Harvard University **Fall 2016**
Department of Economics **Ec2140**

Econometric Methods: Syllabus

Version 9 (December 1, 2016)

Instructor: Mikkel Plagborg-Moller, plagborg@fas.harvard.edu
Lectures: Tue/Thu 2.30–4.00 pm, Littauer M-16
Office hours: Wed 4.00–6.00 pm and by appointment, Littauer M-29

Teaching Fellow: Jann Spiess, jspiess@fas.harvard.edu
Sections: Mon 3.00–4.00 pm, Science Center 216; Wed 6.00–7.00 pm, Littauer M-17
Office hours: Wed 7.00–8.00 pm, Littauer M-17, and by appointment

Description. Econometric methods for cross-sectional data. Topics include extremum estimators, discrete choice models, the bootstrap, nonparametric methods, quantile regression, panel data, treatment effect estimation, and Bayesian computation.

Prerequisites. Ec2120 (or equivalent with the explicit permission of the instructor).

Material. The course material is self-contained and there is no required textbook for the course. Handouts covering most of the material will be available on the website. Like the basic structure of the course, the handouts borrow heavily from material generously shared by Professor Alberto Abadie, although any errors are the sole responsibility of the instructor. Some students might find it useful to have a textbook as an additional reference. Good reference books are:

Cameron, A. C. and Trivedi, P. K., (2005) *Microeconometrics: Methods and Applications*, Cambridge University Press.

Wooldridge, J. M. (2010), *Econometric Analysis of Cross Section and Panel Data*, 2nd edition, MIT Press.

This syllabus also includes a list of additional readings that are useful for a deeper understanding of the material. Many of these are available electronically.

Homework. There will be four work-intensive problem sets dealing with practical data analysis and theoretical calculations. Students will have about three weeks to finish each set. Students may use any software, including Stata where appropriate, but are encouraged to try out Matlab, R, Python, Julia, or equivalent. Students may collaborate in groups of at most four people but must write up and hand in answers and code independently. Please indicate the group of collaborators on each problem set. Problem sets should be printed out and handed in to the Teaching Fellow by the start of class on the due dates.

Final paper. In lieu of a midterm and final exam, each student must hand in a final paper. Collaboration is not allowed, except for big-picture discussion. The paper is expected to be about 20–25 double-spaced pages and its outline should be one or more of the following three types: (1) a theoretical econometrics paper (could be a careful simulation study), (2) an empirical paper employing advanced econometrics, (3) a replication and robustness check of a recent econometrically advanced empirical study. The insight(s) of the final paper must be marginally new to the literature. Clarity of argument and execution of the main specification are more important than robustness/extensions. The instructor must explicitly approve the paper topic in advance. The final paper should be submitted to the instructor by email (PDF format) or during office hours.

Grading. The course grade is determined by the average point score on the four problem sets (weight 70%) and the point score on the final paper (weight 30%). During grading of problem sets, a penalty of up to 25% of a given problem's total point value may be deducted due to bad exposition (e.g., missing documentation, impractical presentation of empirical results, hard-to-read code). If you find a grading error, please resubmit the problem set along with a 1-paragraph explanation; we will then regrade the entire problem set.

Code of conduct. All course activities, including class meetings and homework assignments, are subject to the university's academic code and code of conduct as detailed in the GSAS Handbook's Regulations and Standards of Conduct.

Accommodations for students with disabilities. Students needing academic adjustments or accommodations because of a documented disability must present their Faculty Letter from the Accessible Education Office (AEO) and speak with the instructor by the end of the second week of the term, September 9. Failure to do so may result in the in-

structor's inability to respond in a timely manner. All discussions will remain confidential, although the instructor may contact the AEO to discuss appropriate implementation.

Important dates. These dates are preliminary. Any changes will be announced via course email.

Aug 31 (Wed): First instructor office hours

Sep 1 (Thu): First class

Sep 7 (Wed): First section

Sep 22 (Thu): **No class**

Sep 29 (Thu): Problem Set 1 due at start of class

Oct 27 (Thu): Problem Set 2 due at start of class

Nov 10 (Thu): Problem Set 3 due at start of class

Nov 16 (Wed): Deadline for getting instructor's approval on paper topic

Nov 23 (Wed) and Nov 24 (Thu): Thanksgiving break

Nov 30 (Wed): Last instructor office hours

Dec 6 (Tue): Last class, Problem Set 4 due at start of class (note: during reading week)

Dec 15 (Thu): Final paper due by 2.00 pm

Course outline. This outline is preliminary and may change without warning.

1. Statistical preliminaries
    (a) Population and sample characteristics, the analogy principle
    (b) Causality, identification, inference
    (c) Stochastic convergence
    (d) Regression analysis, clustered standard errors
2. Extremum estimators

   (a) Maximum likelihood, generalized methods of moments, nonlinear least squares, minimum distance
   (b) Numerical optimization
   (c) Large sample properties, two-step estimators
   (d) Efficiency, testing
3. Bootstrap
4. Discrete choice models
   (a) Binary choice
   (b) Multinomial choice
5. Nonparametric methods
   (a) Nonparametric density estimation
   (b) Nonparametric regression
   (c) Semiparametric methods
6. Quantile regression
7. Panel data
   (a) Linear static panel data models, fixed effects, random effects, Hausman test
   (b) Linear dynamic panel data models
   (c) Nonlinear panel data models
8. Estimation of treatment effects
   (a) Counterfactuals, potential outcomes
   (b) Randomized experiments
   (c) Observational studies
   (d) Selection on observables, matching, inverse probability weighting
   (e) Robustness, sensitivity, falsification
   (f) Instrumental variables, local average treatment effects

(g) Partial identification of treatment effects

(h) Differences-in-differences, synthetic controls

(i) Regression discontinuity

(j) Distributional treatment effects (time permitting)

9. Bayesian inference and computation

(a) Markov Chain Monte Carlo

(b) Pooling/shrinkage, hierarchical models

(c) Comparison with frequentist inference

10. Weak instruments (time permitting)

11. Model selection (time permitting)

List of readings

Introductory readings are listed first and marked with a star (*). Other readings are included for your reference. The reading list is preliminary and may change without warning.

1 Statistical preliminaries

* Cameron and Trivedi: Chapter 2, Appendix A.

Imbens, G. W., and Kolesár, M. (forthcoming), "Robust Standard Errors in Small Samples: Some Practical Advice," *Review of Economics and Statistics*.

van der Vaart, A. W. (1998), Asymptotic Statistics. Cambridge University Press. Chapters 2–3.

Wooldridge: Chapters 3–4 and 20.

2 Extremum estimators

* Cameron and Trivedi: Chapters 5–6 and 10.

* Newey, W. K., and McFadden, D. (1994), "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics Vol. 4*, ed. by Engle, R. F. and McFadden, D. L. Elsevier. Sections 1–6 and 9.

Wooldridge: Chapters 13–14.

3 Bootstrap

* Cameron and Trivedi: Chapter 11.

Efron, B. and Tibshirani, R. J. (1993), An Introduction to the Bootstrap. Chapman and Hall. Chapters 6 and 12–13.

Horowitz, J. L. (2001), "The Bootstrap," in *Handbook of Econometrics Vol. 5*, ed. by Heckman, J. J. and Leamer, E. E. Elsevier. Sections 1–3.

van der Vaart, A. W. (1998), Asymptotic Statistics. Cambridge University Press. Chapter 23.

4 Discrete choice models

* Cameron and Trivedi: Chapters 14–15.

McFadden, D. (1973), "Conditional Logit Analysis of Qualitative Choice Behavior," in *Frontiers of Econometrics*, ed. by Zarembka, P. Academic Press.

Train, K. E. (2009), Discrete Choice Methods with Simulation, 2nd edition. Cambridge University Press. Chapters 1–6. Available online at:
`http://elsa.berkeley.edu/books/choice2.html`

Wooldridge: Chapters 15–16.

5 Nonparametric methods

* Cameron and Trivedi: Chapter 9.

Chen, X. (2007), "Large Sample Sieve Estimation of Semi-Nonparametric Models," in *Handbook of Econometrics Vol. 6B*, ed. by Heckman, J. J. and Leamer, E. E. Elsevier. Sections 1–2.

DiNardo, J. and Tobias, J. L. (2001), "Nonparametric Density and Regression Estimation," *Journal of Economic Perspectives*, vol. 15, 11–28.

Green, P. J. and Silverman, B. W. (1993), Nonparametric Regression and Generalized Linear Models: A roughness penalty approach. CRC Press.

Härdle, W. and Linton, O. (1994), "Applied Nonparametric Methods," in *Handbook of Econometrics Vol. 4*, ed. by Engle, R. F. and McFadden, D. L. Elsevier.

Ichimura, H. and Todd, P. E. (2007), "Implementing Nonparametric and Semiparametric Estimators," in *Handbook of Econometrics Vol. 6B*, ed. by Heckman, J. J. and Leamer, E. E. Elsevier. Sections 1–6.

Newey, W. K. (1990), "Semiparametric efficiency bounds," *Journal of Applied Econometrics*, vol. 5(2), 99–135.

Pagan, A. and Ullah, A. (1999), Nonparametric Econometrics. Cambridge University Press. Chapters 1–3.

Powell, J. L. (1994), "Estimation of Semiparametric Models," in *Handbook of Econometrics Vol. 4*, ed. by Engle, R. F. and McFadden, D. L. Elsevier.

Wasserman, L. (2006), All of Nonparametric Statistics. Springer. Chapters 4–6.

6 Quantile regression

* Buchinsky, M. (1998), "Recent Advances in Quantile Regression Models: A Practical Guideline for Empirical Research," *Journal of Human Resources*, vol. 33, 88–126.

* Cameron and Trivedi: Section 4.6.

Angrist, J. D., Chernozhukov, V., and Fernández-Val, I. (2006), "Quantile Regression under Misspecification, with an Application to the U.S. Wage Structure," *Econometrica*, vol. 74, 539–563.

Buchinsky, M. (1994), "Changes in the U.S. Wage Structure 1963-1987: Application of Quantile Regression," *Econometrica*, vol. 62, 405–458.

Chamberlain, G. (1994), "Quantile Regression, Censoring, and the Structure of Wages," in *Advances in Econometrics, Sixth World Congress, Vol. 1*, ed. by C. Sims. Cambridge University Press.

Chernozhukov, V. and Hansen, C. (2008), "Instrumental variable quantile regression: A robust inference approach," *Journal of Econometrics*, vol. 142(1), 379–398.

Koenker, R. (2005), Quantile Regression. Cambridge University Press.

Koenker, R. and Hallock, K. F. (2001), "Quantile Regression," *Journal of Economic Perspectives*, vol. 15(4), 143–156.

7 Panel data

* Cameron and Trivedi: Chapters 21–23.

Arellano, M. (2003), Panel Data Econometrics. Oxford University Press.

Arellano, M. and Honoré, B. (2001), "Panel Data: Some Recent Developments," in *Handbook of Econometrics Vol. 5*, ed. by Heckman, J. J. and Leamer, E. E. Elsevier.

Chamberlain, G. (1984), "Panel Data," in *Handbook of Econometrics Vol. 2*, ed. by Grilliches, Z. and Intrilligator, M. D. Elsevier.

Wooldridge: Chapters 10–11 and 15.8.

8 Estimation of treatment effects

The following three readings are overviews of the material that we will cover in this section.

* Imbens, G. W. and Wooldridge, J. M. (2009), "Recent Developments in the Econometrics of Program Evaluation," *Journal of Economic Literature*, vol. 47(1), 5–86.

Angrist, J. D. and Pischke, J. S. (2009), Mostly Harmless Econometrics: An Empiricist's Companion. Princeton University Press.

Imbens, G. W. and Rubin, D. B. (2015), Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction. Cambridge University Press.

For a thorough discussion with many opposing viewpoints on causal empirical methodology, see the symposium "Con Out of Econometrics" in the *Journal of Economic Perspectives*, 2010, vol. 24(2), pages 3–94. Available online at: `https://www.aeaweb.org/issues/126`

Counterfactuals, potential outcomes

Holland, P. W. (1986), "Statistics and Causal Inference" (with comments and rejoinder), *Journal of the American Statistical Association*, vol. 81, 945–970.

Randomized experiments

Abadie, A., Athey, S., Imbens, G. W., and Wooldridge, J. M. (2014), "Finite Population Causal Standard Errors," working paper. Available online at: `http://economics.mit.edu/files/11855`

Duflo, E., Glennerster, R., and Kremer, M. (2008), "Using Randomization in Development Economics Research: A Toolkit," in *Handbook of Development Economics Vol. 4*, ed. by Schultz, T. P. and Strauss, J. A. Elsevier.

Rosenbaum, P. R. (2002), Observational Studies, 2nd edition. Springer. Chapter 2.

Observational studies

* Rosenbaum, P. R. (2010), Design of Observational Studies. Springer. Chapters 1.1–1.6.

Selection on observables, matching, inverse probability weighting

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

Abadie, A. and Imbens, G. W. (2006), "Large Sample Properties of Matching Estimators for Average Treatment Effects," *Econometrica*, vol. 74, 235–267.

Abadie, A. and Imbens, G. W. (2008), "On the Failure of the Bootstrap for Matching Estimators," *Econometrica*, vol. 76, 1537–1557.

Dehejia, R. H. and Wahba, S. (1999), "Causal Effects in Non-Experimental Studies: Re-Evaluating the Evaluation of Training Programs," *Journal of the American Statistical Association*, vol. 94, 1053–1062.

Heckman, J. J., Ichimura, H., and Todd, P. E. (1997), "Matching as an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Programme," *Review of Economic Studies*, vol. 64, 605–654.

Pearl, J. (2009), Causality, 2nd edition. Cambridge University Press.

Rosenbaum, P. R. and Rubin, D. B. (1983), "The Central Role of the Propensity Score in Observational Studies for Causal Effects," *Biometrika*, vol. 70, 41-55.

Rubin, D.B. (1977), "Assignment to Treatment Group on the Basis of a Covariate," *Journal of Educational Statistics*, vol. 2, 1–26.

Robustness, sensitivity, falsification

Athey, S. and Imbens, G. W. (2016), "The State of Applied Econometrics – Causality and Policy Evaluation," working paper. Section 3. Available online at:
`http://arxiv.org/abs/1607.00699v1`

Imbens, G. W. (2003), "Sensitivity to Exogeneity Assumptions in Program Evaluation," *American Economic Review (Papers & Proceedings)*, vol. 93(2), 126–132.

Instrumental variables, local average treatment effects

* Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996), "Identification of Causal Effects Using Instrumental Variables," *Journal of the American Statistical Association*, vol. 91, 444–472.

Abadie, A. (2003), "Semiparametric Instrumental Variable Estimation of Treatment Response Models," *Journal of Econometrics*, vol. 113, 231–263.

Heckman, J. J. and Vytlacil, E. J. (1999), "Local instrumental variables and latent variable models for identifying and bounding treatment effects," *Proceedings of the National Academy of Sciences*, vol. 96(8), 4730–4734.

Imbens, G. W and Angrist, J. D. (1994), "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, vol. 62, 467–475.

Partial identification of treatment effects

* Manski, C. F. (1995), Identification Problems in the Social Sciences. Harvard University Press. Chapter 2.

Imbens G. W. and Manski, C. F. (2004), "Confidence Intervals for Partially Identified Parameters," *Econometrica*, vol. 72, 1845–1857.

Manski, C. F. (1990), "Nonparametric Bounds on Treatment Effects," *American Economic Review (Papers & Proceedings)*, vol. 80, 319–323.

Manski, C. F. (2003), Partial Identification of Probability Distributions. Springer. Chapters 2 and 7.

Tamer, E. (2010), "Partial Identification in Econometrics," *Annual Review of Economics*, vol. 2, 167–195.

Differences-in-differences, synthetic controls

Abadie, A. (2005), "Semiparametric Difference-in-Differences Estimators," *Review of Economic Studies,* vol. 72, 1–19.

Abadie, A., Diamond, A., and Hainmueller, J. (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, 493–505.

Abadie, A. and Gardeazabal, J. (2003), "The Economic Costs of Conflict: A Case Study of the Basque Country," *American Economic Review*, vol. 93(1), 113–132.

Athey, S. and Imbens, G. W. (2006), "Identification and Inference in Nonlinear Difference-in-Differences Models," *Econometrica*, vol. 74, 431–497.

Card, D. (1990), "The Impact of the Mariel Boatlift on the Miami Labor Market," *Industrial and Labor Relations Review*, vol. 44, 245–257.

Card, D. and Krueger, A. B. (1994), "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania," *American Economic Review*, vol. 84, 772–793.

Regression discontinuity

* Lee, D. S., and Lemieux, T. (2010), "Regression Discontinuity Designs in Economics," *Journal of Economic Literature*, vol. 48, 281–355.

Armstrong, T. and Kolesár, M. (2016), "Simple and Honest Confidence Intervals in Nonparametric Regression," working paper. Section 4. Available online at: `https://arxiv.org/abs/1606.01200`

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014), "Robust Nonparametric Confidence Intervals for Regression Discontinuity Designs," *Econometrica*, vol. 86(2), 2295–2326.

Hahn, J., Todd, P., and Van der Klaauw, W. (2001), "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design," *Econometrica*, vol. 69(1), 201–209.

Imbens, G. W. and Kalyanaraman, K. (2012), "Optimal Bandwidth Choice for the Regression Discontinuity Estimator," *Review of Economic Studies*, vol. 79, 933–959.

Imbens, G. W. and Lemieux, T. (2008), "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, vol. 142, 615–635.

Kolesár, M. and Rothe, C. (2016), "Inference in Regression Discontinuity Designs with a Discrete Running Variable," working paper. Available online at: `https://arxiv.org/abs/1606.04086`

Distributional treatment effects

Abadie, A. (2002), "Bootstrap Tests for the Effects of a Treatment on the Distribution of an Outcome Variable," *Journal of the American Statistical Association*, vol. 97, 284–292.

Abadie, A., Angrist, J. D., and Imbens, G. W. (2002), "Instrumental Variable Estimation of the Effects of Subsidized Training on the Quantiles of Trainee Earnings," *Econometrica*, vol. 70, 91–117.

Abadie, A., Chingos, M. M., and West, M. R. (2016), "Endogenous Stratification in Randomized Experiments," working paper. Available online at: `http://economics.mit.edu/files/11852`

Rothe, C. (2010), "Nonparametric estimation of distributional policy effects," *Journal of Econometrics*, vol. 155, 56–70.

9 Bayesian inference and computation

* Cameron and Trivedi: Chapter 13.

Chib, S. (2001), "Markov Chain Monte Carlo Methods: Computation and Inference," in *Handbook of Econometrics Vol. 5*, ed. by Heckman, J. J. and Leamer, E. E. Elsevier.

Efron, B. (2010), Large-Scale Inference: Empirical Bayes Methods for Estimation, Testing, and Prediction. Cambridge University Press. Chapter 1. Available online at: `http://statweb.stanford.edu/~ckirby/brad/LSI/monograph_CUP_edited2.pdf`

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., and Rubin, D. B. (2014), Bayesian Data Analysis, 3rd edition. CRC Press. Chapters 1, 4–5, 11–12, 15, and Appendix C.

Neal, R. M. (2011), "MCMC Using Hamiltonian Dynamics," in *Handbook of Markov Chain Monte Carlo*, ed. by Brooks, S., Gelman, A., Jones, G. L., and Meng, X.-L. CRC Press.

van der Vaart, A. W. (1998), Asymptotic Statistics. Cambridge University Press. Chapter 10.

10 Weak instruments

* Mikusheva, A. (2013), "Survey on statistical inferences in weakly-identified instrumental variable models," *Applied Econometrics*, vol. 29(1), 117–131.

Stock, J. H., Wright, J. H., and Yogo, M. (2002), "A Survey of Weak Instruments and Weak Identification in Generalized Method of Moments," *Journal of Business & Economic Statistics*, vol. 20(4), 518–529.

Stock, J. H. and Yogo, M. (2005), "Testing for Weak Instruments in Linear IV Regression," in *Identification and Inference for Econometric Models*, ed. by. Andrews, D. W. K. Cambridge University Press.

11 Model selection

* Claeskens, G. and Hjort, N. L. (2008), Model Selection and Model Averaging. Cambridge University Press. Chapters 1–4.

Leeb, H. and Pötscher, B. M. (2005), "Model Selection and Inference: Facts and Fiction," *Econometric Theory*, vol. 21(01), 21–59. Sections 1–2.