17.S953 New Methods for Causal Inference

Spring 2021

MIT

Instructor:	Teppei Yamamoto
Office (Zoom #):	550 928 2449
Email:	teppei@mit.edu
Office Hours:	By Appointment
Class Time:	Thursdays 9-11AM
Class (Zoom #):	955 5597 8923

Purpose and Goals

This is a graduate-level seminar class on recent advancements in the field of statistical methods for causal inference. The purpose of this class is to provide students with experience and skills that are necessary to conduct research on methodological topics professionally. Although the class focuses on methods for causal inference, many of the research skills students will learn in the class will be transportable to methodological research in other subfields. Specifically, the goal of this class is threefold:

1. *Getting used to reading technical journal articles in statistics and related methodological fields.* In typical methods classes taught in social science departments (e.g. the quant sequence in the Political Science Department), the instructor distills the raw materials and provides lecture notes or slides that are easier for students to digest. However, to write methodological papers, it is necessary to go beyond textbook or lecture materials and read the original journal articles, and students who come from social science backgrounds typically find the latter impenetrable. This class aims to fill this gap.

2. *Learning to how to write a methods paper.* Transitioning from a consumer to a producer of methodology requires specific training which is not simply learning a lot of technically difficult methods. Identifying a good research question is key like in any other research, but the rest of the process involves both art and science specific to methodological research. In this class, students who choose to write a methods final paper will learn how to do those well.

3. *Become proficient in causal inference methods that go beyond standard textbook materials.* Causal inference was once a highly specialized topic only a small subset of political methodologists ever worked on. Nowadays, every methodologist needs to know at least its basics, and it is a fertile ground for innovative research that is empirically relevant. In fact, it is one of the few areas of statistics where political methodologists have been part of the main drivers of innovation. This class exposes students to many of the recent developments in this field that are the most important.

After taking this class, students will be able to read typical articles from journals like the *Journal of the American Statistical Association* and *Political Analysis* quite comfortably. They will also be ready to embark on a methodological research project independently, particularly in the field of causal inference. Finally, they will also have built familiarity with cutting-edge causal inference methods potentially useful for their applied work.

Prerequisites

To enroll in this course, you need to have taken 17.800 (Quant I), 17.802 (Quant II) and 17.804 (Quant III) or their equivalents in other departments. It is also highly recommended that you either have taken 17.806 (Quant IV) or are enrolled in the said subject concurrently.

Organization

This class is a seminar-style class, meaning that there will be no regular lecture or problem sets. Instead, you read one or two articles *very closely* each week. We will meet once every week for two hours, and each class session will consist of three components of roughly equal duration. First, one student will present the assigned paper(s) to the class. You are encouraged to present the paper as if you were presenting your own work in a workshop: prior to the class, you should try to distill the content of the paper to the extent possible and create an effective presentation. You are encouraged to both use Beamer slides and show key derivations on a whiteboard. **(Note: Check if everyone has access to digital handwriting equipment if the class is taught virtually.)**

Second, after the presentation, everyone in the class will go through their own notes and we will collectively resolve any questions about the content of the paper. Due to the nature of the readings, the questions are expected to be technical, but we will prioritize questions that concern the core result or argument of the paper as opposed to minor details in mathematical derivations.

Finally, we will take a step back and discuss broader implications of the paper. We will discuss the paper's advantages, limitations, connections with other methods, and importantly, how it can be applied to substantive questions in political science. We will make sure to spend at least some time on potential applications since political methodology is an *applied statistics* subfield: we exist because we know both statistics and politics!

In addition to the class meetings, you will be working on a research project throughout the semester. Your project is typically expected to be methodological, and the goal is to have everyone publish a paper in *Political Analysis* based on the project before they graduate. Alternatively, you can work on an applied project, but only if it involves some innovative use of an advanced causal inference methodology. You will be giving one in-class presentation on your project toward the end of the semester, as well as submitting a final write-up at the end of the semester.

Since this is a small seminar class organized around class discussion, **your individual preparation and participation in each class is absolutely mandatory**. Preparation will involve submitting notes on the required readings on the day before each class (see below).

Requirements

The final grades are based on the following:

- **Class presentations** (40%): As described above, one student will be designated as the presenter in each class meeting. The presenter will also perform the role of a co-moderator (with the instructor) for the remainder of the class, where they are expected to lead the class discussion. Note that there are likely more class meetings than the number of enrolled students, meaning each student will be presenting multiple papers over the semester.

- **Class preparation, attendance and participation** (20%): This includes **uploading annotated PDF copies of the assigned reading** by 5pm two days before each class meeting (i.e., typically Tuesday). You should use Adobe Reader (free) or a similar PDF annotation tool (Notability for iOS is recommended and should be available for free via MIT IS&T) to add questions and comments directly on the paper PDFs. You should highlight the parts of mathematical derivations or text that you do not understand. In addition to those notes on specific parts of the paper, you should also include general, "big picture" comments on the paper's broader implications, promising application ideas, etc. Prior to the class, you should **look through everyone else's uploaded notes** and prepare for the class discussion, where we will be collectively resolving everyone's questions that still remain after the presentation.

- **Final project presentation and write-up** (40%): Your project for this class will typically be methodological, though an applied project involving innovating use of advanced causal inference methodology is also accepted.

  Students are expected to adhere to the following deadlines:

  - February: **Start** thinking about possible topics. Make an appointment with the instructor and meet with him to obtain his reactions.
  - March 11: Turn in a **<u>brief</u> description of your proposed project**. Meet again with the instructor to discuss your proposal. You may be asked to revise and resubmit the proposal in two weeks from the meeting.
  - May 6 (and 13): Students will give **presentations on Zoom** during the regular class time. Students should share screen of their electronic slides to accompany their presentation. Performance on this presentation will be counted toward the class participation grade. Make final revisions to your paper based on the feedback.
  - May 20: **Paper due** in the last class meeting. If you need to request an extension, please contact the instructor.

Course Website

You can find the Canvas website for this course at:

https://canvas.mit.edu/courses/6708

We will distribute course materials on this website. It will also contain pages to upload your notes and other materials.

Auditing Policy

If you are interested in auditing this class, you are welcome, provided that you are willing to *actively participate in the class on a regular basis*. Specifically, I ask you that you make no other commitment that will make you miss the class more than a couple of times and that you will not regularly leave the class early. In addition, just like enrolled students, you will be required to submit annotated paper PDFs two days prior to each class you are attending. Please do not come just to sit and listen. Notetaking does not count as active participation. You are exempted from the class presentations and the final project requirements. However, you are very much welcome to volunteer for one class presentation during the semester.

Topics Covered (Tentative and Incomplete)

1. Machine Learning Methods

   (a) Causal Forest

   - ∗ Wager and Athey (2018, JASA), "Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests."
   - • Wager, Hastie and Efron (2014, JMLR), "Confidence Intervals for Random Forests: The Jackknife and the Infinitesimal Jackknife." (technical foundation for the inference technique)
   - • Athey and Imbens (2016, PNAS), "Recursive Partitioning for Heterogenous Causal Effects." (earlier work on CART for causal inference)
   - • Athey, Tibshirani and Wager (2019, AoS), "Generalized Random Forests." (inferential technique underlying the most recent implementation of CF)

   (b) Bayesian Additive Regression Trees (BART)

   - ∗ Chipman, George and McCulloch (2010, AoAS), "BART: Bayesian Additive Regression Trees."
   - ∗ Hill (2011, JCGS), "Bayesian Nonparametric Modeling for Causal Inference."
   - • Green and Kern (2012, POQ), "Modeling Heterogenous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees."
   - • Hill, Linero and Murray (2020, Annual Reviews), "Bayesian Additive Regression Trees: A Review and Look Forward." (most recent review on BART)
   - • Hahn, Murray and Carvalho (2020, Bayesian Analysis), "Bayesian Regression Tree Models for Causal Inference: Regularization, Confounding, and Heterogeneous Effects."
   - • Ročkoà and Saha (2019, AISTATS Proceedings), "On Theory for BART."

   (c) Generalized Boosted Models (GBM)

   - ∗ McCaffrey, Ridgeway and Morral (2004, Psych Methods), "Propensity Score Estimation with Boosted Regression for Evaluating Causal Effects in Observational Studies."
   - • McCaffrey, Griffin, Almirall, Slaughter, Ramchand and Burgette (2013, Stat in Med), "A Tutorial on Propensity Score Estimation for Multiple Treatments Using Generalized Boosted Models."

2. Robins' G Methods

   (a) G-computation and Marginal Structural Models

   - ∗ Robins, Hernán and Brumback (2000, Epi). "Marginal Structural Models and Causal Inference in Epidemiology."
   - • Robins (1986, Math. Modeling), "A New Approach to Causal Inference..."
   - • Robins (1987), "Addendum to 'A New Approach to Causal Inference..."
   - • Hernán and Robins (2020), *Causal Inference: What If.* Ch. 12–13.

   (b) G-estimation and Structural Nested Models

   - ∗ Robins (2000), "Marginal Structural Models versus Structural Nested Models as Tools for Causal Inference."
   - • Hernán and Robins (2020), *Causal Inference: What If.* Ch. 14.

   (c) Doubly Robust Estimation

   - ∗ Bang and Robins (2005, Biometrics), "Doubly Robust Estimation in Missing Data and Causal Inference Models."
   - ∗ Kang and Schafer (2007, Stat Sci), "Demystifying Double Robustness: A Comparison of Alternative Strategies for Estimating a Population Mean from Incomplete Data."

- Scharfstein, Rotnitzky and Robins (1999, JASA), “Adjusting for Nonignorable Drop-Out Using Semiparametric Nonresponse Models.”
- Glynn and Quinn (2010, PA), “An Introduction to the Augmented Inverse Propensity Weighted Estimator.” (an accessible review article)
- Funk, Westreich, Wiesen, Stürmer, Brookhart and Davidian (2011, Epi), “Doubly Robust Estimation of Causal Effects.” (an accessible tutorial)
- Hernán and Robins (2020), *Causal Inference: What If.* Ch. 21.

3. Ensemble Learning Methods

   (a) Super Learner and Targeted Maximum Likelihood Estimation (TMLE)

   ∗ van der Laan and Rose (2001), *Targeted Learning: Causal Inference for Observational and Experimental data*, Springer. Ch. 3–5.
   - van der Laan and Rose (2001), *Targeted Learning: Causal Inference for Observational and Experimental data*, Appendix A.
   - van der Laan, Polley and Hubbard (2007), “Super Learner,” *Statistical Applications in Genetics and Molecular Biology.*
   - van der Laan and Rubin (2006, IJB), “Targeted Maximum Likelihood Learning.”

   (b) Double Machine Learning

   ∗ Chernozhukov, Chetverikov, Demirer, Duflo, Hansen, Newey and Robins (2018, Econometrics J), “Double/Debiased Machine Learning for Treatment and Structural Parameters.”

   (c) X-learner

   ∗ Künzel, Sekhon, Bickel and Yu (2019, PNAS), “Metalearners for Estimating Heterogeneous Treatment Effects Using Machine Learning.”

4. Heckman’s Framework for Econometric Causal Modeling

   (a) A Preliminary – Control Functions

   ∗ Heckman and Robb (1985, J Econometrics), “Alternative Methods for Evaluating the Impact of Interventions: An Overview.”
   - Heckman and Robb (1986), “Alternative Methods for Solving the Problem of Selection Bias in Evaluating the Impact of Treatments on Outcomes,” in Wainer eds. *Drawing Inferences from Self-Selected Samples,* Springer.
   - Heckman (1979, Econometrica), “Sample Selection Bias as a Specification Error.”
   - Wooldridge (2015, J of Human Resources), “Control Function Methods in Applied Econometrics.”

   (b) Marginal Treatment Effects (MTE)

   ∗ Heckman and Vytlacil (2005, Econometrica), “Structural Equations, Treatment Effects, and Econometric Policy Evaluation.”
   - Heckman and Navarro-Lozano (2004, REStat), “Using Matching, Instrumental Variables, and Control Functions to Estimate Economic Choice Models.”
   - Heckman (2008, Int. Stat. Rev.), “Econometric Causality.”
   - Heckman and Vytlacil (2007, Handbook of Econometrics), “Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation.”
   - Heckman and Vytlacil (2007, Handbook of Econometrics), “Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast Their Effects in New Environments.”

5. Interference, Spillover Effects, and Network Causation

    * Hudgens and Halloran (2008, JASA), "Toward Causal Inference with Interference."
    - Sobel (2006, JASA), "What Do Randomized Studies of Housing Mobility Demonstrate? Causal Inference in the Face of Interference."
    - Rosenbaum (2007, JASA), "Interference between Units in Randomized Experiments."
    - Tchetgen Tchetgen and VanderWeele (2012, Stat Methods Med Res.), "On Causal Inference in the Presence of Interference."
    - Aronow and Samii (2017, AoAS), "Estimating Average Causal Effects under General Interference, with Application to a Social Network Experiment." (generalization to arbitrary known interference)
    - Egami (2020, PA), "Spillover Effects in the Presence of Unobserved Networks." (network-specific causal effects)
    - Sävje, Aronow and Hudgens (2020, AoS), "Average Treatment Effects in the Presence of Unknown Interference." (what can be done with arbitrary and unknown interference)

6. Sensitivity Analysis

    (a) E-value and related alternatives

        * VanderWeele and Ding (2017, AoIM), "Sensitivity Analysis in Observational Research: Introducing the E-value."
        * Ding and VanderWeele (2016, Epi), "Sensitivity Analysis without Assumptions.'
        - Blackwell (2013, PA), "A Selection Bias Approach to Sensitivity Analysis for Causal Effects." (confounding function)
        - Cinelli and Hazlett (2020, JRSS-B), "Making Sense of Sensitivity: Extending Omitted Variables Bias." (robustness value)
        - Zhao (2019, JASA), "On Sensitivity Value of Pair-Matched Observational Studies." (sensitivity value)

    (b) Bayesian sensitivity analysis

        - Gustavson, McCandless, Levy and Richardson (2010, Biometrics), "Simplified Bayesian Sensitivity Analysis for Mismeasured and Unobserved Confounders."

7. Partial Identification

    (a) Frequentist methods
    (b) Bayesian methods

8. Transportability and Extrapolation

9. Mediation

    (a) Fundamentals
    (b) Recent Advances