PAM 6091: Empirical Strategies for Policy Research II
Spring 2017

Zhuan Pei
Office location: MVR 134
E-mail: zhuan.pei@cornell.edu

Office Hours:
Tuesday
2:50-4:50pm

Lecture: Tuesday, Thursday 1:25-2:40pm, MRV Hall G73

Lecture: Luoyi (Roy) Su; E-mail: ls495@cornell.edu. Office Hours: 3-5pm (435 Uris Hall).

Course description: This course is the second of a two-course sequence. Both PAM 6090 and this course are, for the most part, targeted at students looking to do empirical research into the effects of some X on some Y. Both courses require students to complete problem sets that involve "hands-on" exercises – some based on real data and some using Monte Carlo simulations. The hope is that this "learning by doing" will reinforce what is taught in class. Usually, the first course covers core methods, specifically regression adjustment, matching and instrumental variables. This second course covers regression discontinuity designs, panel data methods, standard error issues and other miscellaneous topics.

Learning goals: Upon successful completion of the course, you will

- Have familiarity with the core methods in empirical research;
- Have proficiency in using Stata;
- Be able to critically evaluate empirical analyses and carry out data-oriented research.

Prerequisites: Ideally, students will have met the requirements for the first course (familiarity with matrix algebra, basic statistics/econoometrics and inference procedures) and will have taken the first course. I will not enforce these requirements, but probabilistically you will get more out of the course if you meet these.

Texts and materials: There is no required textbook for the class, but several books serve as good references for empirical researchers, especially Angrist and Pischke.

1. [AP] Joshua D. Angrist and Jörn-Steffen Pischke, *Mostly Harmless Econometrics*, Princeton University Press, 2009.
2. [W] Jefferey Wooldridge, *Econometric Analysis of Cross-Sectional and Panel Data*, MIT Press, 2010.
3. [CT] A. Colin Cameron and Pravin K. Trivedi, *Microeconometrics: Methods and Applications*, Cambridge University Press, 2005.

Statistical Software: We will mainly use Stata for this class, and the class should have access to CISER.

Evaluation: Students will be evaluated on four problem sets and a group presentation. Problem sets will be posted every two or three weeks during the semester. You are encouraged to work on

the problem sets in groups of up to four students and to hand in a single *typed* answer sheet for the group. Complete problem sets must be submitted to Blackboard by 1:25pm on the due date. Late problem sets will not be accepted under any circumstances. Problem sets will count for 50% of the final course grade.

Students will be randomly divided into teams and assigned a paper to present during the semester. The presenters should aim for a thirty-minute talk, and their performance will be assessed by their peers. I will provide a rubric for the other students to score the presentation, and their scoring needs to be supported by written comments, the quality of which will contribute to their course grades. Within a week after the presentation, the presenting team will also need to submit to me 1) a Stata program that implements the main procedure of the paper if the paper does not have accompanying programs available or 2) documentations of existing Stata program that informs the user what each part of the program does in as much detail as possible. Overall, the presentation component accounts for 50% of the course grade, with the presentation itself accounting for 35%, the Stata porgram and documentations accounting for 10%, and the quality of peer review accounting for 5%.

Special accommodations: In compliance with section 504 of Rehabilitation Act of 1973 and the Americans with Disabilities Act, reasonable accommodation will be provided to students with documented disabilities. Students with disabilities must provide the College with appropriate documentation of their disability before any accommodation can be made. Reasonable accommodation will be provided, on a case-by-case basis.

Academic integrity statement: Absolute integrity is expected of every Cornell student in all academic undertakings. Integrity entails a firm adherence to a set of values, and the values most essential to an academic community are grounded on the concept of honesty with respect to the intellectual efforts of oneself and others. Academic integrity is expected not only in formal coursework situations, but in all University relationships and interactions connected to the educational process, including the use of University resources. A Cornell student's submission of work for academic credit indicates that the work is the student's own. All outside assistance should be acknowledged, and the student's academic position truthfully reported at all times. In addition, Cornell students have a right to expect academic integrity from each of their peers. For further information regarding the Cornell Code of Academic Integrity see: http://cuinfo.cornell.edu/aic.cfm. Unless you have the express permission of the instructor, you should not buy or sell course materials. Such unauthorized behavior constitutes academic dishonesty.

TurnItIn.com Acknowledgement: Students agree that by taking this course that all papers submitted for the course may be subject to submission for textual similarity review to Turnitin.com for the detection of plagiarism. All submitted papers will be included as source documents in the Turnitin.com reference database solely for the purpose of detecting plagiarism of such papers. Use of the Turnitin.com service is subject to the Usage Policy posted on the Turnitin.com site.

Tentative Course Outline

Class	Date	Topic	Student Presentation Paper #	Problem Set Due
1	1/26	Introduction and review		
2	1/31	Introduction and review		
3	2/2	Nonparametric basics		
4	2/7	Nonparametric basics		
5	2/9	Nonparametric basics		
6	2/14	Nonparametric basics		
	2/16	No class		
7	2/23	Regression discontinuity designs		#1
8	2/28	Regression discontinuity designs		
9	3/2	Regression discontinuity designs		
10	3/7	Regression discontinuity designs	#1	
11	3/9	Regression discontinuity designs	#2	
	3/14	Cancelled due to snow storm		
12	3/16	Regression discontinuity designs	#3	
13	3/21	Regression discontinuity designs	#4	
14	3/23	Regression discontinuity designs		
15	3/28	Regression discontinuity designs		#2
16	3/30	Regression discontinuity designs	#5	
17	4/11	Panel methods		
18	4/13	Panel methods		
19	4/18	Panel methods		
20	4/20	Panel methods		
21	4/25	Standard error issues		
22	4/27	Standard error issues		
23	5/2	Standard error issues	#6	
24	5/4	Standard error issues	#7	
25	5/9	Catch up and Summary	#8	#3, #4

Course Readings (tentative; will be updated as the semester progresses) The main readings for each topic are listed below, with the most important references starred.

Introduction and review

Greene, *Econometric Analysis*, 7th edition, Prentice Hall 2012: Appendix B, D and Chapters 2-4, 8, 13.

AP, Chapters 2-4

CT, Chapter 4, 6.

Nonparametric Basics

(*) CT, Chapter 9.

Silverman, B. W., *Density Estimation for Statistics and Data Analysis,* Chapman & Hall, 1986: Chapters 3-4.

Fan, Jianqing and Irene Gijbels, *Local Polynomial Modelling and Its Applications,* Chapman & Hall, 1996: Chapters 2-4.

Cheng, M., Jianqing Fan and J. S. Marron (1996), "On automatic boundary corrections," *The Annals of Statistics*, 25(4): 1691-1708.

Cleveland, William S. (1979), "Robust Locally Weighted Regression and SmoSmooth Scatterplots", *Journal of the American Statistical Association*, 74(368), 829-836.

DiNardo, J., N. Fortin and T. Lemieux (1996), "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach", *Econometrica*, 64(5): 1001-1044.

Fan, Jianqing and Irene Gijbels (1992), "Variable Bandwidth and Local Linear Regression Smoothers," *Annals of Statistics*, 20(4), 2008-2036.

Nadaraya, E. A. (1964), "On Estimating Regression," *Theory Prob. Appl.*, 141-142.

P.M. Robinson (1988), "Root-N-Consistent Semiparametric Regression", *Econometrica*, 56(4): 931-954.

Watson, G S. (1964), "Smooth Regression Analysis," *Sankhya: The Indian Journal of Statistic*s, Series A, 26 (4), 359-372.

Regression Discontinuity Methods

Calonico, Sebastian and Matias D. Cattaneo and Rocio Titiunik (2014), "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs", *Econometrica* 82(6): 2295-2626.

Calonico, Sebastian and Matias D. Cattaneo and Rocio Titiunik (2015), "Optimal Data-driven Regression Discontinuity Plots", *Journal of the American Statistical Association,* 110(512), 1755-1769.

Card, David, David Lee, Zhuan Pei and Andrea Weber (2015), "Inference on Causal Effects in a Generalized Regression Kink Design", *Econometrica*, 86(6): 2453-2483.

Card, David, David Lee, Zhuan Pei and Andrea Weber, Forthcoming, "Regression Kink Design: Theory and Practice", *Advances in Econometrics,* volumne 38 *(Regression Discontinuity Designs:*

Theory and Applications) edited by Matias D. Cattaneo and Juan Carlos Escanciano.

Card, David, Andrew Johnston, Pauline Leung, Alex Mas and Zhuan Pei (2015), "The Effect of Unemployment Benefits on the Duration of Unemployment Insurance Receipt: New Evidence from a Regression Kink Design in Missouri, 2003-2013", NBER Working Paper 20869.

Imbens, Guido W. and Karthik Kalyanaraman (2012), "Optimal Bandwidth Choice for the Regression Discontinuity Estimator", *Review of Economic Studies* 79(3): 933-959.

(*) Lee, David S. and Thomas Lemieux, "Regression Discontinuity Designs in Economics," *Journal of Economic Literature* 48: 281–355.

(*) Lee, David (2008), "Randomized Experiments from Non-random Selection in U.S. House Elections," *Journal of Econometrics* 142(2): 675-697.

Lee, David and David Card (2008), "Regression Discontinuity Inference with Specification Error", *Journal of Econometrics* 142(2): 655-674.

McCrary, Justin (2008), "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test", *Journal of Econometrics* 142(2): 698-714.

Pei, Zhuan and Yi Shen, Forthcoming, "The Devil is in the Tails: Regression Discontinuity Design with Measurement Error in the Assignment Variable", *Advances in Econometrics,* volumne 38 *(Regression Discontinuity Designs: Theory and Applications)* edited by Matias D. Cattaneo and Juan Carlos Escanciano.

Ruppert, David (1997), "Empirical-Bias Bandwidths for Local Polynomial Nonparametric Regression and Density Estimation", *Journal of the American Statistical Association*, 92: 439, 1049-1062

Panel Data

(*) AP Chapter 5

(*) W Chapter 10, 17

(*) Chamberlain, Gary (1984), "Panel data", *Handbook of Econometrics*, 2: 1247-1318.

Angrist, Joshua D. and Withney K. Newey (1991), "Over-identication tests in earnings functinos with fixed effects," *Journal of Business and Economic Statistcs*, 9(3), 317-323.

Behrman, Jere R. and Barbara L. Wolfe (1989), "Does More Schooling Make Women Better Nourished and Healthier? Adult Sibling Random and Fixed Eects Estimates for Nicaragua," *Journal of Human Resources*, 24(4), 644-663.

Bond, Stephen (2002), "Dynamic Panel Data Models: A Guide to Micro Data Methods and Practice", Cemmap Working Paper CWP 09/02.

Card, David and Daniel Sullivan (1988), "Measuring the Effect of Subsidized Training Programs on Movements in and out of Employment", *Econometrica*, 56(3), 497-530.

Card, David, Joerg Heining and Patrick Kline (2013), "Workplace Heterogeneity and the Rise of West German Wage Inequality, Quarterly Journal of Economics, 128(3), 967-1015.

Card, David, Ana Rute Cardoso, Joerg Heining and Patrick Kline (Forthcoming), "Firms and Labor Market Inequality: Evidence and Some Theory", Journal of Labor Economics.

Card, David, Ana Rute Cardoso, and Patrick Kline (2016), "Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women", Quarterly Journal of Economics, 131(2), 633-686.

Chamberlain, Gary (1980), "Analysis of Covariance with Qualitative Data", *Review of Economic Studies*, 47(1), 225-238.

Griliches, Zvi and Jerry A. Hausman (1986), "Errors in variables in panel data," *Journal of Econometrics*, 31, 93-118.

Lancaster, Tony (2000), "The incidental parameter problem since 1948," *Journal of Econometrics*, 2000, 391-413.

Mundlak, Yair (1978), "On the Pooling of Time Series and Cross Section Data," *Econometrica*, 46(1), 69-85.

Neyman, J and Elizabeth L. Scott (1948), "Consistent Estimates Based on Partially Consistent Observations," *Econometrica*, 16(1), 1-32.

Wooldridge, Jefferey M. (2005), "Simple solutions to the initial conditions problem in dynamic, nonlinear panel data models with unobserved heterogeneity," *Journal of Applied Econometrics*, 20(1), 39-54.

Standard Errors

(*) AP Chapter 9

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* 105: 493-505.

Bertrand, Marianne, Esther Duflo and Sendhil Mullainathan (2004), "How much should we trust differences-in-differences estimates?" *Quarterly Journal of Economics* 119(1): 249-275.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2008), "Bootstrap-based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, 90(3): 414–427.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller (2011), "Robust inference with multiway clustering", *Journal of Business & Economic Statistics* 29(2): 238-249.

Cameron, A. Colin and Douglas L. Miller (2013), "A Practitioner's Guide to Cluster-Robust Inference", working paper.

Chesher, A., and I. Jewitt (1987), "The Bias of a Heteroskedasticity-consistent Covariance Matrix Estimator," Econometrica 55(5): 1217–1222.

Conley, Timothy G., and Taber, Christopher R. (2011), "Inference with 'difference in differences' with a small number of policy changes", *Review of Economics and Statistics* 93(1): 113-125.

Moulton, Brent (1986), "Random Group Effects and the Precision of Regression Estimates," *Journal of Econometrics* 32(3): 385-397.

Hansen, Christian (2007), "Asymptotic Properties of a Robust Variance Estimator for Panel Data When T is Large," *Journal of Econometrics*, 141(2): 597–620.

Hansen, Christian (2007), "Generalized Least Squares Inference in Panel and Multilevel Models with Serial Correlation and Fixed Effects," *Journal of Econometrics,* 141(2): 670–694.

Imbens, Guido, and Michel Kolesar (2012), "Robust Standard Errors in Small Samples: Some Practical Advice," *NBER Working Paper no. 18478.*

Young, Alwyn (2015), "Channelling Fisher: Randomization Tests and the Statistical Insignificance of Seemingly Significant Experimental Results," London School of Economics, mimeo.

Young, Alwyn (2015), "Improved, Nearly Exact, Statistical Inference with Robust and Clustered Covariance Matrices using Effective Degrees of Freedom Corrections," London School of Economics, mimeo.

Summary and Overview: Perspectives on Program Evaluation

DiNardo, John and David Lee. 2011. Program Evaluation and Research Designs. In Orley Ashenfelter and David Card, eds., *Handbook of Labour Economics*, Volume 4A. Amsterdam: North-Holland.

Heckman, J. J. (2010). Building Bridges between Structural and Program Evaluation Approaches to Evaluating Policy, *Journal of Economic Literature* 48(2): 356-98.

Imbens, G. W. (2010), Better LATE Than Nothing: Some Comments on Deaton (2009) and Heckman and Urzua (2009). *Journal of Economic Literature* 48(2): 399-423.

Deaton, A. (2010). Instruments, Randomization, and Learning about Development, *Journal of Economic Literature* 48(2): 424-55.

Angrist, J. D. and J-S. Pischke (2010). The Credibility Revolution in Empirical Economics: How Better Research Design Is Taking the Con out of Econometrics, *Journal of Economic Perspectives* 24(2): 3-30.

Leamer, E. E. (2010). Tantalus on the Road to Asymptopia, *Journal of Economic Perspectives* 24(2): 31-46.

Keane, M. P. (2010). A Structural Perspective on the Experimentalist School, *Journal of Economic Perspectives* 24(2): 47-58.

Sims, C. A. (2010). But Economics Is Not an Experimental Science, *Journal of Economic Perspectives* 24(2): 59-68.

Nevo, A. and M. D. Whinston (2010). Taking the Dogma out of Econometrics: Structural Modeling and Credible Inference, *Journal of Economic Perspectives* 24(2): 69-82.