Plan for DRE7006 - 2023
PhD course in Panel Data / Microeconometrics at BI
preliminary

Christian Brinch

January 2023

Introduction

This course will consist of approximately 12 2 hour lectures on microeconometrics and applied microeconomics. Some lectures will be expanded to three hours, usually by having guest lecturers, and there will be some digital material as well.

Workload: You are supposed to spend about 160 hours on the course. This means that in addition to attending lectures, you should spend maybe 10 hours a week working with the course. This means reading and working with problems.

The course has the dual aims of bringing your basic microeconometric competence up to a certain level - and then to teach some topics that are vital to applied empirical economic research.

We cover roughly Chapters 1-6, 10-11 and 20 in Wooldridge (2010) and and most of Angrist and Pischke (2008). I will also use a lot of examples and motivations from Angrist and Pischke (2014). If you are not familiar with the topics, just pick up and read this (bachelor level) book.

We will spend a lot of time using STATA. We use STATA because this is the industry standard in microeconometrics. Typically, published papers that come with data and programs come in STATA-format. I want to emphasize that I am not particularly competent in or particularly fond of STATA. STATA is a useful tool and it is in many respects less quirky than R.[1]

I will run STATA within a Jupyter notebook. You can set this up for yourself as well. (It is pretty straightforward to install Anaconda and install a Jupyter notebook, and then you will either need to look here: stata-kernel to get STATA up and running within the notebook or use the PYSTATA module shipped with STATA 17, that I am currently using.) It is not necessarily best for you to run STATA through a Jupyter notebook, even though I do that. It is often better to play around with stuff within STATA and then to export what you did to a notebook afterwards. Or just keep it in STATA.

Most STATA-user collect their code in do-files. These do-files are not really programs in any sense, just collections of commands. There is no need for do-files when we use a notebook.

I have put up a STATA course (designed for a Master level course) at itslearning. The course is useful for those without a STATA background. This course runs without the notebooks and with do-files.

[1] None of them are good languages for learning how to program, both of them are useful tools to get things done.

Simplified overview of topics and literature

date	time	topic	literature
Fri Jan 6	10:00-11:45	Intro, OLS	W4 AP3
Fri Jan 13	10:00-11:45	Maximum Likelihood	W12-W13 (simplified)
Fri Jan 20	10:00-11:45	OLS continued	W4 AP3
Fri Jan 27	10:00-11:45	Instrumental variables	W5 AP4
Fri Feb 3	10:00-11:45	Instrumental variables continued	W5 AP4
Tue Feb 6	12:00-14:00	Panel Data Basics, clustered standard errors	W10 AP8
Fri Feb 17	10:00-11:45	Quantile regression	W12.10 AP7
Fri Mar 3	10:00-11:45	Selection on observables	W21.1-W21.3 AP3
Fri Mar 10	9:00-11:45	Difference in differences (and variations)	AP5
Fri Mar 17	9:00-11:45	Regression Discontinuity and Regression kink designs	W21.1-W21.3 AP6
Fri Mar 24	9:00-11:45	Structural Estimation, Bunching and Suff. stat.	
Fri Mar 31	9:00-11:45	Structural Estimation and Marginal treatment effects	

Lecture 1: Intro, OLS basics, bootstrap.

Topics

- Empirical research in economics and the credibility revolution.
- Course outline, which relates to the above point.
- OLS refresher (reasonably advanced, including "robust" standard errors)
- Non-parametric bootstrap for standard errors in regression.
- Introduction to simulation in STATA.

In this lecture, I will present the course outline and discuss some developments in microeconometrics and applied microeconomics over the past 40 years. The discussion will be based on the following papers: Leamer (1983); LaLonde (1986); Angrist and Pischke (2010); Panhans and Singleton (2017) I hope you will take the time to read these papers before we start. This discussion takes us to the current state of applied micro and motivates why we need to know the techniques covered in this course.

We then go on to cover the basic econometrics of the linear regression model and ordinary least squares at the PhD level. We will go through both the asymptotic theory of OLS on the level of Wooldridge (2010), Chapter 4 and how we estimate the same objects as we do using the asymptotic theory using bootstrap techniques.

Preparations before lecture (or after)

Read the mentioned papers. Work through chapters 1-3 in Wooldridge. Pick up and read the book Mastering metrics. Look at the following Ted talk by Duflo, to get a flavor of the use of randomized controlled trials in economics.

The econometrics that we go through covers approximately Chapter 3.1 in Angrist and Pischke (2008) and Chapter 4.1-4.2.3 in Wooldridge (2010). In addition, we will go through material that corresponds to 12.8.

Problems to work with after lecture

A Wooldridge problems 4.2-4.4

B Download a small dataset, for example to Excel. Do OLS (one X is fine) and standard errors manually, both the robust and the not-so-robust version of standard errors.

C Generate data from a linear regression model, do parametric bootstrap, nonparametric bootstrap and standard OLS to assess standard errors. In addition to using the bootstraps to compute standard errors, also use them to assess approximate normality is estimates (graphically).

D Generate data for a treatment effect, with random treatment and heterogeneous treatment effects. Show that you estimate the average treatment effect through linear regression.

Lecture 2: Maximum likelihood estimation.

Topics

- The likelihood principle.
- Consistency of maximum likelihood.
- Sampling variation of maximum likelihood estimates.
- Asymptotic normality of maximum likelihood estimates.
- Cramer-Rao bounds to unbiased estimators.
- The trinity of asymptotically equivalent tests.

This course goes through the analysis of (large N) cross-sectional and large-N panel data. It is important to be aware of the beautiful theory of maximum likelihood estimation and the relationship of what we are doing to that theory.

Lecture 3: OLS continued.

Topics

-
- The OVB (omitted variable bias) formula.
- The Frisch-Waugh Theorem (or regression anatomy).
- What happens if we false impose linearity, result 1: "best" linear approximation.
- What happens if we false impose linearity, result 2: weights.
- Kernel density estimation
- Scatterplot smoothing and Local linear regression

In this lecture, we go through the rest of the main topics from OLS and the linear regression model. This includes the omitted variable bias formula, the Frisch/Waugh-theorem or regression anatomy as it is called in Mostly Harmless econometrics. Finally, we go through two results covered in Mostly Harmless Econometrics that covers what happens if we estimate a linear regression model when the data generating process in not linear.

Preparations before lecture

Work through chapter 4 in Wooldridge (2010) and Chapters 3.1-3.2 and 3.5 in Angrist and Pischke (2008).

Problem set to work with after lecture

A Wooldridge problems 4.8, 4.9. 4.11-4.14

B We estimate a regression and we are interested in the effect of x on y. We also have a control variable z that is just a binary control variable. In reality, the effect of x on y depends on z. What happens if we estimate the regression without an interaction term?

C Use some data and a regression program to verify the exact relationships in the Frisch-Waugh theorem and the OVB formula.

D Generate data from a non-linear regression model. Estimate a linear regression from the data. Validate the two formulas that characterize the estimand

Lecture 4: Instrumental variables

- The beauty of instrumental variables.
- The asymptotics of IV and TSLS.
- IV and treatment effect heterogeneity.
- Imbens and Rubin (1997)
- The Heckman selection model.

In this lecture, we motivate instrumental variables and provide the asymptotics of two stage least squares. We cover 5.1-5.2.2 of Wooldridge (2010) and 4.1-4.2 of Angrist and Pischke (2008).

In addition, we work with the interpretation of IV under treatment effect heterogeneity. We will also go through Imbens and Rubin (1997) and discuss control function methods.

Preparations before lecture

Read the relevant chapter in Mastering metrics.

Problem set to work with after lecture

A Wooldridge problems 5.3 and 5.4. These problems involve estimation.

B Wooldridge problems 5.11. This is theory.

C Consider the simple TSLS problem with one instrument and one endogenous covariate. (Assume that both variables have mean zero if that is useful). Try to make the complete argument why the estimator is asymptotically normal.

D In class I used a dataset from Angrist and Evans (1998) measuring the effect of family size on the labor supply of mothers. I put out the dataset on its learning. It is In this dataset we have two family size instruments. Try to do a specification test of the model exploiting the fact that the model is overidentified.

Lecture 5: Instrumental variables continued

- More on treatment effect heterogeneity.
- IV weights.
- The forbidden regression.
- Bad controls.
- Controls in both stages of TSLS.
- The weak instruments problem.
- Solution number 1 to weak instruments.
- Solution number 2 to weak instruments.

More IV. Heterogeneous effects. Abadie's method. The weak instruments problem. 5.2.6 of Wooldridge (2010) and 4.5-4.7 of Angrist and Pischke (2008).

In this lecture, we covered how IV estimation weights different parts of the slope in the equation of interest depending on where the instrument shifts the distribution of the endogeneous covariate. This is material from Løken et al. (2012) and an unpublished (?) paper by Lochner and Moretti.

We also covered issues related to the robustness of IV, why we should include control variables in both stages in our IV routines and what is called the forbidden regression. All of this material is covered in the IV chapter in Angrist and Pischke (2008).

Preparations before lecture

Read Angrist and Keueger (1991).

Read Chernozhukov and Hansen (2008) and browse Andrews et al. (2019).

Problem set to work with after lecture

In this problem, we will study the returns to education using instrumental variable methods based on a dataset from Wooldridge (2010) The dataset is called CARD.DTA. It is uploaded on its learning.

A Do a simple IV analysis with log wage (lwage) as dependent variable, nearc2 and nearc4 (two college proximity dummies) as instruments and exper expersq black south smsa reg661-reg668 smsa66 as control variables. Make sure standard errors do not rely on homoskedasticity. Inspect the first stage and test for weak instruments using the F-statistic ¿ 10 rule of thumb. Interpret the results.

B Do the same analysis based on "manually" doing two stage least squares and compare the standard errors you get in the two cases. Embed the manual two stage least squares procedure in a program and bootstrap the standard errors. Compare with the other standard errors.

C Do a two stage least squares analysis using only nearc2 as an instrument. Inspect the first stage results and the rule of thumb. Make a "weak instruments"-robust test of whether there is an effect of education on earnings based on the reduced form. Finally, make a 95 percent confidence interval on the effect of education on earnings based on the reduced form and the tests suggested in the paper by Chernozhukov and Hansen. (You will have to make tests for various values of beta and make a 95 percent confidence interval using the set of values that can be rejected at the 5 percent level - this is called inverting a test). Compare your results with the two stage least squares confidence interval.

D When using two stage least squares, the predicted endogenous covariate in the second stage has a certain variance. Part of the variation of the predicted endogenous covariate is just noise generated by the estimation error in the first stage. Try to find some way to assess what part of the variance that can be attributed to noise. Apply this to the setting using both nearc2 and nearc4 as instruments. (I will post some hint later.)

E Generate 10 standard normal variables and include these variables as well as nearc2 and nearc4 as instruments in the two stage least squares analysis. Do the same with 100 standard normal variables. Apply the method used in part D to the different analyses. In addition, assess the claim that the IV estimate will be more biased towards the OLS estimate as we include more "rubbish instruments".

F Do an instrumental variable analysis without control variables using only nearc4 as an instrument. Discretize the education variable as well. Use the methods from Imbens and Rubin (1997) to estimate potential outcomes as treated and untreated for the compliers, never-takers and always-takers (only potential outcome as treated for always and untreated for nevertakers). Is the instrument likely to be invalidated by the discretization of the endogenous covariate?

G Generate a model with endogenous treatment (so that OLS fails), and treatment generated by a logit or probit model. Estimate the model using the forbidden regression. Also estimate the model using standard 2sls and the combination between the forbidden regression and 2sls (first predicting treatment using a probit/logit, then using predicted treatment as an instrument in 2sls).

H Do the same as above, but now with a model where treatment is generated by a different model than the model used for estimation (include an interaction term or generate treatment by a linear probability model).

I Use the Card-data (college proximity instrument). Estimate the weight that this instrument (preferably with controls) puts on different margins of educational length. Estimate an OLS model of the effects of educational length on log earnings that allow for changes in the slope. Compare the results of an appropriately IV-weights-reweighted OLS estimate with the IV estimate. Should these be equal if education is exogeneous?

Lecture 7: Panel data

- Measurement error and regression (Wooldridge 4.4)
- Basics of panel data and the omitted variables motivation (Wooldridge 10.1-10.3)
- Clustered standard errors (Wooldridge 7.8.4)
- Random Effects Models and Feasible GLS (Wooldridge 10.4)
- The fixed effects model (Wooldridge 10.5)
- First-differencing (Wooldridge 10.6)

Preparations before lecture

Problem set to work with after lecture

A Set up a small simulation study. Generate a model with measurement errors in the dependent variable. Show that OLS is not affected. Generate a model with classical measurement error in the right hand side variable. Show that we get attenuation bias and that IV solves the problem.

B Find or generate a panel data set and estimate a model using a random effects command (e.g. xtreg, re). Try to figure out if you can estimate the random effects model using FGLS manually somehow. Also estimate the re model the indirect way through estimating the within and between estimators and take a precision-weighted average.

C Again, find or make some panel data set. Estimate the fixed effects model. Show numerically that the slope coefficient in the fixed effects model is approximately equal to the average of the unit specific slope coefficients (

Lecture 7: Quantile regression

In this lecture we go through material from the quantile regression chapter in Angrist and Pischke (2008). The main topics are

- quantile regression
- interpretation of treatment effects in multivariate quantile regression
- unconditional quantile treatment effects (Firpo et al., 2009).
- estimation of quantile treatment effects with an instrument, from Angrist and Pischke (2008) - using the Abadie kappa.
- estimation of quantile treatment effects with an instrument, using Chernozhukov and Hansen (2005)

Problems to work with after lecture

We will exploit the following approximation

$$E(y) \approx \sum_{j=1}^{m-1} d_j Pr(c_j < y < c_{j+1})$$

where $c_1, \ldots, c_n$ is an increasing sequence, d_j are numbers such that $c_j < d_j < c_{j+1}$ and y is bounded between c_1 and c_n. (This holds exact in the limit av n grows, provided a suitable specification for how the $c's$ are chosen when n grows.)

One variation of this approximation is

$$E(y) \approx \frac{1}{m} \sum_{i=1}^{m} F_y^{-1}(\frac{i}{m+1}),$$

where F_y^{-1} is the quantile function of y. (The quantiles play the role of the $d's$ and the probabilities can be thought about as equal.

Hence, when we do a mean regression (a standard regression), the results (slope coefficients) can be decomposed into distributional effects using quantile regression based on the approximation

$$\frac{\partial E(y;x)}{\partial x} \approx \frac{1}{m} \sum_{j=1}^{m} \frac{\partial F_y^{-1}(\frac{j}{m+1})}{\partial x}$$

inducating that the mean regression slope coefficient should be approximately the average of the quantile regression slope coefficients, or we could just use a series of prespecified intervals

$$\frac{\partial E(y;x)}{\partial x} \approx \frac{1}{m} \sum_{i=1}^{m} \frac{c_{j+1} + c_j}{2} \frac{\partial Pr(c_j < y < c_{j+1})}{\partial x}.$$

a. Use the data set "beauty.dta". This dataset contains a variable "look" that is an index of the attractiveness of the person involved in addition to log wages, education etc. We will look into the effect of beauty on wages. Run quantile regressions for each of the nine deciles, finding the effect of "looks" on log wages. Predict what the effect will be in a standard linear regression. Compare the results.

b. Use the deciles of log wages to define 10 groups and find the average log wages within each of these groups. Define ten outcome dummy variables based on which group the log wage falls within. (Or based on whether the log wage is above the 1st decile, the 2nd decile etc.) Use these regression to predict the outcome in the mean regression as in a. (Hint, you can use the STATA commands "xtile lwagecat=lwage, nq(10)" to define a variable that specifies which quantile we are in, use "bysort lwagecat: summarize lwage" to find the quantile specific means and define new outcome variables based on "gen lwagecat10=lwagecat==10" etc.

c. Try to do the same job using unconditional quantile regressions, as in Firpo et al. (2009). That is, estimate regressions based on the dependent variables $q_\tau + \frac{\tau + 1(y \leq q_\tau)}{f(q_\tau)}$, where q_τ are quantiles. You can find deciles and kernel density estimates for use in the denominator by using the command "pctile deciles = lwage, n(10)", which will give a list of the nine deciles and "kdensity lwage, at(deciles) generate(deciles densities)" which will provide a list of values to use for the denominator.

d. Give an assessment of the pros and cons of the different approaches used to decompose the mean regression. Why are the results in a. and c. not identical? Does it matter that we use log wage and not wage as dependent variable? What happens when we want to control for more covariates, like education and experience?

Lecture 8: Selection on observables

- The unconfoundedness / selection on observables assumption.
- Dale and Krueger (2002)
- inverse propensity score weighting methods
- matching methods
- double robust methods
- Pischke paper on testing.
- Altonji et al. (2005)

Estimating treatment effects under the unconfoundedness assumption (selection on observables). 21.1-21.3 of Wooldridge (2010) and 3.3 of Angrist and Pischke (2008).

Preparations before lecture

Read the chapter 2 in Mastering metrics or Dale and Krueger (2002)

Read Altonji et al. (2005).

Problem set to work with after lecture

This problem involves more sorting out different methods of estimation using matching methods and inverse propensity score weighting software-wise. It is possible to implement the commands in STATA using the "teffects" command. Use the data set jtrain3.dta. This is the same dataset that is used for examples in the Wooldridge textbook. We are interested in the effects of job training (in 1977) on real earnings in 1978. As a list of control variables, use "age educ black hisp married".

A Start out by comparing the means without any controls. Follow up with a standard regression adjust (that is, just using linear regression).

B Estimate a propensity score model (a logit model) using the list of controls as explanatory variables. Investigate the propensity scores in the treated and untreated groups.

C Estimate an average treatment effect based on a. treatment effects estimation with regression adjustment (what is the difference from A?) b. treatment effects estimation with inverse propensity score weighting c. treatment effects estimation controlling for an estimated propensity score d. matching estimator with matching on covariates e. matching estimator with propensity score matching

D In C, we have some trouble because the treatment and control groups are very different. Make some effort to trim the sample so that the two groups are more comparable, Redo the analysis above.

E In some of the treatment effect estimation procedures, it is necessary to correctly model the propensity score, while in others, it is necessary to correctly model the relationship between outcomes and control variables. It is also possible to combine the two approaches so that the estimator works (estimates the average treatment effect consistently) if only one of the two are correctly specified. Set up a small simulation study where you generate two data sets - one where the outcome is a simple known function of controls, and one where the propensity score is a simple known function of control variable (make sure the other assumption fails). Illustrate how the different methods work on these three datasets.

Lecture 9: Difference-in-differences

Chapter 5 in Angrist and Pischke (2008) and Section 6.5 in Wooldridge (2010). In addition, we will go through Abadie et al. (2010) and parts of Callaway and Sant'Anna (2020). We will also look into the application McCrary (2007) for some event study analyses.

Problem set to work with after lecture

Check out Favara and Imbs (2015). This paper can be interpreted as a diff-in-diff / IV using easing of restrictions on bank operations (by out-of-state banks) as exogenous variation. It is a very interesting paper.

Download the data from AER. (It is not entirely clear to me that I am allowed to download and distribute it.) Try to analyze "the first stage" using a diff-in-diff. Note that the models in the paper use lagged dependent variables here. You will not be allowed to have them in the models as they violate the strict exogeneity assumption. (We will get back to the lagged dependent variables in the next lecture). Estimate "the reduced form" in a symmetric manner so that we can generate an ILS estimate.

Note that there are several ways to do this. In particular, we need to choose one measure from the list of credit supply measures that are available and we need to consider how to use the weights that appear in the reduced form part of the paper. And there is a question of levels vs differences.

Please also play around a bit with leads and lags so we can get event-study estimates from the diff-in-diff-models.

Lecture 10: Regression discontinuity and kinks

The lecture will be based on the Chapter on regression discontinuity designs in Angrist and Pischke (2008) and the exposition in Lee and Lemieux (2010). We will also go through the test in McCrary (2008).

Preparations before lecture

For those who are not acquainted with regression discontinuity designs at all, read the relevant chapter in Mastering 'metrics. Browse the chapters in Angrist and Pischke (2008) and Lee and Lemieux (2010).

Problem set to work with after lecture

Read/browse the paper by Buser (2015) Download the datafiles and run the main do-script. You will now have available data for a regression discontinuity analysis. The running variable is score2_1. Alternatively, download the dataset buser_processed.dta from itslearning. The main outcome variables are "religiousness", "evangelical" and "attendance".

A Perform parametric RDDs (linear and quadratic slopes, different slopes on either side of the threshold) on the different outcomes.

B Experiment with the bandwidth in the sense of excluding data far from the threshold.

C Assess the assumption that the running variable is continuous around the threshold by inspecting histograms.

D Do RDDs with local linear regressions. How do you choose your bandwidth?

E Run McCrary's test.

F Create a small simulation study: Generate data where the outcome is not linear (but quadratic) in the running variable. Show that (global) linear RDDs fail and that global quadratic and local linear RDDs do not fail.

Lecture 11: Structural estimation, bunching and sufficient statistics

Lecture 12: Structural estimation and Marginal treatment effects.

Literature

We are going through Heckman (2010). We draw on quite a bit of material from Brinch et al. (2017). we talked quite a bit about the model from Björklund and Moffitt (1987).

Problems to work with after lecture

a. Use the data from the paper by Angrist and Evans on its learning. This is an analysis of the effects of having more children on the labor supply of mothers. In short, the data consists of women with at least two children. There are two instruments in the data, twin births (at second parity), $z2$ and the same-sex instrument $z1$ (the two first kids have the same sex, this increases the probability of having a third child.) The endogenous treatment variable is x, having at least three children. In this analysis, we use the outcome variable $y1$, the number of weeks the mother worked during a certain year.

b. Estimate the local average treatment effects based on the two different instruments separately, using the Wald estimator.

c. Estimate the generalized Roy model of Bjorklund and Moffit, using the same-sex instrument. Derive the marginal treatment effects (average treatment effects conditional on $U_D = p$, where U_D is the random term in the choice equation.

d. The LATE in the generalized Roy model is the integral over the marginal treatment effect for the values of U_D that defines the group of compliers. Show (numerically) that this LATE is identical to the Wald estimates of the LATE.

e. Test the assumption of constant marginal treatment effects in the Bjorklund-Moffit model.

f. Why can't you do the above analysis with the twins instrument?

g. Assume instead a linear mte model. Estimate the marginal treatment effects with local instrumental variables (you will need both instruments) and the separate estimation approach.

h. Try to use both instruments and assess the linear mte model (can we reject the linear model vs a quadratic alternative).

Hint: (When using both instrument, the first stage is really just to estimate the propensity to have three kids conditional on the two instruments. There will be only three cases: twins (prop score 1), no twins, no same-sex, and no twins, but same-sex. The propensities are of course estimated by the sample proportions.

References

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller (2010) "Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program," *Journal of the American statistical Association*, Vol. 105, No. 490, pp. 493–505.

Altonji, Joseph G, Todd E Elder, and Christopher R Taber (2005) "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools," *Journal of political economy*, Vol. 113, No. 1, pp. 151–184.

Andrews, Isaiah, James H Stock, and Liyang Sun (2019) "Weak instruments in instrumental variables regression: Theory and practice," *Annual Review of Economics*, Vol. 11, pp. 727–753.

Angrist, Joshua and William Evans (1998) "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size," *American Economic Review*, Vol. 88, No. 3, pp. 450–77.

Angrist, Joshua D and Alan B Keueger (1991) "Does compulsory school attendance affect schooling and earnings?" *The Quarterly Journal of Economics*, Vol. 106, No. 4, pp. 979–1014.

Angrist, Joshua D and Jörn-Steffen Pischke (2008) *Mostly harmless econometrics: An empiricist's companion*: Princeton university press.

——— (2010) "The credibility revolution in empirical economics: How better research design is taking the con out of econometrics," *Journal of economic perspectives*, Vol. 24, No. 2, pp. 3–30.

——— (2014) *Mastering'metrics: The path from cause to effect*: Princeton University Press.

Björklund, Anders and Robert Moffitt (1987) "The estimation of wage gains and welfare gains in self-selection models," *The Review of Economics and Statistics*, pp. 42–49.

Brinch, Christian N, Magne Mogstad, and Matthew Wiswall (2017) "Beyond LATE with a discrete instrument," *Journal of Political Economy*, Vol. 125, No. 4, pp. 985–1039.

Buser, Thomas (2015) "The effect of income on religiousness," *American Economic Journal: Applied Economics*, Vol. 7, No. 3, pp. 178–95.

Callaway, Brantly and Pedro HC Sant'Anna (2020) "Difference-in-differences with multiple time periods," *Journal of Econometrics*.

Chernozhukov, Victor and Christian Hansen (2005) "An IV model of quantile treatment effects," *Econometrica*, Vol. 73, No. 1, pp. 245–261.

——— (2008) "The reduced form: A simple approach to inference with weak instruments," *Economics Letters*, Vol. 100, No. 1, pp. 68–71.

Dale, Stacy Berg and Alan B Krueger (2002) "Estimating the payoff to attending a more selective college: An application of selection on observables and unobservables," *The Quarterly Journal of Economics*, Vol. 117, No. 4, pp. 1491–1527.

Favara, Giovanni and Jean Imbs (2015) "Credit supply and the price of housing," *American Economic Review*, Vol. 105, No. 3, pp. 958–92.

Firpo, Sergio, Nicole M Fortin, and Thomas Lemieux (2009) "Unconditional quantile regressions," *Econometrica*, Vol. 77, No. 3, pp. 953–973.

Heckman, James J (2010) "Building bridges between structural and program evaluation approaches to evaluating policy," *Journal of Economic literature*, Vol. 48, No. 2, pp. 356–98.

Imbens, Guido W and Donald B Rubin (1997) "Estimating outcome distributions for compliers in instrumental variables models," *The Review of Economic Studies*, Vol. 64, No. 4, pp. 555–574.

LaLonde, Robert J (1986) “Evaluating the econometric evaluations of training programs with experimental data,” *The American economic review*, pp. 604–620.

Leamer, Edward E (1983) “Let’s take the con out of econometrics,” *The American Economic Review*, Vol. 73, No. 1, pp. 31–43.

Lee, David S and Thomas Lemieux (2010) “Regression discontinuity designs in economics,” *Journal of economic literature*, Vol. 48, No. 2, pp. 281–355.

Løken, Katrine V, Magne Mogstad, and Matthew Wiswall (2012) “What linear estimators miss: The effects of family income on child outcomes,” *American Economic Journal: Applied Economics*, Vol. 4, No. 2, pp. 1–35.

McCrary, Justin (2007) “The effect of court-ordered hiring quotas on the composition and quality of police,” *American Economic Review*, Vol. 97, No. 1, pp. 318–353.

——— (2008) “Manipulation of the running variable in the regression discontinuity design: A density test,” *Journal of econometrics*, Vol. 142, No. 2, pp. 698–714.

Panhans, Matthew T and John D Singleton (2017) “The empirical economist’s toolkit: From models to methods,” *History of Political Economy*, Vol. 49, No. Supplement, pp. 127–157.

Wooldridge, Jeffrey M (2010) *Econometric analysis of cross section and panel data*: MIT press.