Syllabus for Causal Inference

Ye Wang

Course description

This is the second/third course in the method sequence of a Political Science department. The aim of the course is to equip students with a current perspective on the progress of causal inference and necessary skills to analyze datasets in social sciences. The course is organized by the structures of data at hand, from the simplest case with only one outcome and one treatment to more complex scenarios with the time dimension involved or multiple outcomes. It covers a series of popular methods that social scientists rely on to draw causal conclusions, ranging from randomized experiments to machine learning techniques. Throughout the course, we will investigate the strength and weakness of each method from a design-based perspective and apply them to analyze real-world examples.

Texts and software

The course will draw a lot from the following textbooks:

1. Angrist and Pischke (2008)
2. Imbens and Rubin (2015)

We will also refer to the following textbooks for several topics:

1. Pearl and Mackenzie (2018)
2. Hansen (2016)
3. Dunning (2012)
4. Aronow and Miller (2019)
5. Friedman, Hastie and Tibshirani (2001)
6. Gerber and Green (2012)

We will working with R in this course, which is an open-source computing language that is very widely used in statistics. You can download it for free from www.r-project.org. We also encourage you to use Rmarkdown for your homework.

Requirements

Students who take the course should have working knowledge of probability theory, matrix algebra, and calculus, as well as some background in writing scripts in R.

Grading of the course is based on class participation (10%), eight assignments (40%), a mid-term exam (20%) and a final exam (30%). You are expected to submit your completed assignment before the deadlines listed on the assignment. Both the mid-term and the final exams will take place in the classroom. Their exact dates will be notified later in the semester.

Course outline

Section I: Data with Y and D

This section first introduces basic concepts in causal inference, such as the Rubin model, the directed acyclic graph (DAG), and the "fundamental problem of causal inference." Then, we turn to simple experiments, which produce datasets with only an outcome Y and a treatment variable D. We will learn the design, estimation, and statistical inference of simple experiments, as well as the general idea of design-based causal inference. We end the section with the discussion of more complicated designs, including patient preference trials and conjoint analysis.

Lecture 1 The fundamental problem of causal inference

The Rubin model
The fundamental problem of causal inference: two solutions
DAG: pros and cons
The tradition of econometrics and the identification of models
Design-based perspective vs. model-based perspective

References: Holland (1986), Freedman (1991), Imbens (2019), Samii (2016), Intro of Pearl and Mackenzie (2018), Ch1 of Angrist and Pischke (2008), Ch1 and Ch2 of Imbens and Rubin (2015)

Lecture 2 The basics of experiment: design and estimation

From sampling to design
Complete experiment vs. Bernoulli trial
Horvitz-Thompson estimator vs. Hajek estimator
The bias-variance tradeoff
Unbiasedness and consistency

References: Ch3, Ch4 and Ch6 of Imbens and Rubin (2015), Hahn (1998), Middleton (2018), Delevoye and Sävje (2020)

Lecture 3 The basics of experiment: inference based on asymptotics

Design-based vs. sampling-based uncertainty
Variance estimation in experiments
Asymptotic normality of the estimators
Finite population CLT
Stein's method

References: Abadie et al. (2014), Aronow et al. (2014), Imbens and Menzel (2018), Li and Ding (2017), Ross et al. (2011)

Lecture 4 The basics of experiment: inference based on resampling

Sharp null and Fisher's randomization test
Bootstrap: basic idea
Bootstrap: different methods
Compare various approaches

References: Ch5 of Imbens and Rubin (2015), Morgan (2017), Young (2019a), Ding et al. (2017), Ch 10 of Hansen (2016)

Lecture 5 More complicated designs

Conjoint analysis: setup and estimation
Patient preference trials
Factorial designs

References: Hainmueller, Hopkins and Yamamoto (2014), Egami and Imai (2018), Bansak et al. (2018), Knox et al. (2019), Li et al. (2020)

Applications

References: Ch8 and Ch9 of Dunning (2012), Gerber, Green and Larimer (2008), Butler and Broockman (2011), Hainmueller and Hopkins (2015), De Benedictis-Kessner et al. (2019)

Section II: Data with Y, D, and $\mathbf{X}$

In spite of the outcome and the treatment, we may also possess auxiliary information in the dataset, the covariates $\mathbf{X}$. Their existence allows us to enhance the efficiency of our estimators or eliminate the estimation bias caused by confounders. This section starts from scenarios where $\mathbf{X}$ are "exogenous," meaning that they are independent to the treatment (but not the outcome). We discuss how to use these covariates to reduce the uncertainty in estimation and generate estimates of heterogeneous treatment effects. We then move to the case where $\mathbf{X}$ are "endogenous" and confounding the causal identification. We review popular tools of de-confounding: regression, weighting, balancing, and matching.

Lecture 6 Regression in causal inference

A quick review
The bias of the regression estimator
Regression adjustment in simple experiments
Bias amplification

References: Samii and Aronow (2012), Ch2 of Angrist and Pischke (2008), Ch7 of Imbens and Rubin (2015), Lin et al. (2013), Aronow and Samii (2016), Middleton et al. (2016)

Lecture 7 From regression to machine learning

Prediction vs. causal inference
Sampling splitting, cross-validation and penalization
Linear models: LASSO, Ridge, SVM
Tree-based models: CART, Boost, Random forest

References: Mullainathan and Spiess (2017), Athey and Imbens (2019), Montgomery and Olivella (2018), Belloni, Chernozhukov and Hansen (2014)

Lecture 8 Heterogeneous treatment effects

Interaction terms
Causal tree and causal forest
Ensemble methods

References: Hainmueller, Mummolo and Xu (2019), Athey and Imbens (2016), Wager and Athey (2018), Athey et al. (2019), Imai and Strauss (2011), Imai, Ratkovic et al. (2013), Grimmer, Messing and Westwood (2017), Künzel et al. (2019), Nie and Wager (2017), Zhou and Xie (2019)

Lecture 9 External validity and the optimal assignment

From SATE to PATE
Classification of external validities
Optimal assignment

References: Stuart et al. (2011), Hartman et al. (2015), Egami and Hartman (2020), Dehejia, Pop-Eleches and Samii (2019), Bisbee et al. (2017), Kitagawa and Tetenov (2018), Athey and Wager (2017), Nie, Brunskill and Wager (2020)

Lecture 10 Control the confounders: weighting and balancing

Rerandomization
IPW: pros and cons
CBPS
Permutation weighting
Entropy balancing
Kernel balancing

References: Li and Ding (2020), Li, Ding and Rubin (2018), Rosenbaum and Rubin (1983), Hirano, Imbens and Ridder (2003), Ma and Wang (2020), Imai and Ratkovic (2014), Arbour and Dimmery (2019), Hainmueller (2012), Hazlett (2018), Harshaw et al. (2019)

Lecture 11 Control the confounders: blocking and matching

From blocking to matching
The asymptotics of matching estimators
Bootstrap and matching
PS matching vs. NN matching
Genetic matching and other varieties

References: Higgins, Sävje and Sekhon (2016), Abadie and Imbens (2006), Abadie and Imbens (2008), Abadie and Imbens (2012), Abadie and Imbens (2011), Abadie and Imbens (2016), Imbens (2015), Otsu and Rai (2017), Diamond and Sekhon (2013), Iacus, King and Porro (2011), King and Nielsen (2019), Iacus, King and Porro (2012), King, Lucas and Nielsen (2017), Roberts, Stewart and Nielsen (2020)

Lecture 12 Synthesis

Doubly robustness
Nuisance parameters and semiparametric estimators
Cluster standard errors

References: *Chernozhukov et al. (2017), Belloni et al. (2017), Chernozhukov et al. (2016), Ratkovic (2019), Abadie et al. (2017), Cameron, Gelbach and Miller (2008), Cameron and Miller (2015), Young (2016)*

Lecture 13 Control the unobservables

Placebo tests
Manski bound
Trimming bound
Sensitivity analysis

References: Ch21 and Ch22 of Imbens and Rubin (2015), Manski (1990), Imbens and Manski (2004), Lee (2009), Imbens (2003), Blackwell (2014), Cinelli and Hazlett (2020)

Applications

References: Fisman and Wei (2004), LaLonde (1986), Sekhon and Titiunik (2012), Lyall (2010), Boyd, Epstein and Martin (2010), Tsai and Xu (2018)

Section III: Data with Y, D, and Z

Treatment assignment Z often differs from the actual treatment exposure D in practice. Based on how Z is assigned, we can identify various types of local effects generated by D. When the difference between the two comes from non-compliance, we say Z is an instrument for D and we are able to estimate the causal effect for a specific group called compliers. When D's value changes over a threshold of Z, we have the regression discontinuity design. Multiple methods are available for us to infer the causal effect of D at Z's threshold. Finally, interference appears when one observation's outcome is affected by another's treatment assignment. Using design-based approaches, we can identify both the direct effect caused by Z and the indirect effect caused by D (relative to Z).

Lecture 14 From non-compliance to instruments

Non-compliance in experiments
The history of IV
2SLS and the Wald estimator
IV in practice

*References: Ch3 of Angrist and Pischke (2008), Ch23 and Ch24 of Imbens and Rubin (2015), Angrist, Imbens and Rubin (1996), Ding, VanderWeele and Robins (2017), Young (2019*b*)*

Lecture 15 Principal strata

Multiple IVs
Principal strata score
IV tests

References: Frangakis and Rubin (2002), Ding and Lu (2016), Aronow and Carnegie (2013), Miratrix et al. (2018), Blackwell (2017), Blackwell and Pashley (2020), Kitagawa (2015), Huber and Mellace (2015), Lee et al. (2020)

Lecture 16 Regression discontinuity design: the classics

Local regression
Bandwidth selection
Bias-correction estimator
Fuzzy RD
Covariates and multiple cutoffs

References: Hahn, Todd and Van der Klaauw (2001), Porter (2003), Lee and Lemieux (2010), McCrary (2008), Imbens and Kalyanaraman (2012), Imbens and Lemieux (2008), Gelman and Imbens (2019), Calonico, Cattaneo and Titiunik (2014), Calonico et al. (2019), Cattaneo et al. (2016), Cattaneo et al. (2020), Cattaneo, Titiunik and Vazquez-Bare (2019)

Lecture 17 Regression discontinuity design: recent development

Local experiment
Discrete running variable
Kink
Donut RD
Geographic RD
RD and convex optimization

References: Sekhon and Titiunik (2017), Cattaneo, Titiunik and Vazquez-Bare (2016), Kolesár and Rothe (2018), Card et al. (2015), Keele and Titiunik (2015), Imbens and Wager (2019), Eckles et al. (2020)

Lecture 18 Interference and diffusion

Partial interference
Exposure mapping
General interference: direct effects
General interference: spillover effects
Contagion

References: Hudgens and Halloran (2008), Tchetgen and VanderWeele (2012), Sinclair, McConnell and Green (2012), Aronow and Samii (2017), Ogburn et al. (2020), Savje, Aronow and Hudgens (2018), Li and Wager (2020), Aronow, Samii and Wang (2020), Wang (2020), Egami (2018), Li et al. (2019)

Applications

References: Angrist, Evans et al. (1998), Clingingsmith, Khwaja and Kremer (2009), Lee (2008), Hall (2015), Dell, Lane and Querubin (2018), Nickerson (2008), Paluck, Shepherd and Aronow (2016)

Section IV: Data with Y, D, and T

We focus on causal inference in panel data, or time-series cross-sectional (TSCS) data, in this section. Time plays a prominent role in these datasets. We observe each unit for multiple times, which provides us with extra information to relax the identification assumption along different directions. Instead of strong/weak ignorability, we often rely on either strict exogeneity or sequential ignorability to draw causal conclusions in panel/TSCS data. We first review classic methods in this field, including fixed effects models and synthetic control. We then proceed to approaches based on flexible models, such as generalized synthetic control and matrix completion. Finally, we introduce some recent methods that take the design-based perspective to understand causal inference with the time dimension.

Lecture 19 Panel data I: classic methods

Classic DID
The two-way fixed effects model (and its bias)
Synthetic control
Sequential ignorability vs. strict exogeneity

References: Ch3 of Hsiao (2014), Bertrand, Duflo and Mullainathan (2004), Abadie, Diamond and Hainmueller (2010), Abadie, Diamond and Hainmueller (2015), Blackwell (2013), Blackwell and Glynn (2018), Imai and Kim (2019), de Chaisemartin and D'Haultfœuille (2020), Strezhnev (2017)

Lecture 20 Panel data II: flexible models

Generalized synthetic control
Matrix completion
SC-DID
Trajectory balancing

References: Xu (2017), Liu, Wang and Xu (2020), Gobillon and Magnac (2016), Athey et al. (2018), Bai and Ng (2019), Hazlett and Xu (2018), Arkhangelsky et al. (2019)

Lecture 21 Panel data III: designed-based methods

Staggered adoption
Augmented synthetic control
Doubly robust panel estimator
Localized factor models

References: Athey and Imbens (2018), Ben-Michael, Feller and Rothstein (2018), Sant'Anna and Zhao (2020), Arkhangelsky and Imbens (2019), Feng (2020), Doudchenko and Imbens (2016), Lu, Nie and Wager (2019)

Applications

References: Tomz, Goldstein and Rivers (2007), Acemoglu et al. (2019), Acemoglu et al. (2016), Sances (2016), Anzia and Berry (2011)

Section V: Data with more than one Y

The last section of the course is dedicated to datasets with multiple outcomes. When these Ys are parallel to each other, hypothesis testing must be conducted with caution due to the well-known problem of multiple comparisons. Sometimes, we are also interested in the causal relationship between the outcome variables. Mediation analysis and structural estimation are two useful tools for this purpose. We conclude this section by examining the connection between empirical results and theories.

Lecture 22 Multiple testing

Test the balance of covariates
The problem of multiple comparisons
Solutions

References: *Hartman and Hidalgo (2018)*, *Casey, Glennerster and Miguel (2012)*, *Anderson (2008)*

Lecture 23 Mediation analysis

Regression with residuals
Mediation analysis from a modern perspective
Multiple mediators

References: *Zhou and Wodtke (2019)*, *Imai, Keele and Tingley (2010)*, *Imai et al. (2011)*, *Acharya, Blackwell and Sen (2016)*, *Zhou and Yamamoto (2020)*

Lecture 24 Structural estimation

The basic idea of structural estimation
GMM
Combine empirics with theories

References: *Hartman and Hidalgo (2018)*, *Acemoglu (2010)*, *Chassang et al. (2012)*, *De Mesquita and Tyson (2020)*, *Slough (2019)*

Applications

References: *Cantoni et al. (2017)*, *Graham and Svolik (2020)*, *Ahlfeldt et al. (2015)*

References

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(490):493–505.

Abadie, Alberto, Alexis Diamond and Jens Hainmueller. 2015. "Comparative politics and the synthetic control method." *American Journal of Political Science* 59(2):495–510.

Abadie, Alberto and Guido W Imbens. 2006. "Large sample properties of matching estimators for average treatment effects." *econometrica* 74(1):235–267.

Abadie, Alberto and Guido W Imbens. 2008. "On the failure of the bootstrap for matching estimators." *Econometrica* 76(6):1537–1557.

Abadie, Alberto and Guido W Imbens. 2011. "Bias-corrected matching estimators for average treatment effects." *Journal of Business & Economic Statistics* 29(1):1–11.

Abadie, Alberto and Guido W Imbens. 2012. "A martingale representation for matching estimators." *Journal of the American Statistical Association* 107(498):833–843.

Abadie, Alberto and Guido W Imbens. 2016. "Matching on the estimated propensity score." *Econometrica* 84(2):781–807.

Abadie, Alberto, Susan Athey, Guido W Imbens and Jeffrey M Wooldridge. 2014. Finite population causal standard errors. Technical report National Bureau of Economic Research.

Abadie, Alberto, Susan Athey, Guido W Imbens and Jeffrey Wooldridge. 2017. When should you adjust standard errors for clustering? Technical report National Bureau of Economic Research.

Acemoglu, Daron. 2010. "Theory, general equilibrium, and political economy in development economics." *Journal of Economic Perspectives* 24(3):17–32.

Acemoglu, Daron, Simon Johnson, Amir Kermani, James Kwak and Todd Mitton. 2016. "The value of connections in turbulent times: Evidence from the United States." *Journal of Financial Economics* 121(2):368–391.

Acemoglu, Daron, Suresh Naidu, Pascual Restrepo and James A Robinson. 2019. "Democracy does cause growth." *Journal of Political Economy* 127(1):47–100.

Acharya, Avidit, Matthew Blackwell and Maya Sen. 2016. "Explaining causal findings without bias: Detecting and assessing direct effects." *American Political Science Review* 110(3):512–529.

Ahlfeldt, Gabriel M, Stephen J Redding, Daniel M Sturm and Nikolaus Wolf. 2015. "The economics of density: Evidence from the Berlin Wall." *Econometrica* 83(6):2127–2189.

Anderson, Michael L. 2008. "Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects." *Journal of the American statistical Association* 103(484):1481–1495.

Angrist, Joshua D, Guido W Imbens and Donald B Rubin. 1996. "Identification of causal effects using instrumental variables." *Journal of the American statistical Association* 91(434):444–455.

Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion.* Princeton university press.

Angrist, Joshua D, William N Evans et al. 1998. "Children and Their Parents' Labor Supply: Evidence from Exogenous Variation in Family Size." *American Economic Review* 88(3):450–477.

Anzia, Sarah F and Christopher R Berry. 2011. "The Jackie (and Jill) Robinson effect: why do congresswomen outperform congressmen?" *American Journal of Political Science* 55(3):478–493.

Arbour, David and Drew Dimmery. 2019. "Permutation Weighting." *arXiv preprint arXiv:1901.01230* .

Arkhangelsky, Dmitry and Guido W Imbens. 2019. "Double-robust identification for causal panel data models." *arXiv preprint arXiv:1909.09412* .

Arkhangelsky, Dmitry, Susan Athey, David A Hirshberg, Guido W Imbens and Stefan Wager. 2019. Synthetic difference in differences. Technical report National Bureau of Economic Research.

Aronow, Peter M and Allison Carnegie. 2013. "Beyond LATE: Estimation of the average treatment effect with an instrumental variable." *Political Analysis* pp. 492–506.

Aronow, Peter M and Benjamin T Miller. 2019. *Foundations of agnostic statistics.* Cambridge University Press.

Aronow, Peter M and Cyrus Samii. 2016. "Does regression produce representative estimates of causal effects?" *American Journal of Political Science* 60(1):250–267.

Aronow, Peter M and Cyrus Samii. 2017. "Estimating average causal effects under general interference, with application to a social network experiment." *The Annals of Applied Statistics* 11(4):1912–1947.

Aronow, Peter M, Cyrus Samii and Ye Wang. 2020. "Design-based inference for spatial experiments with interference." *arXiv preprint arXiv:2010.13599* .

Aronow, Peter M, Donald P Green, Donald KK Lee et al. 2014. "Sharp bounds on the variance in randomized experiments." *The Annals of Statistics* 42(3):850–871.

Athey, Susan and Guido Imbens. 2016. "Recursive partitioning for heterogeneous causal effects." *Proceedings of the National Academy of Sciences* 113(27):7353–7360.

Athey, Susan and Guido W Imbens. 2018. Design-based analysis in difference-in-differences settings with staggered adoption. Technical report National Bureau of Economic Research.

Athey, Susan and Guido W Imbens. 2019. "Machine learning methods that economists should know about." *Annual Review of Economics* 11:685–725.

Athey, Susan, Julie Tibshirani, Stefan Wager et al. 2019. "Generalized random forests." *The Annals of Statistics* 47(2):1148–1178.

Athey, Susan, Mohsen Bayati, Nikolay Doudchenko, Guido Imbens and Khashayar Khosravi. 2018. Matrix completion methods for causal panel data models. Technical report National Bureau of Economic Research.

Athey, Susan and Stefan Wager. 2017. "Efficient policy learning." *arXiv preprint arXiv:1702.02896* .

Bai, Jushan and Serena Ng. 2019. "Matrix completion, counterfactuals, and factor analysis of missing data." *arXiv preprint arXiv:1910.06677* .

Bansak, Kirk, Jens Hainmueller, Daniel J Hopkins and Teppei Yamamoto. 2018. "The number of choice tasks and survey satisficing in conjoint experiments." *Political Analysis* 26(1):112–119.

Belloni, Alexandre, Victor Chernozhukov and Christian Hansen. 2014. "High-dimensional methods and inference on structural and treatment effects." *Journal of Economic Perspectives* 28(2):29–50.

Belloni, Alexandre, Victor Chernozhukov, Iván Fernández-Val and Christian Hansen. 2017. "Program evaluation and causal inference with high-dimensional data." *Econometrica* 85(1):233–298.

Ben-Michael, Eli, Avi Feller and Jesse Rothstein. 2018. "The augmented synthetic control method." *arXiv preprint arXiv:1811.04170* .

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

Bisbee, James, Rajeev Dehejia, Cristian Pop-Eleches and Cyrus Samii. 2017. "Local instruments, global extrapolation: External validity of the labor supply–fertility local average treatment effect." *Journal of Labor Economics* 35(S1):S99–S147.

Blackwell, Matthew. 2013. "A framework for dynamic causal inference in political science." *American Journal of Political Science* 57(2):504–520.

Blackwell, Matthew. 2014. "A selection bias approach to sensitivity analysis for causal effects." *Political Analysis* 22(2):169–182.

Blackwell, Matthew. 2017. "Instrumental variable methods for conditional effects and causal interaction in voter mobilization experiments." *Journal of the American Statistical Association* 112(518):590–599.

Blackwell, Matthew and Adam N Glynn. 2018. "How to make causal inferences with time-series cross-sectional data under selection on observables." *American Political Science Review* 112(4):1067–1082.

Blackwell, Matthew and Nicole E Pashley. 2020. "Noncompliance and instrumental variables for 2 factorial experiments.".

Boyd, Christina L, Lee Epstein and Andrew D Martin. 2010. "Untangling the causal effects of sex on judging." *American journal of political science* 54(2):389–411.

Butler, Daniel M and David E Broockman. 2011. "Do politicians racially discriminate against constituents? A field experiment on state legislators." *American Journal of Political Science* 55(3):463–477.

Calonico, Sebastian, Matias D Cattaneo, Max H Farrell and Rocio Titiunik. 2019. "Regression discontinuity designs using covariates." *Review of Economics and Statistics* 101(3):442–451.

Calonico, Sebastian, Matias D Cattaneo and Rocio Titiunik. 2014. "Robust nonparametric confidence intervals for regression-discontinuity designs." *Econometrica* 82(6):2295–2326.

Cameron, A Colin and Douglas L Miller. 2015. "A practitioner's guide to cluster-robust inference." *Journal of human resources* 50(2):317–372.

Cameron, A Colin, Jonah B Gelbach and Douglas L Miller. 2008. "Bootstrap-based improvements for inference with clustered errors." *The Review of Economics and Statistics* 90(3):414–427.

Cantoni, Davide, Yuyu Chen, David Y Yang, Noam Yuchtman and Y Jane Zhang. 2017. "Curriculum and ideology." *Journal of Political Economy* 125(2):338–392.

Card, David, David S Lee, Zhuan Pei and Andrea Weber. 2015. "Inference on causal effects in a generalized regression kink design." *Econometrica* 83(6):2453–2483.

Casey, Katherine, Rachel Glennerster and Edward Miguel. 2012. "Reshaping institutions: Evidence on aid impacts using a preanalysis plan." *The Quarterly Journal of Economics* 127(4):1755–1812.

Cattaneo, Matias D, Luke Keele, Rocío Titiunik and Gonzalo Vazquez-Bare. 2020. "Extrapolating treatment effects in multi-cutoff regression discontinuity designs." *Journal of the American Statistical Association* pp. 1–12.

Cattaneo, Matias D, Rocio Titiunik and Gonzalo Vazquez-Bare. 2016. "Inference in regression discontinuity designs under local randomization." *The Stata Journal* 16(2):331–367.

Cattaneo, Matias D, Rocio Titiunik and Gonzalo Vazquez-Bare. 2019. "The Regression Discontinuity Design." *arXiv preprint arXiv:1906.04242* .

Cattaneo, Matias D, Rocío Titiunik, Gonzalo Vazquez-Bare and Luke Keele. 2016. "Interpreting regression discontinuity designs with multiple cutoffs." *The Journal of Politics* 78(4):1229–1248.

Chassang, Sylvain, Padró I Miquel, Erik Snowberg et al. 2012. "Selective trials: A principal-agent approach to randomized controlled experiments." *American Economic Review* 102(4):1279–1309.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen and Whitney Newey. 2017. "Double/debiased/neyman machine learning of treatment effects." *American Economic Review* 107(5):261–65.

Chernozhukov, Victor, Juan Carlos Escanciano, Hidehiko Ichimura, Whitney K Newey and James M Robins. 2016. "Locally robust semiparametric estimation." *arXiv preprint arXiv:1608.00033* .

Cinelli, Carlos and Chad Hazlett. 2020. "Making sense of sensitivity: Extending omitted variable bias." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 82(1):39–67.

Clingingsmith, David, Asim Ijaz Khwaja and Michael Kremer. 2009. "Estimating the impact of the Hajj: religion and tolerance in Islam's global gathering." *The Quarterly Journal of Economics* 124(3):1133–1170.

De Benedictis-Kessner, Justin, Matthew A Baum, Adam J Berinsky and Teppei Yamamoto. 2019. "Persuading the enemy: Estimating the persuasive effects of partisan media with the preference-incorporating choice and assignment design." *American Political Science Review* 113(4):902–916.

de Chaisemartin, Clément and Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* .

De Mesquita, Ethan Bueno and Scott A Tyson. 2020. "The commensurability problem: Conceptual difficulties in estimating the effect of behavior on behavior." *American Political Science Review* 114(2):375–391.

Dehejia, Rajeev, Cristian Pop-Eleches and Cyrus Samii. 2019. "From local to global: External validity in a fertility natural experiment." *Journal of Business & Economic Statistics* pp. 1–27.

Delevoye, Angèle and Fredrik Sävje. 2020. "Consistency of the Horvitz–Thompson estimator under general sampling and experimental designs." *Journal of Statistical Planning and Inference* 207:190–197.

Dell, Melissa, Nathan Lane and Pablo Querubin. 2018. "The historical state, local collective action, and economic development in Vietnam." *Econometrica* 86(6):2083–2121.

Diamond, Alexis and Jasjeet S Sekhon. 2013. "Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies." *Review of Economics and Statistics* 95(3):932–945.

Ding, Peng and Jiannan Lu. 2016. "Principal stratification analysis using principal scores." *arXiv preprint arXiv:1602.01196* .

Ding, Peng, TJ VanderWeele and James M Robins. 2017. "Instrumental variables as bias amplifiers with general outcome and confounding." *Biometrika* 104(2):291–302.

Ding, Peng et al. 2017. "A paradox from randomization-based causal inference." *Statistical science* 32(3):331–345.

Doudchenko, Nikolay and Guido W Imbens. 2016. Balancing, regression, difference-in-differences and synthetic control methods: A synthesis. Technical report National Bureau of Economic Research.

Dunning, Thad. 2012. *Natural experiments in the social sciences: a design-based approach.* Cambridge University Press.

Eckles, Dean, Nikolaos Ignatiadis, Stefan Wager and Han Wu. 2020. "Noise-Induced Randomization in Regression Discontinuity Designs." *arXiv preprint arXiv:2004.09458* .

Egami, Naoki. 2018. "Identification of Causal Diffusion Effects Using Stationary Causal Directed Acyclic Graphs." *arXiv preprint arXiv:1810.07858* .

Egami, Naoki and Erin Hartman. 2020. "Elements of External Validity: Framework, Design, and Analysis.".

Egami, Naoki and Kosuke Imai. 2018. "Causal interaction in factorial experiments: Application to conjoint analysis." *Journal of the American Statistical Association* .

Feng, Yingjie. 2020. "Causal Inference in Possibly Nonlinear Factor Models." *arXiv preprint arXiv:2008.13651* .

Fisman, Raymond and Shang-Jin Wei. 2004. "Tax rates and tax evasion: evidence from "missing imports" in China." *Journal of political Economy* 112(2):471–496.

Frangakis, Constantine E and Donald B Rubin. 2002. "Principal stratification in causal inference." *Biometrics* 58(1):21–29.

Freedman, David A. 1991. "Statistical models and shoe leather." *Sociological methodology* pp. 291–313.

Friedman, Jerome, Trevor Hastie and Robert Tibshirani. 2001. *The elements of statistical learning.* Vol. 1 Springer series in statistics New York.

Gelman, Andrew and Guido Imbens. 2019. "Why high-order polynomials should not be used in regression discontinuity designs." *Journal of Business & Economic Statistics* 37(3):447–456.

Gerber, Alan S and Donald P Green. 2012. *Field experiments: Design, analysis, and interpretation.* WW Norton.

Gerber, Alan S, Donald P Green and Christopher W Larimer. 2008. "Social pressure and voter turnout: Evidence from a large-scale field experiment." *American political Science review* pp. 33–48.

Gobillon, Laurent and Thierry Magnac. 2016. "Regional policy evaluation: Interactive fixed effects and synthetic controls." *Review of Economics and Statistics* 98(3):535–551.

Graham, Matthew H and Milan W Svolik. 2020. "Democracy in America? Partisanship, Polarization, and the Robustness of Support for Democracy in the United States." *American Political Science Review* 114(2):392–409.

Grimmer, Justin, Solomon Messing and Sean J Westwood. 2017. "Estimating heterogeneous treatment effects and the effects of heterogeneous treatments with ensemble methods." *Political Analysis* 25(4):413–434.

Hahn, Jinyong. 1998. "On the role of the propensity score in efficient semiparametric estimation of average treatment effects." *Econometrica* pp. 315–331.

Hahn, Jinyong, Petra Todd and Wilbert Van der Klaauw. 2001. "Identification and estimation of treatment effects with a regression-discontinuity design." *Econometrica* 69(1):201–209.

Hainmueller, Jens. 2012. "Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies." *Political analysis* pp. 25–46.

Hainmueller, Jens and Daniel J Hopkins. 2015. "The hidden American immigration consensus: A conjoint analysis of attitudes toward immigrants." *American Journal of Political Science* 59(3):529–548.

Hainmueller, Jens, Daniel J Hopkins and Teppei Yamamoto. 2014. "Causal inference in conjoint analysis: Understanding multidimensional choices via stated preference experiments." *Political analysis* 22(1):1–30.

Hainmueller, Jens, Jonathan Mummolo and Yiqing Xu. 2019. "How much should we trust estimates from multiplicative interaction models? Simple tools to improve empirical practice." *Political Analysis* 27(2):163–192.

Hall, Andrew B. 2015. "What happens when extremists win primaries?" *American Political Science Review* 109(1):18–42.

Hansen, Bruce. 2016. "Econometrics." *A textbook draft available online at www. ssc. wisc. edu/˜ bhansen/econometrics/Econometrics. pdf* .

Harshaw, Christopher, Fredrik Sävje, Daniel Spielman and Peng Zhang. 2019. "Balancing covariates in randomized experiments using the Gram-Schmidt walk." *arXiv preprint arXiv:1911.03071* .

Hartman, Erin and F Daniel Hidalgo. 2018. "An equivalence approach to balance and placebo tests." *American Journal of Political Science* 62(4):1000–1013.

Hartman, Erin, Richard Grieve, Roland Ramsahai and Jasjeet S. Sekhon. 2015. "From sample average treatment effect to population average treatment effect on the treated: combining experimental with observational studies to estimate population treatment effects." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 178(3):757–778.

Hazlett, Chad. 2018. "Kernel Balancing: A flexible non-parametric weighting procedure for estimating causal effects." *Available at SSRN 2746753* .

Hazlett, Chad and Yiqing Xu. 2018. "Trajectory balancing: A general reweighting approach to causal inference with time-series cross-sectional data." *Available at SSRN 3214231* .

Higgins, Michael J, Fredrik Sävje and Jasjeet S Sekhon. 2016. "Improving massive experiments with threshold blocking." *Proceedings of the National Academy of Sciences* 113(27):7369–7376.

Hirano, Keisuke, Guido W Imbens and Geert Ridder. 2003. "Efficient estimation of average treatment effects using the estimated propensity score." *Econometrica* 71(4):1161–1189.

Holland, Paul W. 1986. "Statistics and causal inference." *Journal of the American statistical Association* 81(396):945–960.

Hsiao, Cheng. 2014. *Analysis of panel data*. Number 54 Cambridge university press.

Huber, Martin and Giovanni Mellace. 2015. “Testing instrument validity for LATE identification based on inequality moment constraints.” *Review of Economics and Statistics* 97(2):398–411.

Hudgens, Michael G and M Elizabeth Halloran. 2008. “Toward causal inference with interference.” *Journal of the American Statistical Association* 103(482):832–842.

Iacus, Stefano M, Gary King and Giuseppe Porro. 2011. “Multivariate matching methods that are monotonic imbalance bounding.” *Journal of the American Statistical Association* 106(493):345–361.

Iacus, Stefano M, Gary King and Giuseppe Porro. 2012. “Causal inference without balance checking: Coarsened exact matching.” *Political analysis* pp. 1–24.

Imai, Kosuke and Aaron Strauss. 2011. “Estimation of heterogeneous treatment effects from randomized experiments, with application to the optimal planning of the get-out-the-vote campaign.” *Political Analysis* 19(1):1–19.

Imai, Kosuke and In Song Kim. 2019. “When should we use unit fixed effects regression models for causal inference with longitudinal data?” *American Journal of Political Science* 63(2):467–490.

Imai, Kosuke, Luke Keele and Dustin Tingley. 2010. “A general approach to causal mediation analysis.” *Psychological methods* 15(4):309.

Imai, Kosuke, Luke Keele, Dustin Tingley and Teppei Yamamoto. 2011. “Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies.” *American Political Science Review* pp. 765–789.

Imai, Kosuke and Marc Ratkovic. 2014. “Covariate balancing propensity score.” *Journal of the Royal Statistical Society: Series B: Statistical Methodology* pp. 243–263.

Imai, Kosuke, Marc Ratkovic et al. 2013. “Estimating treatment effect heterogeneity in randomized program evaluation.” *The Annals of Applied Statistics* 7(1):443–470.

Imbens, Guido. 2019. Potential outcome and directed acyclic graph approaches to causality: Relevance for empirical practice in economics. Technical report National Bureau of Economic Research.

Imbens, Guido and Karthik Kalyanaraman. 2012. “Optimal bandwidth choice for the regression discontinuity estimator.” *The Review of economic studies* 79(3):933–959.

Imbens, Guido and Konrad Menzel. 2018. A causal bootstrap. Technical report National Bureau of Economic Research.

Imbens, Guido and Stefan Wager. 2019. “Optimized regression discontinuity designs.” *Review of Economics and Statistics* 101(2):264–278.

Imbens, Guido W. 2003. “Sensitivity to exogeneity assumptions in program evaluation.” *American Economic Review* 93(2):126–132.

Imbens, Guido W. 2015. “Matching methods in practice: Three examples.” *Journal of Human Resources* 50(2):373–419.

Imbens, Guido W and Charles F Manski. 2004. “Confidence intervals for partially identified parameters.” *Econometrica* 72(6):1845–1857.

Imbens, Guido W and Donald B Rubin. 2015. *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Imbens, Guido W and Thomas Lemieux. 2008. “Regression discontinuity designs: A guide to practice.” *Journal of econometrics* 142(2):615–635.

Keele, Luke J and Rocio Titiunik. 2015. "Geographic boundaries as regression discontinuities." *Political Analysis* 23(1):127–155.

King, Gary, Christopher Lucas and Richard A Nielsen. 2017. "The balance-sample size frontier in matching methods for causal inference." *American Journal of Political Science* 61(2):473–489.

King, Gary and Richard Alexander Nielsen. 2019. "Why propensity scores should not be used for matching.".

Kitagawa, Toru. 2015. "A test for instrument validity." *Econometrica* 83(5):2043–2063.

Kitagawa, Toru and Aleksey Tetenov. 2018. "Who should be treated? empirical welfare maximization methods for treatment choice." *Econometrica* 86(2):591–616.

Knox, Dean, Teppei Yamamoto, Matthew A Baum and Adam J Berinsky. 2019. "Design, identification, and sensitivity analysis for patient preference trials." *Journal of the American Statistical Association* 114(528):1532–1546.

Kolesár, Michal and Christoph Rothe. 2018. "Inference in regression discontinuity designs with a discrete running variable." *American Economic Review* 108(8):2277–2304.

Künzel, Sören R, Jasjeet S Sekhon, Peter J Bickel and Bin Yu. 2019. "Metalearners for estimating heterogeneous treatment effects using machine learning." *Proceedings of the national academy of sciences* 116(10):4156–4165.

LaLonde, Robert J. 1986. "Evaluating the econometric evaluations of training programs with experimental data." *The American economic review* pp. 604–620.

Lee, David L, Justin McCrary, Marcelo J Moreira and Jack Porter. 2020. "Valid t-ratio Inference for IV." *arXiv preprint arXiv:2010.05058* .

Lee, David S. 2008. "Randomized experiments from non-random selection in US House elections." *Journal of Econometrics* 142(2):675–697.

Lee, David S. 2009. "Training, wages, and sample selection: Estimating sharp bounds on treatment effects." *The Review of Economic Studies* 76(3):1071–1102.

Lee, David S and Thomas Lemieux. 2010. "Regression discontinuity designs in economics." *Journal of economic literature* 48(2):281–355.

Li, Shuangning and Stefan Wager. 2020. "Random Graph Asymptotics for Treatment Effect Estimation under Network Interference." *arXiv preprint arXiv:2007.13302* .

Li, Xinran and Peng Ding. 2017. "General forms of finite population central limit theorems with applications to causal inference." *Journal of the American Statistical Association* 112(520):1759–1769.

Li, Xinran and Peng Ding. 2020. "Rerandomization and regression adjustment." *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* .

Li, Xinran, Peng Ding and Donald B Rubin. 2018. "Asymptotic theory of rerandomization in treatment–control experiments." *Proceedings of the National Academy of Sciences* 115(37):9157–9162.

Li, Xinran, Peng Ding, Donald B Rubin et al. 2020. "Rerandomization in 2^K factorial experiments." *The Annals of Statistics* 48(1):43–63.

Li, Xinran, Peng Ding, Qian Lin, Dawei Yang and Jun S Liu. 2019. "Randomization inference for peer effects." *Journal of the American Statistical Association* 114(528):1651–1664.

Lin, Winston et al. 2013. "Agnostic notes on regression adjustments to experimental data: Reexamining Freedman's critique." *The Annals of Applied Statistics* 7(1):295–318.

Liu, Licheng, Ye Wang and Yiqing Xu. 2020. "A practical guide to counterfactual estimators for causal inference with time-series cross-sectional data." *Available at SSRN 3555463* .

Lu, Chen, Xinkun Nie and Stefan Wager. 2019. "Robust nonparametric difference-in-differences estimation." *arXiv preprint arXiv:1905.11622* .

Lyall, Jason. 2010. "Are coethnics more effective counterinsurgents? Evidence from the second Chechen war." *American Political Science Review* pp. 1–20.

Ma, Xinwei and Jingshen Wang. 2020. "Robust inference using inverse probability weighting." *Journal of the American Statistical Association* pp. 1–10.

Manski, Charles F. 1990. "Nonparametric bounds on treatment effects." *The American Economic Review* 80(2):319–323.

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of econometrics* 142(2):698–714.

Middleton, Joel A. 2018. "A unified theory of regression adjustment for design-based inference." *arXiv preprint arXiv:1803.06011* .

Middleton, Joel A, Marc A Scott, Ronli Diakow and Jennifer L Hill. 2016. "Bias amplification and bias unmasking." *Political Analysis* 24(3):307–323.

Miratrix, Luke, Jane Furey, Avi Feller, Todd Grindal and Lindsay C Page. 2018. "Bounding, an accessible method for estimating principal causal effects, examined and explained." *Journal of Research on Educational Effectiveness* 11(1):133–162.

Montgomery, Jacob M and Santiago Olivella. 2018. "Tree-Based Models for Political Science Data." *American Journal of Political Science* 62(3):729–744.

Morgan, Kari Lock. 2017. "Reallocating and Resampling: A Comparison for Inference." *arXiv preprint arXiv:1708.02102* .

Mullainathan, Sendhil and Jann Spiess. 2017. "Machine learning: an applied econometric approach." *Journal of Economic Perspectives* 31(2):87–106.

Nickerson, David W. 2008. "Is voting contagious? Evidence from two field experiments." *American political Science review* pp. 49–57.

Nie, Xinkun, Emma Brunskill and Stefan Wager. 2020. "Learning when-to-treat policies." *Journal of the American Statistical Association* (just-accepted):1–58.

Nie, Xinkun and Stefan Wager. 2017. "Quasi-oracle estimation of heterogeneous treatment effects." *arXiv preprint arXiv:1712.04912* .

Ogburn, Elizabeth L, Oleg Sofrygin, Ivan Diaz and Mark J van der Laan. 2020. "Causal inference for social network data." *arXiv preprint arXiv:1705.08527* .

Otsu, Taisuke and Yoshiyasu Rai. 2017. "Bootstrap inference of matching estimators for average treatment effects." *Journal of the American Statistical Association* 112(520):1720–1732.

Paluck, Elizabeth Levy, Hana Shepherd and Peter M. Aronow. 2016. "Changing climates of conflict: A social network experiment in 56 schools." *Proceedings of the National Academy of Science* 113(3):566–571.

Pearl, Judea and Dana Mackenzie. 2018. *The book of why: the new science of cause and effect*. Basic Books.

Porter, Jack. 2003. "Estimation in the regression discontinuity model." *Unpublished Manuscript, Department of Economics, University of Wisconsin at Madison* 2003:5–19.

Ratkovic, Marc. 2019. "Rehabilitating the Regression: Honest and Valid Causal Inference through Machine Learning.".

Roberts, Margaret E, Brandon M Stewart and Richard A Nielsen. 2020. "Adjusting for confounding with text matching." *American Journal of Political Science* 64(4):887–903.

Rosenbaum, Paul R and Donald B Rubin. 1983. "The central role of the propensity score in observational studies for causal effects." *Biometrika* 70(1):41–55.

Ross, Nathan et al. 2011. "Fundamentals of Stein's method." *Probability Surveys* 8:210–293.

Samii, Cyrus. 2016. "Causal empiricism in quantitative research." *The Journal of Politics* 78(3):941–955.

Samii, Cyrus and Peter M Aronow. 2012. "On equivalencies between design-based and regression-based variance estimators for randomized experiments." *Statistics & Probability Letters* 82(2):365–370.

Sances, Michael W. 2016. "The distributional impact of greater responsiveness: Evidence from New York towns." *The Journal of Politics* 78(1):105–119.

Sant'Anna, Pedro HC and Jun Zhao. 2020. "Doubly robust difference-in-differences estimators." *Journal of Econometrics* 219(1):101–122.

Savje, Fredrik, Peter M. Aronow and Michael G. Hudgens. 2018. "Average treatment effects in the presence of unknown interference." *arXiv:1711.06399 [math.ST]* .

Sekhon, Jasjeet S and Rocio Titiunik. 2012. "When natural experiments are neither natural nor experiments." *American Political Science Review* pp. 35–57.

Sekhon, Jasjeet S and Rocío Titiunik. 2017. "On Interpreting the Regression Discontinuity Design as a Local Experiment', Regression Discontinuity Designs (Advances in Econometrics, Volume 38).".

Sinclair, Betsy, Margaret McConnell and Donald P Green. 2012. "Detecting spillover effects: Design and analysis of multilevel experiments." *American Journal of Political Science* 56(4):1055–1069.

Slough, Tara. 2019. "On Theory and Identification: When and Why We Need Theory for Causal Identification.".

Strezhnev, Anton. 2017. "Generalized Difference-in-Differences Estimands and Synthetic Controls." *Unpublished manuscript* .

Stuart, Elizabeth A, Stephen R Cole, Catherine P Bradshaw and Philip J Leaf. 2011. "The use of propensity scores to assess the generalizability of results from randomized trials." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 174(2):369–386.

Tchetgen, Eric J Tchetgen and Tyler J VanderWeele. 2012. "On causal inference in the presence of interference." *Statistical methods in medical research* 21(1):55–75.

Tomz, Michael, Judith L Goldstein and Douglas Rivers. 2007. "Do we really know that the WTO increases trade? Comment." *American Economic Review* 97(5):2005–2018.

Tsai, Lily L and Yiqing Xu. 2018. "Outspoken insiders: Political connections and citizen participation in authoritarian China." *Political Behavior* 40(3):629–657.

Wager, Stefan and Susan Athey. 2018. "Estimation and inference of heterogeneous treatment effects using random forests." *Journal of the American Statistical Association* 113(523):1228–1242.

Wang, Ye. 2020. "Causal Inference under Temporal and Spatial Interference.".

Xu, Yiqing. 2017. "Generalized synthetic control method: Causal inference with interactive fixed effects models." *Political Analysis* 25(1):57–76.

Young, Alwyn. 2016. "Improved, nearly exact, statistical inference with robust and clustered covariance matrices using effective degrees of freedom corrections." *Manuscript, London School of Economics* .

Young, Alwyn. 2019*a*. "Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results." *The Quarterly Journal of Economics* 134(2):557–598.

Young, Alwyn. 2019*b*. "Consistency without inference: Instrumental variables in practical application.".

Zhou, Xiang and Geoffrey T Wodtke. 2019. "A regression-with-residuals method for estimating controlled direct effects." *Political Analysis* 27(3):360–369.

Zhou, Xiang and Teppei Yamamoto. 2020. "Tracing Causal Paths from Experimental and Observational Data.".

Zhou, Xiang and Yu Xie. 2019. "Marginal treatment effects from a propensity score perspective." *Journal of Political Economy* 127(6):3070–3084.