Syllabus

Overview: This is a "shared course", which is like a cross-listed course, except that it has different course numbers and course titles.[1] The two different courses being taught are ECON250C (Labor Economics) and LAW209.6 (Topics in Quantitative Methods). *Speaking to the economics students*: there is not much in the way of labor economics theory covered in this course (except possibly for some simple job search theory). Labor theory has been covered more than adequately in ECON250A and ECON250B. My impression is that students have figured out something about theory, but know less about practical issues such as computation. If you are more interested in pure labor theory (e.g., questions about optimal contracts) then this course might be interesting, but isn't exactly what are you looking for. This course is more focused on empirical labor economics. In particular, what I am interested in making sure of in this course is that you know how to actually *do* things with honest-to-goodness data sets. *Speaking to the law/JSP students*: you should have already taken LAW209.3 (QM1), and LAW209.33 (QM2) or a course like it. If you are worried about your level of preparation, come see me and we'll talk about it. *Speaking to students from neither economics nor law*: If you are worried about your level of preparation, come see me.

Office Hours: Th 12-2pm, 657 Evans. Note: My main office is in the law school, 586 Simon Hall, which may be relevant if you are trying to track me down at a random time (I have kids, so "random time" doesn't mean nights or weekends, but rather a random time during the workweek). Also: I tend to be behind on my e-mails. I usually think it is going to get better soon. See the behavioral economists, or develop some empathy, if you are having a hard time understanding this phenomenon. So the chances are that you will send me an e-mail, see me 2 days later and say "I sent you an e-mail" and I will look at you blankly. You have been warned. If you need to find me, contact my assistant Ida Ng (`ing@law.berkeley.edu`). She actually does check her e-mail, answers her phone, and is deeply reliable. She also controls my calendar, knows when I am around, and has my cell phone number in case I've forgotten an appointment.

Goals for the Course: I want you all to write an empirical paper that I'm interested in reading. It doesn't have to be of publishable quality, although that is obviously a plus. There are (at least) 2 hard aspects to writing an interesting empirical paper: (1) coming up with a good question, and (2) answering it well. It is very hard to teach students how to come up with good research questions. It's a bit like trying to explain how to tie your shoes. Most professors "have the knack" (that is why they have jobs). But that doesn't mean they can teach it to others. In contrast, it is relatively easy to teach students how to answer a good research question. (On the other hand, this is a wee bit tautological, because part of what we mean by a good research question is one that is answerable.) The first class session we will discuss research questions, and it will come up again and again throughout the term.

Grading: The final grade for the course will be calculated as follows: 37.2% problem sets, 62.8% final paper.

Problem Sets: There will be a bunch of problem sets, perhaps even one a week. They will mostly (all? I haven't decided) be computational. Most of them will involve downloading a data set used in a particular paper and torturing it in various ways. Some of them will involve optimization of a function (you need to know how to do this). Some of them will involve simulation (you need to know how to do this). You will fail at many of them. That's ok. I will generally be asking you to do things you don't know how to do and that you haven't even been trained to do. The idea is to help you realize what you don't know and then to help you to figure out how to do it.[2] The grading of the problem sets will mainly be on effort and on your ability to learn from problem set to problem set (the early ones will develop skills that will be drawn on in subsequent problem sets). There are a small enough number of students that I should be able to keep track of your intellectual progression and/or lack thereof. Don't spend more than 10 hours on any given problem set unless you need to distract yourself from the fact that it is Saturday night and you are lonely.

Readings: There will be various readings, drawing from both textbooks and from the academic literature. You don't have to buy a textbook. On the other hand, you are presumably planning on becoming an academic. That usually means that if I allude to a textbook, it will be assumed by those interviewing you when you are on the job market in 3 (20?) years that you have read it already and picked it clean of all good ideas. Alternatively, and by continuity, at some point your adviser will begin to assume a similar level of knowledge. If he/she does not detect that level of knowledge, eventually he/she will start avoiding you.

[1] I have no idea why the bureaucrats wanted a shared course instead of a cross-listed course, but let's humor them.

[2] Become comfortable with acknowledging your ignorance. Figure out how to rectify it. That is the essence of becoming an academic: be ruthless in detecting your intellectual weaknesses, and be ruthless in fixing them. Nine times out of ten, when I detect ruthlessness in a student, that person is a scholar in four years and has a good academic job in five.

Topics

1. What is a good research question? (It's like the Supreme Court's definition of pornography.) How do you find one?
   (a) Start with a data set?
   (b) Start with a research area?
   (c) Iterate?
   (d) When you figure out the formula, write it down and give me a call
   (e) I can propose good research questions for you to work on, but somehow that never seems to work. David Card's theory is that students are rarely interested in working on somebody else's idea. If you are the exception, ask.
2. Regression, or the Artist Formerly Known As Ordinary Least Squares (Ruud 2000)
   (a) Prediction
   (b) Estimating Coefficients (not the same as the above)
   (c) Testing hypotheses about coefficients
   (d) Do we believe that we observe a random sample on (Y_i, X_i), or do we believe that $Y_i = X_i'\beta + \varepsilon_i$ with $E[\varepsilon_i|X_i] = 0$?
   (e) To weight or not to weight (Deaton 1997)
   (f) Hausman tests (Hausman 1978, Davidson and MacKinnon 1993)
3. Standard Errors and Hypothesis Tests (Davidson and MacKinnon 1993, Davidson and MacKinnon 2004)
   (a) Independence ("independence and Taylor's theorem are all that separate us from the animals" –JED)
   (b) Various large sample approximations: relationship to differing motivations for regression
   (c) Small sample problems with Huber-Eicker-White standard errors (White 1980, White 1982, MacKinnon and White 1985)
   (d) Flavors of the bootstrap and jackknife (size) (Hansen 2006)
   (e) Empirical likelihood (power)
   (f) Serial correlation
4. Selection on Observables (Imbens 2004)
   (a) Regression controls (Blinder 1973, Oaxaca 1973)
   (b) Reweighting (Horvitz and Thompson 1952, DiNardo, Fortin and Lemieux 1996)
   (c) Matching (Rosenbaum and Rubin 1983)
5. Instrumental Variables (Angrist and Krueger 1999, Angrist, Imbens and Rubin 1996)
   (a) Good instruments are hard to find
   (b) Constant coefficients versus random coefficients
   (c) Pure statistical problems that are usually second-order can be first-order here
   (d) Hausman tests revisited
   (e) Regression discontinuity as special case (Lee 2008)
6. Maximum Likelihood (Ruud 2000, White 1982, Efron 1988)
   (a) Information Matrix Test
   (b) Censoring
   (c) Sample selection (Heckman 1976)
7. Method of Moments (Newey and McFadden 1994)
   (a) Unifying Framework for almost all estimators of interest
   (b) Figuring out standard errors for unusual problems
   (c) Testing fixed effects (Chamberlain 1984, Ashenfelter and Krueger 1994)
   (d) Relationship to minimum chi-square
8. Nonparametrics (Pagan and Ullah 1999, Fan and Gijbels 1996)
   (a) Nice flexible techniques, but... "Real" asymptotics (good) versus "promise" asymptotics (misleading)
   (b) Back to regression discontinuity (McCrary and Royer 2011)
9. Getting good data
   (a) Probably more important than any of the above topics
   (b) Completely different skill set required to get it—need to be able to smooth-talk administrators, need to do endless Google searching, need to start with a good research question, need to also know your public-use data sets
   (c) Often need to merge various data sets (easy to make mistakes here, and need to get crosswalks if different data sets use different identifiers), often need to investigate quality of measurement by comparing measurements in various data sets
   (d) Can be tedious, but incredibly important. Part of the job where you get out of the clouds and into the muck. I won't take you seriously unless you are willing to get in the muck and stay there until you get the job done.
   (e) CCRDC

References

Angrist, Joshua and Alan B. Krueger, "Empirical Strategies in Labor Economics," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics*, Vol. 3, Amsterdam: North Holland, 1999.

Angrist, Joshua D., Guido W. Imbens, and Donald B. Rubin, "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, June 1996, *91* (434), 444–455.

Ashenfelter, Orley and Alan Krueger, "Estimates of the Economic Returns to Schooling from a New Sample of Twins," *American Economic Review*, December 1994, *84* (5), 1157–1173.

Blinder, Alan, "Wage Discrimination: Reduced Form and Structural Estimates," *Journal of Human Resources*, 1973, *8* (4), 436–455.

Chamberlain, Gary, "Panel Data," in Zvi Griliches and Michael D. Intriligator, eds., *The Handbook of Econometrics*, Vol. 2, Amsterdam: North-Holland, 1984, pp. 1247–1318.

Davidson, Russell and James G. MacKinnon, *Estimation and Inference in Econometrics*, New York: Oxford University Press, 1993.

____ and ____ , *Econometric Theory and Methods*, New York: Oxford University Press, 2004.

Deaton, Angus, *The Analysis of Household Surveys : A Microeconomic Approach to Development Policy*, Washington, D.C.: World Bank, 1997.

DiNardo, John E., Nicole M. Fortin, and Thomas Lemieux, "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," *Econometrica*, September 1996, *64* (5), 1001–1044.

Efron, Bradley, "Logistic Regression, Survival Analysis, and the Kaplan-Meier Curve," *Journal of the American Statistical Association*, June 1988, *83* (402), 414–425.

Fan, Jianqing and Irene Gijbels, *Local Polynomial Modelling and Its Applications*, New York: Chapman and Hall, 1996.

Hansen, Bruce E., *Econometrics*, Unpublished manuscript, 2006.

Hausman, Jerry, "Specification Tests in Econometrics," *Econometrica*, November 1978, *46* (6), 1251–1271.

Heckman, James J., "The Common Structure of Statistical Models of Truncation, Sample Selection, and Limited Dependent Variables and a Simple Estimator for Such Models," *Annals of Economic and Social Measurement*, 1976, *5* (4), 475–492.

Horvitz, D. and D. Thompson, "A Generalization of Sampling Without Replacement from a Finite Population," *Journal of the American Statistical Association*, 1952, *47*, 663–685.

Imbens, Guido W., "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review," *Review of Economics and Statistics*, February 2004, *86* (1), 4–29.

Lee, David S., "Randomized Experiments from Non-random Selection in U.S. House Elections," *Journal of Econometrics*, February 2008, *142* (2).

MacKinnon, James G. and Halbert White, "Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties," *Journal of Econometrics*, 1985, *29*, 305–325.

McCrary, Justin and Heather N. Royer, "The Effect of Female Education on Fertility and Infant Health: Evidence from School Entry Policies Using Exact Date of Birth," *American Economic Review*, February 2011.

Newey, Whitney K. and Daniel L. McFadden, "Large Sample Estimation and Hypothesis Testing," in Robert F. Engle and Daniel McFadden, eds., *Handbook of Econometrics*, Vol. 4, New York: North–Holland, 1994, pp. 2111–2245.

Oaxaca, R., "Male-female wage differentials in urban labor markets," *International Economic Review*, 1973, *14*, 693–709.

Pagan, Adrian and Aman Ullah, *Nonparametric Econometrics*, New York: Cambridge University Press, 1999.

Rosenbaum, Paul R. and Donald B. Rubin, "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, April 1983, *70* (1), 41–55.

Ruud, Paul A., *An Introduction to Classical Econometric Theory*, New York: Oxford University Press, 2000.

White, Halbert, "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, 1980, *48* (4), 817–838.

____ , "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, 1982, *50* (1), 1–25.