ED 795 – Quasi-Experimental Analysis of Observational Data in Education (Version 2.1)
Thursdays, 4:30-7:15
Poe 209

Instructor: Dr. Stephen Porter, Dept. of Leadership, Policy, and Adult and Higher Education
Email: srporter@ncsu.edu (best way to contact me)
Skype: stephen.porter50010
Phone: 515-428-0008
Office and hours: Poe 616A, by appointment

> *In the eyes of a growing number of knowledgeable observers, the practice of regression analysis and its extensions is a disaster.*
> Berk, *Regression Analysis: A Critique*, p. 203

The main issue we face when studying human behaviour is that it is difficult to randomly assign individuals to control and treatment conditions for many program and policy issues (e.g., being held back a grade in elementary school, or obtaining a college degree). The goal of this course is to supplement your statistical training with advanced research methods used in education and the social sciences to estimate causal effects on observational, versus experimental, data.

By observational, I mean any kind of data where units (students and their families, teachers, schools) <u>choose</u> to engage in behaviors, programs or policies, rather than being randomly assigned by researchers. As should be obvious, these kinds of data are prevalent throughout education, and you will likely analyze such data in your dissertation. Although many educational researchers are loath to admit it, in general OLS and HLM can't be used to analyze observational data in any meaningful way, other than to generate reams of correlational results that no one really cares about.

Although some of the techniques in this course may seem complicated, they are in essence applications of OLS and logistic regression, material which was covered in ED 711. The main foci of this course are 1) understanding conceptually what these techniques are doing, and 2) identifying research problems where they can be used.

Learning objectives:
- Understand the potential outcomes approach to causality
- Understand and implement the following techniques to estimate causal effects with observational data:
    - Matching
    - Difference-in-difference and fixed effects models
    - Instrumental variables
    - Regression discontinuity
- Have a much deeper understanding of issues around internal validity when analyzing observational data

We will be covering these topics as a way of introducing you to these techniques. To use any of these techniques in your dissertation will likely require additional study on your part.

To succeed in this course you must have taken ED 711 or a similar semester-long course on multiple regression, as well as having received a grade of A- or above.

All parts of this syllabus are subject to change as needed. Any changes will be discussed in class.

Evaluation

Your grade will be determined as follows:

1. *Participation* (20%) – Participation in class discussions is vital for this class, and part of your grade will be the frequency and quality of your contributions to our discussions. I will provide a series of discussion questions for each section of the course. You should come to class prepared to answer these questions.

2. *Take-home exams* (60%) – You will be given four assignments during the semester that cover various topics in the course; you will have around one week to complete each assignment. These assignments will generally involve analyzing datasets and reporting results from your analyses. You may use your notes and assigned readings to do the assignments, but you may not consult with any other people or sources (e.g., the Internet). I refer to these short papers as exams to stress that you should complete these assignments completely on your own, without any computer or statistical assistance from anyone else (other than me, of course). Late assignments will not be accepted, because I review the assignment answers during class on the day the assignment is due; you will receive a 0 for that assignment.

3. *Literature review* (20%) – You will write a 15-20 double-spaced page review of the literature on one of the topics below, discussing how researchers have used the techniques in this class to understand the topic. The focus will be on how researchers have used these techniques to take into account selection and the strengths and weaknesses of their approaches. You may investigate another area of research once you have cleared it with me (it should have a decent body of literature using quasi-experimental techniques, at least 5 or 6 studies).

    - Effect of charter schools on K-12 students or schools/districts
    - Effect of grade retention policies on K-12 students
    - Effect of teacher quality on K-12 students
    - Effect of class size on K-12 students
    - Effect of developmental (remedial) education on college students
    - Effect of financial aid on college students
    - Effect of attending college on life outcomes

Please use Google Scholar and EconLit as you search for articles. Educational research using these techniques is interdisciplinary and ERIC is in general a woefully poor resource.

Use APA style in terms of citations and how you order the pages of your paper. I advise learning how to use bibliographic software (which formats your cites and reference list for you), so you do not clutter your brain with useless details such as APA style.

I will use this grading scale for the course:

97	≤	A+	≤	100	73	≤	C	<	77
93	≤	A	<	97	70	≤	C-	<	73
90	≤	A-	<	93	67	≤	D+	<	70
87	≤	B+	<	90	63	≤	D	<	67
83	≤	B	<	87	60	≤	D-	<	63
80	≤	B-	<	83	0	≤	F	<	60
77	≤	C+	<	80					

Academic misconduct

Academic misconduct in any form is in violation of North Carolina State University policy and will not be tolerated. This includes, but is not limited to: copying or sharing answers on assignments, plagiarism, and having someone else do or assist you with your academic work. See http://policies.ncsu.edu/policy/pol-11-35-01 for more information.

Submitting an assignment to me electronically is the equivalent to signing the university Honor Pledge: "I have neither given nor received unauthorized aid on this test or assignment."

Disabilities

Reasonable accommodations will be made for students with verifiable disabilities. In order to take advantage of available accommodations, students must register with Disability Services for Students at 1900 Student Health Center, Campus Box 7509, 515-7653. For more information on NC State's policy on working with students with disabilities, please see the Academic Accommodations for Students with Disabilities Regulation (REG02.20.01).

Software

We will be using Stata IC for the course. Besides R, Stata is the only package that allows researchers to fully, as well as easily, use the techniques that we will be studying. Please contact me immediately if you have not used Stata before.

You have three options for using Stata:

1. Virtual Computing Lab (VCL) - With the VCL, you make an advance reservation to have Stata run virtually on your home computer, during class and whenever you need it to work on an assignment. Information on using the VCL can be found here:
   - https://vcl.ncsu.edu/help
   - https://sites.google.com/a/ncsu.edu/cedstata
2. Rent – you can rent Stata for either six-months ($65) or one year ($98).
3. Purchase – you can purchase a copy of Stata for $179; this is a perpetual license, meaning that you will only pay this fee once.

You can rent or purchase Stata here: http://www.stata.com/order/new/edu/gradplans/direct-ship-pricing.

The college is in the midst of updating its license for Stata 13. We may end up using version 13 for one part of the course, because it looks like they have a new set of procedures for estimating treatment effects. We will discuss how to handle this in our first meeting, for those of you who have purchased version 12 and don't want to purchase version 13.

Miscellaneous background material

If you need a quick refresher of logistic regression (useful for the matching section), see this webpage: http://www.ats.ucla.edu/stat/stata/webbooks/logistic/chapter1/statalog1.htm

If you are still confused about the expectation operator, see this Khan Academy video and pay careful attention to Chapter 2 of *Field Experiments*:
https://www.khanacademy.org/math/probability/random-variables-topic/random_variables_prob_dist/v/expected-value--e-x

The term "selection bias" is used in a variety of ways by researchers. Technically, *selection bias* refers to some sample selection process on the part of the researcher that biases the resulting analyses. For example, you might study time to degree, but you exclude students who drop out or who have not yet graduated by the time you collect data. Here, you have selected your sample on the basis of your dependent variable, which is generally a poor idea. However, many people use the term "selection bias" to refer to *self-selection bias*, in which units select themselves into treated and control groups. We will be focusing on the latter in this course. For more information see:

- http://en.wikipedia.org/wiki/Selection_bias
- http://en.wikipedia.org/wiki/Self-selection_bias

Readings and schedule

I would prefer to cover 75% of the syllabus very well, rather than cover 100% and have students leaving the course scratching their heads, so I will change the schedule as needed. Any changes to the course schedule will be discussed in class.

The main text for this course is

Khandker, S.R. et al. (2010). *Handbook on impact evaluation: quantitative methods and practices*. Washington, DC: World Bank. Also available online: http://www-wds.worldbank.org/external/default/WDSContentServer/WDSP/IB/2009/12/10/000333037_20091210014322/Rendered/PDF/520990PUB0EPI1101Official0Use0Only1.pdf

The course is designed to alternate weeks. The first one to two weeks will cover the theoretical background of a technique, with Stata applications in class. The following week will be a critical discussion of several research articles using the technique; the take-homes will usually be posted after that class. I have also listed some further readings; these <u>optional</u> readings are included for your future reference.

Important dates:

- October 10th – no class, Fall Break
- November 14th – no class, ASHE conference
- November 28th – no class, Thanksgiving
- December 5th – last class
- December 12th – literature reviews due at noon

I. Causality and educational research

A. Background

1. Stock, J.H. & Watson, M.W. (2006). *Introduction to Econometrics*, chapter 9, "Assessing studies based on multiple regression," sections 9.1 & 9.2
2. Rudalevige, A. (2009). Juggling act: the politics of science in education research. *Education Next*, (9)1, 34-41.
3. IES *Education Research RFA 84.305A*, pp. 57-68.
4. Khandker – chapters 1 & 2.
5. Gerber, A.S. & Green, D.P. (2012). *Field Experiments*, chapter 2, "Causal inference and experimentation."
6. Khandker – chapter 3.
7. West, S.G. and Thoemmes, F. (2011). Campbell's and Rubin's perspectives on causal inference. *Psychological Methods*, 15(1), 18-37.
8. Murnane, R.J. and Willett, J.B. (2011). *Methods Matter*, chapter 8, "Using natural experiments to provide 'arguably exogenous' treatment variability."

B. Applications

1. Bettinger, E. and Baker, R. (2011). The effects of student coaching in college: An evaluation of a randomized experiment in student mentoring. National Bureau of Economic Research.
2. Schnell, C.A. and Doetkott, C.D. (2003). First-year seminars produce long-term impact. *Journal of College Student Retention*, 4(4), 377-391.

3. Strayhorn, T.L. (2009). An examination of the impact of first-year seminars on correlates of college student retention. *Journal of the First-Year Experience & Students in Transition*, 21(1), 9-27.

C. Further readings

1. Newey, W. (2007). *Treatment effects*. Course materials for 14.386 New Econometric Methods (http://ocw.mit.edu/courses/economics/14-386-new-econometric-methods-spring-2007/readings/treatment_effect.pdf).
2. Rubin, D.B. (2010). Reflections stimulated by the comments of Shadish (2010) and West and Thoemmes (2010). *Psychological Methods, 15*(1), 38-46.

II. Matching

A. Background

1. Khandker, chapter 4.
2. Ho, D.E., Imai, K., King, G., & Stuart, E.A. (2007). Matching as nonparametric preprocessing for reducing model dependence in parametric causal inference. *Political Analysis, 15*(3), 199-236
3. Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for implementation of propensity score matching. *Journal of Economic Surveys, 22*(1), 31-72.
4. Lee, J.C. & Staff, J. (2007). When work matters: The varying impact of work intensity on high school dropout. *Sociology of Education, 80*(2), 158-178. **[We will be using their dataset in class]**

B. Applications

1. Wilde, E.T., & Hollister, R. (2007). How close is close enough? Evaluating propensity score matching using data from a class size reduction experiment. *Journal of Policy Analysis and Management*, 26(3), 455-477.
2. Shadish,W.R. et al. (2008). Can nonrandomized experiments yield accurate answers? A randomized experiment comparing random and nonrandom assignments. *Journal of the American Statistical Association, 103(*484), 1334-1356.
3. Retelsdorf et al. (2012). Reading development in a tracked school system: A longitudinal study over 3 years using propensity score matching. *British Journal of Educational Psychology, 82*, 647–671.
4. Berends et al. (2010). Instructional conditions in charter schools and students' mathematics achievement gains. *American Journal of Education*, 116(3), 303-335.
5. Schudde, L.T. (2011). The causal effect of campus residency on college student retention. *Review of Higher Education, 34*(4), 581-610.

C. Further readings

1. Bia, M., & Mattei, A. (2008). A Stata package for the estimation of the dose-response function through adjustment for the generalized propensity score. *Stata Journal*, 8(3), 354-373.

2. DuGoff, E.H. et al. (2013). Generalizing observational study results: Applying propensity score methods to complex surveys. *Health Services Research*.
3. Iacus, S.M., King, G. & Porro, G. (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis, 20*(1), 1-24.
4. Stata documentation: *teffects intro — introduction to treatment effects for observational data.*
5. Stata documentation: *teffects intro advanced — advanced introduction to treatment effects for observational data.*

Difference-in-difference and fixed effects models

A. Background

1. Khandker, chapter 5.
2. Wooldridge, J.M. (2013) *Introductory Econometrics*, chapter 13, "Pooling cross sections across time: Simple panel data methods," and chapter 14, "Advanced panel data methods."
3. Halaby, C.N. (2004). Panel models in sociological research: Theory into practice. *Annual Review of Sociology*, 30, 507–544.
4. Schilt, K., & Wiswall, M. (2008). Before and after: Gender transitions, human capital, and workplace experiences. *B.E. Journal of Economic Analysis & Policy, 8*(1), 1-26. **[We will be using their dataset in class.]**
5. Stata documentation: *xt — introduction to xt commands*
6. Stata documentation: *xtset — declare data to be panel data.*

B. Applications

1. Dynarski, S.M. (2003). Does aid matter? Measuring the effect of student aid on college attendance and completion. *American Economic Review*, *93*(1), 279-288.
2. Garces, E., Thomas, D. & Currie, J. (2002) Longer-term effects of Head Start. *American Economic Review*, *92*(4), 999-1012.
3. Dee, T.S., Jacob, B. & Schwartz, N.L. (2012). The effects of NCLB on school resources and practices. *Educational Evaluation and Policy Analysis, 35*(2), 252-279.
4. Figlio, D.N. (2006). Testing, crime and punishment. *Journal of Public Economics, 90*, 837-851.

C. Further readings

6. Lechner, M. (2010). The estimation of causal effects by difference-in-difference methods. *Foundations and Trends in Econometrics*, 4(3), 165-224.
7. Arai, M., & Thoursie, P.S. (2009). Renouncing personal names: An empirical examination of surname change and earnings. *Journal of Labor Economics, 27*(1), 127-147. [Just a cool application of panel methods to investigate discrimination.]

Instrumental variables

A. Background

1. Khandker, chapter 6.
2. Murray, M.P. (2006). Avoiding invalid instruments and coping with weak instruments. *Journal of Economic Perspectives*, 20(4), 111-132.
3. Porter, S.R. (2013). Understanding the LATE approach to instrumental variables.
4. Card. D. (1993). Using geographic variation in college proximity to estimate the return to schooling. *NBER Working Paper* No. 4483. **[Read only pp. 1-13 – we will be using his dataset in class]**
5. Stata documentation: *ivregress postestimation — postestimation tools for ivregress*

B. Applications

1. Stinebrickner, R., & Stinebrickner, T.R. (2008). The causal effect of studying on academic performance. *B.E. Journal of Economic Analysis & Policy, 8*(1), 1-53.
2. Long, B.T. & Kurlaender, M. (2009). Do community colleges provide a viable pathway to a baccalaureate degree? *Educational Evaluation and Policy Analysis, 31*(1), 30-53.
3. Currie, J. & Moretti, E. (2003). Mother's education and the intergenerational transmission of human capital: Evidence from college openings. *Quarterly Journal of economics*, 1495-1532
4. Rees, D.I. & Sabia, J.J. (2010). Sports participation and academic performance: Evidence from the National Longitudinal Study of Adolescent Health. *Economics of Education Review, 29*, 751-759.

C. Further reading

1. Angrist, J.D. & Krueger, A.B. (2001). Instrumental variables and the search for identification: From supply and demand to natural experiments. *Journal of Economic Perspectives, 15*(4) 69-85.
2. Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American Statistical Association, 91*:444–455. [I try to translate this article into everyday language in my paper]

Regression discontinuity

A. Background

1. Khandker, chapter 7.
2. Imbens, G.W., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics, 142*, 615-635.
3. Lee, D.S., & Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of Economic Literature, 48*, 281-355.

B. Applications

1. Carrell, S.E., Hoekstra, M. & West, J.E. (2011). Does drinking impair college performance? Evidence from a regression discontinuity approach. *Journal of Public Economics, 95*, 54-62.
2. Schanzenbach, D. (2009). Do school lunches contribute to childhood obesity? *Journal of Human Resources, 44*(3), 684-709.

3. Mariano, L.T., & Martorell, Paco. (2013). The academic effects of summer instruction and retention in New York City. *Educational Evaluation and Policy Analysis*, 35(1), 96-117.
4. Ludwig, J. & Miller, D. (2007). Does Head Start improve children's life chances? Evidence from a regression discontinuity design. *Quarterly Journal of Economics, 122*(1), 159-208
5. Lindo, J.M., Sanders, N.J., & Oreopoulos,P. (2010). Ability, gender, and performance standards: Evidence from academic probation. *American Economic Journal: Applied Economics, 2*(2), 95-117.

C. Further reading

1. Van der Klaauw, W. (2008). Regression-discontinuity analysis: A survey of recent developments in economics. *Labour, 22*(2), 219-245.

Disproving someone else's causal analysis (we will cover this topic only if time permits)

1. Cohen-Cole et al. (2008). Detecting implausible social network effects in acne, height, and headaches: longitudinal analysis. *BMJ, 337*(a2533), 1-5.
2. Simonsohn, U. (2011). Spurious? Name similarity effects (implicit egotism) in marriage, job, and moving decisions. *Journal of Personality and Social Psychology, 101*(1), 1-24.