and location of the final exam will be available at https://registrar.berkeley.edu/scheduling/academic-scheduling/final-exam-guide-schedules. **Please do not ask me when or where the final is.** I assume **no responsibility** for erroneous information if you ask me when/where the final is, as any information I give you on this matter can only be weakly *less* accurate than what is on the Registrar website.

Statistical Software

You may use any software that you wish, but solutions for problem sets will be handed out in Stata and R. Demonstrations during lectures will be conducted in Stata. In the long run, if you are doing applied microeconometrics, you will almost surely end up using one of these two packages. However, in some problem sets it will be recommended that you use Stata's or R's more primitive commands, or the Mata language, rather than the "canned" commands.

Textbooks and Notes

The course is not based on any one text, but draws heavily on Bruce Hansen's *Econometrics* and Angrist and Pischke's *Mostly Harmless Econometrics*. Guido Imbens' and Jeff Wooldridge's formal econometrics notes (from a NBER econometrics course) also form a useful reference.

- [*BH*] Hansen, Bruce (2021). *Econometrics*. Princeton University Press.

- [*AP*] Angrist, Joshua and Jorn-Steffen Pischke (2009). *Mostly Harmless Econometrics*. Princeton University Press.

- [*WNE*] Imbens, Guido and Jeffrey Wooldridge (2007). *What's New In Econometrics*, NBER Summer Course.

Classroom Climate

We are all responsible for creating a learning environment that is welcoming, inclusive, equitable, and respectful. If you feel that these expectations are not being met, please consult your instructors, or seek assistance from campus resources (please see Academic Accommodations).

Accommodations

Students with DSP accommodations should have the DSP office inform the instructor within the first three weeks of classes. In general it is logistically infeasible to grant last-minute requests for accommodations just prior to exams or assignment due dates. The purpose of academic accommodations is to ensure that all students have a fair chance at academic success. Disability, or hardships such as basic needs insecurity, uncertain documentation and immigration status, medical and mental health concerns, pregnancy and

parenting, significant familial distress, and experiencing sexual violence or harassment, can affect one's ability to satisfy particular course requirements. Students have the right to reasonable academic accommodations, without having to disclose personal information to instructors, and thus arrangements should be made via DSP. For more information about accommodations, scheduling conflicts related to religious creed or extracurricular activities, please see Academic Accommodations.

Course Outline

1. All about linear regression

A. The conditional expectation function (CEF) and ordinary least squares (OLS)

AP Chapter 3.1.

BH Chapter 2.

Geiser, Saul and Maria Veronica Santelices. "Validity Of High-School Grades In Predicting Student Success Beyond The Freshman Year: High-School Record vs. Standardized Tests as Indicators of Four-Year College Outcomes." *Center for Studies in Higher Education Research and Occasional Paper Series CSHE.6.07*, 2007.

Hertenstein, Matt, Carrie Hansel, Alissa Butts, and Sarah Hile. "Smile Intensity in Photographs Predicts Divorce Later in Life." *Motivation and Emotion*, 2009, 33, 99-105.

B. The classical linear regression model (CLRM) — assumptions and algebra

BH Chapters 3 and 4.1-4.9.

C. CLRM — goodness of fit and decompositions

BH Chapters 3.13-3.16 and 4.20-4.21.

Rahimi, Ebrahim, and Seyed Saeed Hashemi Nazari. "A detailed explanation and graphical representation of the Blinder-Oaxaca decomposition method with its application in health inequalities." *Emerging Themes in Epidemiology*, 2021, 18:12.

Fortin, Nicole, Thomas Lemieux, and Sergio Firpo. "Decomposition Methods in Economics." *Handbook of Labor Economics*, 2011, Vol. 4, Part A, 1-102.

D. CLRM — finite sample properties

BH Chapter 5.

E. CLRM — large sample properties

BH Chapter 7.

AP Chapter 3.1.3.

F. CLRM — hypothesis testing

BH Chapter 9.

G. CLRM — problems with the covariance matrix

BH Chapter 4.10-4.25.

AP Chapter 8.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics*, 2004, 119, 249-275.

Cameron, Colin, Jonah Gelbach, and Doug Miller. "Robust Inference with Multi-way Clustering." *Journal of Business and Economic Statistics*, 2011, 29, 238-249.

Sjoquist, David, and John Winters. "Building the Stock of College-Educated Labor Revisited." *Journal of Human Resources*, 2012, 47, 270-285.

H.CLRM — grab bag problems (limited dependent variables and measurement error)

2. Other estimators (primarily) of the CEF

A. Binned scatterplots

BH Chapter 19.2.

Cattaneo, Matias, Richard Crump, Max Farrell, and Yingjie Feng. "On Binscatter." *American Economic Review*, 2024, 114, 1488-1514.

B. Kernel densities and nonparametric regression

BH Chapters 19 and 20.

Blundell, Richard and Alan Duncan. "Kernel Regression in Empirical Microeconomics." *The Journal of Human Resources*, 1998, 33, 62-87.

Cleveland, William. "Robust Locally Weighted Regression and Smoothing Scatterplots." *Journal of the American Statistical Association*, 1979, 74, 829-836.

3. Prediction

A. Machine learning fundamentals

BH Chapter 29.

Mullainathan, Sendhil and Jann Spiess. "Machine Learning: An Applied Econometric Approach." *Journal of Economic Perspectives*, 2017, 31, 87-106.

B. Time series

BH Chapter 14.

4. Causality

A. Potential outcomes

AP Chapters 1 and 2.

BH Chapter 2.30.

WNE Lecture 1, Section 2.

Holland, Paul "Statistics and Causal Inference." *Journal of the American Statistical Association*, 1986, 81, 945-960.

Rubin, Donald "Statistics and Causal Inference: Comment: Which Ifs Have Causal Answers?" *Journal of the American Statistical Association*, 1986, 81, 961-962.

Willer, Robb, Christabel Rogalin, Bridget Conlon, and Michael Wojnowicz. "Overdoing Gender: A Test of the Masculine Overcompensation Thesis." *American Journal of Sociology*, 2013, 118(4), 980-1022.

B. Randomized controlled trials

AP Chapter 2.

C. Cautionary notes

Lalonde, Robert. "Evaluating Econometric Evaluations of Training Programs with Experimental Data." *American Economic Review*, 1986, 76, 604-620.

Freedman, David. "Statistical Models and Shoe Leather." *Sociological Methodology*, 1991, 21, 291-313.

Scheiber, Noam. "Freaks and Geeks: How Freakonomics is Ruining the Dismal Science." *The New Republic*, 2007, April 2, 27-31.

D. Selection on observables— regression adjustment

AP Chapter 3.2.

WNE Lecture 1, Section 3.1.

Yule, G. Udny. "An Investigation into the Causes of Changes in Pauperism in England, Chiefly During the Last Two Intercensal Decades (Part I.)." *Journal of the Royal Statistical Society*, 1899, 62, 249-295. (possibly the first published paper with multiple regression estimates)

Krueger, Alan. "How Computers Have Changed the Wage Structure: Evidence from Micro Data." *Quarterly Journal of Economics*, 1993, 108, 33-60.

DiNardo, John and Jorn-Steffen Pischke. "The Returns to Computer Use Revisited: Have Pencils Changed the Wage Structure Too?" *Quarterly Journal of Economics*, 1997, 112, 291-303.

Altonji, Joseph, Todd Elder, and Christopher Taber. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy*, 2005, 113, 151-184.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. "High-Dimensional Methods and Inference on Structural and Treatment Effects." *Journal of Economic Perspectives*, 2014, 28, 29-50.

Goldsmith-Pinkham, Paul, Peter Hull, and Michal Kolesár. "Contamination Bias in Linear Regressions." *American Economic Review*, 2024, 114, 4015-4051.

E. Selection on observables —propensity score and dimensionality reduction

AP Chapter 3.3.

WNE Lecture 1, Sections 3.2 - 3.4 and 5 - 7.

Rosenbaum, Paul and Donald Rubin. "Reducing Bias in Observational Studies Using Subclassification on the Propensity Score." *Journal of the American Statistical Association*, 1984, 79, 516-524.

Dehejia, Rajeev and Sadek Wahba. "Causal Effects in Non-Experimental Studies: Reevaluating the Evaluation of Training Programs." *Journal of the American Statistical Association*, 94, 1999, 1053-1062.

Arceneaux, Kevin, Alan Gerber, and Donald Green. "Comparing Experimental and Matching Methods Using a Large-Scale Voter Mobilization Experiment." *Political Analysis*, 2006, 14, 37-62.

Griffen, Andrew and Petra Todd. "Assessing the Performance of Nonexperimental Estimators for Evaluating Head Start." *Journal of Labor Economics*, 2017, 35, S7-S63.

Shadish, William, M. H. Clark, and Peter Steiner. "Can Nonrandomized Experiments Yield Accurate Answers? A Randomized Experiment Comparing Random and Nonrandom Assignments." *Journal of the American Statistical Association*, 2008, 103, 1334-1356.

Dale, Stacy and Alan Krueger. "Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables." *Quarterly Journal of Economics*, 2002, 117, 1491-1527.

Anderson, Michael. "The Benefits of College Athletic Success: An Application of the Propensity Score Design." *Review of Economics and Statistics*, 2017, 99, 119-134.

Imbens, Guido. "Matching methods in practice: Three examples." *Journal of Human Resources*, 2015, 50, 373-419.