EKT-816: Microeconometrics
Semester II, 2020

Instructors: Steve Koch (**steve.koch@up.ac.za**) and Jesse Naidoo (**jesse.naidoo@up.ac.za**)
Office Hours: TBA

Lectures are from 4.30 PM - 7.30 PM on Monday evenings, online. We will use the Blackboard Collaborate feature for the lectures and class discussions.

Problem Sets: We will post four problem sets (about one every three weeks). ***These are not for credit.*** We will also provide suggested solutions.

Grades will be determined by a weighted average of your scores on the final exam and the midterm. The respective weights will be either (40% midterm, 60% final), or (25% midterm, 75% final) - whichever is in your favor.

Important Dates

Midterm: September 21

Final Exam: TBA.

Textbooks

For the most part we will follow Angrist and Pischke's *Mostly Harmless Econometrics* closely. Useful supplementary texts are

- *Microeconometrics: Methods and Applications*, by Cameron and Trivedi
- *Econometric Analysis of Cross-Section and Panel Data*, by Wooldridge

We will partly base our lectures on these texts, but may present the content in a slightly different way in class. Study several different sources (the internet may be useful here) and - most importantly! - do the problems to make sure you understand.

Syllabus

Counterfactuals [1 lecture]

Potential outcomes and causality. Varieties of (ex-post) treatment effects. Causal diagrams.

References: Angrist & Pischke (2008) Ch. 1-2.

Additional Readings: Heckman (2000); Heckman & Vytlacil (2001).

Selection on Observables, Conditional Independence, and Regression [2 lectures]

Derivation of the OLS formula; Frisch-Waugh-Lovell theorem. Implications for interpreting regression output. Conditional expectations and best linear predictors. Typical violations of conditional independence: simultaneity, measurement error, self-selection. Basic similarities between matching and regression. Discrete choice models. The propensity score.

References: Angrist & Pischke (2008) Ch. 3.

Additional Readings: Abadie & Imbens (2006); Abadie & Imbens (2008); Abadie & Imbens (2016).

Software References: Keller & Tipton (2016).

Online Application Lecture [Video and Collaborate discussion]

Data management. Markdown. Estimation and matching in R. Tables and Figures.

Instrumental Variables, Selection Models, and LATE [3 lectures]

Just-identified IV; 2SLS and GMM. Identification and (asymptotic) inference. 2SLS bias; misleading inference with weak instruments. Outcome and selection equations in the "Generalized Roy Model". LATE and MTE. Identification via non-standard moments. Control functions.

References: Angrist & Pischke (2008) Ch. 4.

Additional Readings: Pearl (1995); Dong (2010); Heckman & Navarro-Lozano (2004); Heckman *et al.* (2006); Carneiro *et al.* (2010); Lewbel (2012); Lewbel (2018).

Software References: Gui *et al.* (2020); Torgovitsky & Shea (2020).

Online Application Lecture [Video and Collaborate discussion]

Markdown. IV and 2SLS. REndo. Marginal treatment effects. DAGs.

Tentative Midterm [Coverage: First two topics.]

Regression Discontinuity Designs [1 lecture]

Strict RD. Fuzzy RD. Running variable manipulation.

References: Angrist & Pischke (2008) Ch. 6.; McCrary (2008); Imbens & Kalyanaraman (2012); Eggers *et al.* (2015); Dimmery (2016); Gelman & Imbens (2019).

Synthetic Panels and Differences-in-Differences [2 lecture]

An additively separable model for Y_0 and the parallel trends assumption. Pitfalls: heterogeneous trends and serially correlated shocks. Placebo tests and placebo outcomes (DDD). Clustering and bootstrapping standard errors.

References: Angrist & Pischke (2008) Ch. 5, 8.

Additional Readings: Deaton (1985); Nijman & Verbeek (1990); Verbeek & Nijman (1992); Moffit (1993).

Online Application Lecture [Video and Collaborate discussion]

Regression discontinuity. Building synthetic panel. Differences-in-Differences. Placebo testing. Bootstrapping.

Quantile Regression [1 lecture]

Check Function. Location-Scale Models.

References: Angrist & Pischke (2008) Ch. 7.

Software Reference: Koenker (2020).

Nonparametric Densities and Regression [1 lecture]

Location-Scale Models. Density and regression. Bandwidth. Cross-validation. Bias and variance. Simple hypothesis testing.

References: Racine (2008).

Additional Readings: Racine & Li (2017).

Software References: Hayfield & Racine (2008).

item[Decomposition and Recentered Influence Functions] [1 lecture]

Regression decomposition. Recentered influence function.

References: Firpo *et al.* (2018).

Additional Readings: Firpo *et al.* (2009); Heckley *et al.* (2016)

Software References: Schulenberg (2018).

Online Application Lecture [Video and Collaborate discussion]

Quantile regression. Density estimation. NP regression estimation. Hypothesis testing.

Final Exam [Coverage: All]

One class will be taken up by the midterm, one by a pre-midterm review session, and one by a pre-exam review session.

References

Abadie, Alberto, & Imbens, Guido W. 2006. Large Sample Properties of Matching Estimators for Average Treatment Effects. *Econometrica*, **74**(1), 253–267.

Abadie, Alberto, & Imbens, Guido W. 2008. On the Failure of the Bootstrap for Matching Estimators. *Econometrica*, **76**(6), 1537–1557.

Abadie, Alberto, & Imbens, Guido W. 2016. Matching on the Estimated Propensity Score. *Econometrica*, **84**(2), 781–807.

Angrist, Joshua D, & Pischke, Jörn-Steffen. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Carneiro, Pedro, Heckman, James J, & Vytlacil, Edward J. 2010. Evaluating Marginal Policy Changes and the Average Effect of Treatment for Individuals at the Margin. *Econometrica*, **78**(1), 377–394.

Deaton, Angus. 1985. Panel Data from Time Series of Cross-Sections. *Journal of Econometrics*, **30**(1–2), 109–126.

Dimmery, Drew. 2016. *rdd: Regression Discontinuity Estimation*. R package version 0.57.

Dong, Yingyang. 2010. Endogenous Regressor Binary Choice Models Without Instruments, with an Application to Migration. *Economics Letters*, **107**(1), 33–35.

Eggers, Andrew C., Fowler, Anthony, Hainmueller, Jens, Hall, Andrew B., & Snyder, Jr., James M. 2015. On the Validity of the Regression Discontinuity Design for Estimating Electoral Effects: New Evidence from Over 40,000 Close Races. *Amerian Journal of Political Science*, **59**(1), 259–274.

Firpo, Sergio, Fortin, Nicole M., & Lemieux, Thomas. 2009. Unconditional Quantile Regressions. *Econometrica*, **77**(3), 953–973.

Firpo, Sergio P, Fortin, Nicole M, & Lemieux, Thomas. 2018. Decomposing Wage Distributions Using Recentered Influence Function Regressions. *Econometrics*, **6**(2), 28.

Gelman, Andrew, & Imbens, Guido W. 2019. Why High-Order Polynomials Should Not be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, **37**(3), 447–456.

Gui, Raluca, Meierer, Markus, Algesheimer, Rene, & Schilter, Patrik. 2020. *REndo: Fitting Linear Models with Endogenous Regressors using Latent Instrumental Variables*. R package version 2.3.1.

Hayfield, Tristen, & Racine, Jeffrey S. 2008. Nonparametric econometrics: the np package. *Journal of Statistical Software*, **27**(5), 1–32.

Heckley, Gawain, Gerdtham, Ulf-G., & Kjellssond, Gustav. 2016. A general method for decomposing the causes of socioeconomic inequality in health. *Journal of Health Economics*, **48**, 89–106.

Heckman, James J. 2000 (December). *Microdata, Heterogeneity and the Evaluation of Public Policy*. Nobel Prize in Economic Sciences Lecture.

Heckman, James J, & Navarro-Lozano, Salvador. 2004. Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models. *Review of Economics and Statistics*, **86**(1), 30–57.

Heckman, James J, & Vytlacil, Edward J. 2001. Policy-relevant Treatment Effects. *American Economic Review*, **91**(2), 107–111.

Heckman, James J, Urzua, Sergio, & Vytlacil, Edward J. 2006. Understanding Instrumental Variables in Models with Essential Heterogeneity. *Review of Economics and Statistics*, **88**(3), 389–432.

Imbens, Guido W, & Kalyanaraman, Karthik. 2012. Optimal Bandwidth Choice for the Regression Discontinuity Estimator. *The Review of Economic Studies*, **79**(3), 933–959.

Keller, Bryan, & Tipton, Elizabeth. 2016. Propensity Score Analysis in R: A Software Review. *Journal of Educational and Behavioral Statistics*, **41**(3), 326–348.

Koenker, Roger. 2020. *quantreg: Quantile Regression*. R package version 5.55.

Lewbel, Arthur. 2012. Using Heteroscedasticity to Identify and Estimate Mismeasured and Endogenous Regressor Models. *Journal of Business & Economic Statistics*, **30**(1), 67–80.

Lewbel, Arthur. 2018. Identification and Estimation Using Heteroscedasticity Without Instruments: The Binary Endogenous Regressor Case. *Economics Letters*, **165**(1), 10–12.

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

Moffit, Robert. 1993. Identification and Estimation of Dynamic Models with a Time Series of Repeated Cross Sections. *Journal of Econometrics*, **59**(1–2), 99–123.

Nijman, Theo, & Verbeek, Marno. 1990. Estimation of Time-Dependent Parameters in Linear Models Using Cross-Sections, Panels, or Both. *Journal of Econometrics*, **46**(3), 333–346.

Pearl, Judea. 1995. Causal Diagrams for Empirical Research. *Biometrika*, **82**(4), 669–688.

Racine, Jeffrey S. 2008. Nonparametric Econometrics: A Primer. *Foundations and Trends(R) in Econometrics*, **3**(1), 1–88.

Racine, Jeffrey S, & Li, Kevin. 2017. Nonparametric Conditional Quantile Estimation: A Locally Weighted Kernel Approach. *Journal of Econometrics*, **201**(1), 72–94.

Schulenberg, René. 2018. *dineq: Decomposition of (Income) Inequality*. R package version 0.1.0.

Torgovitsky, Alexander, & Shea, Joshua. 2020. *ivmte: Instrumental Variables: Extrapolation by Marginal Treatment Effects*. R package version 1.2.0.

Verbeek, Marno, & Nijman, Theo. 1992. Can Cohort Data Be Treated As Genuine Panel Data? *Empirical Economics*, **17**(1), 9–23.