15.838 Experimental Design & Analysis
Research Seminar in Marketing

MIT Sloan School of Management **Spring 2019** **Professor Dean Eckles**

Wednesdays 10am – 1pm, E51-057

Randomized experiments are important tools for testing theories, exploring design spaces, and evaluating or finding policies. It is increasingly possible to conduct field experiments in many empirical contexts. We will approach learning how to design and analyze experiments through a mix of methodological learning (through readings, lectures, a quiz, an assignment) and critical readings of important experimental work.

Topics include: potential outcomes, Fisherian randomization inference, Fisherian and Neymanian null hypotheses, optimal design of experiments, using covariates to increase precision, planning experiments, heterogeneous treatment effects, evaluating and learning policies.

This year there will be particular emphasis on (a) randomization inference and (b) evaluation and optimization of policies.

This course is open to graduate students from across MIT with some prior statistics, machine learning, or econometrics training. This course is required for Sloan PhD students in Marketing. This year it fulfills the core requirement for the IDSS SES PhD program for social science research design, substituting for 21A.809.

Textbooks

Required:
Gerber, A., & Green, D. (2012). *Field Experiments: Design, Analysis, and Interpretation*.
Imbens, G. W., & Rubin, D. B. (2015). *Causal Inference in Statistics, Social, and Biomedical Sciences*.

We use excerpts from these books, which are also recommended in general:
Shadish, W.R., Cook, T.D., & Campbell, D.T. (2001). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*.
Owen. A. B. (draft). *Monte Carlo Theory, Methods and Examples*. http://statweb.stanford.edu/~owen/mc/
Manski, C. F. (2007). *Identification for Prediction and Decision*.
Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction*.

Schedule with assigned reading

Feb 13: Overview of class, potential outcomes, Fisherian and Neymanian inference
Gerber & Green, Ch. 2 - 3.
Imbens & Rubin, Ch. 6

Optional:
Imbens & Rubin, Ch. 5
Samii, C., & Aronow, P. M. (2012). On equivalencies between design-based and regression-based variance estimators for randomized experiments. *Statistics & Probability Letters*, *82*(2), 365-370.

Abadie, A., Athey, S., Imbens, G. W., & Wooldridge, J. M. (2014). Finite population causal standard errors. NBER Working Paper 20325.

Feb 20: Fisherian randomization inference II; using covariates (blocking and adjustment)

Wu, J., & Ding, P. (2018). Randomization tests for weak null hypotheses. arXiv preprint arXiv:1809.07419.

Peters, J., Bühlmann, P., & Meinshausen, N. (2016). Causal inference by using invariant prediction: identification and confidence intervals. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 78(5), 947-1012.

Gerber & Green, Ch. 4.

Optional:

Imbens, G. W., & Rosenbaum, P. R. (2005). Robust, accurate confidence intervals with a weak instrument: Quarter of birth and education. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 168(1), 109-126.

Chung, E., & Romano, J. P. (2013). Exact and asymptotically robust permutation tests. *The Annals of Statistics, 41*(2), 484-507.

Miratrix, L. W., Sekhon, J. S., & Yu, B. (2013). Adjusting treatment effect estimates by post-stratification in randomized experiments. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *75*(2), 369-396.

Aronow, P. M., & Middleton, J. A. (2013). A class of unbiased estimators of the average treatment effect in randomized experiments. *Journal of Causal Inference*, *1*(1), 135-154.

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

Higgins, M. J., Sävje, F., & Sekhon, J. S. (2016). Improving massive experiments with threshold blocking. *PNAS*.

Feb 27: Quiz; Policy evaluation basic

In-class quiz on Fisherian randomization inference

Review readings for quiz.

Mar 6: Policy evaluation and optimization; bandits

Manski, ch. 10.

Athey, S., & Wager, S. (2017). Efficient policy learning. https://arxiv.org/abs/1702.02896

Optional:

Simester, D., Timoshenko, A., & Zoumpoulis, S. I. (2017). Efficiently evaluating targeting policies using field experiments. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3017384

Agarwal, A. (2017). Off-policy evaluation and learning. Lecture notes. http://alekhagarwal.net/bandits_and_rl/off_policy.pdf

Mar 13: Stochastic policies; dynamic treatment regimes

Owen, ch. 9.

Chin, A., Eckles, D., & Ugander, J. (2018). Evaluating stochastic seeding strategies in networks. arXiv preprint arXiv:1809.09561.

Optional:

Kennedy, E. H. (2018). Nonparametric causal effects based on incremental propensity score interventions. *Journal of the American Statistical Association*. Forthcoming.

Chakraborty, B., & Moodie, E. E. (2013). *Statistical Methods for Dynamic Treatment Regimes*. New York: Springer.

Mar 20: Learning about mechanisms; post-treatment variables
Gerber & Green, Ch. 10

Optional:
Imai, K., Tingley, D., & Yamamoto, T. (2013). Experimental designs for identifying causal mechanisms. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 176(1), 5-51.
Page, L. C., Feller, A., Grindal, T., Miratrix, L., & Somers, M. A. (2015). Principal stratification: A tool for understanding variation in program effects across endogenous subgroups. *American Journal of Evaluation*, 36(4), 514-531.
Pearl, J. (2011). Principal stratification—A goal or a tool?. *The International Journal of Biostatistics, 7*(1), 1-13.
And associated discussion in the same issue: https://www.degruyter.com/view/j/ijb.2011.7.issue-1/issue-files/ijb.2011.7.issue-1.xml

Not covered this year: Combining multiple (experimental and observational) data sets
Rosenman, E., Owen, A. B., & Baiocchi, M. (2018). Propensity score methods for merging observational and experimental datasets. arXiv preprint arXiv:1804.07863.
Künzel, S. R., Stadie, B. C., Vemuri, N., Ramakrishnan, V., Sekhon, J. S., & Abbeel, P. (2018). Transfer learning for estimating causal effects using neural networks. arXiv preprint arXiv:1808.07804.
Chen, A., Owen, A. B., & Shi, M. (2015). Data enriched linear regression. *Electronic Journal of Statistics*, *9*(1), 1078-1112.
Hartman, E., Grieve, R., Ramsahai, R., & Sekhon, J. S. (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.

Empirical papers

The latter part of the course will consist of reading empirical papers through the lens of some of the methods considered in the first part. Students will present an empirical paper.

Banerjee, A. V., Chandrasekhar, A. G., Duflo, E., & Jackson, M. O. (2017). Using gossips to spread information: Theory and evidence from two randomized controlled trials. http://web.stanford.edu/~arungc/BCDJ_gossip.pdf	April 3	Yuting
Sahni, N. S., Wheeler, S. C., & Chintagunta, P. (2018). Personalization in email marketing: The role of noninformative advertising content. Marketing Science, 37(2), 236-258.	April 3	Cathy
Gunter Hitsch and Sanjog Misra. Heterogeneous Treatment Effects and Optimal Targeting Policy Evaluation. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3111957	April 3	Keyan

Allcott, H., Braghieri, L., Eichmeyer, S., & Gentzkow, M. (2019). The Welfare Effects of Social Media. NBER Working Paper No. w25514. http://web.stanford.edu/~gentzkow/research/facebook.pdf	April 10	Eaman
Berman, R., Pekelis, L., Scott, A., & Van den Bulte, C. (2018). p-Hacking and false discovery in A/B testing. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3204791	April 10	Marat
Royer, H., Stehr, M., & Sydnor, J. (2015). Incentives, commitments, and habit formation in exercise: Evidence from a field experiment with workers at a Fortune-500 company. American Economic Journal: Applied Economics, 7(3), 51-84.	April 17	Amir
Bakshy, E., Rosenn, I., Marlow, C., & Adamic, L. (2012). The role of social networks in information diffusion. In Proceedings of WWW. ACM.	April 17	Qi
Chen, Y., & Yang, D. Y. (2018). The Impact of Media Censorship: 1984 or Brave New World? Forthcoming in American Economic Review.	April 17	Madhav
Lambrecht, A., Tucker, C., & Wiertz, C. (2018). Advertising to early trend propagators: Evidence from Twitter. Marketing Science, 37(2), 177-199.	April 24	Leon
Simonov, A., Nosko, C., & Rao, J. M. (2018). Competition and crowd-out for brand keywords in sponsored search. Marketing Science, 37(2), 200-215.	April 24	Artem
Sun, T., Viswanathan, S., & Zheleva, E. (2017). Creating social contagion through firm mediated message design: Evidence from a randomized field experiment. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2543864	April 24	Yuan
Berry, G., & Taylor, S. J. (2017). Discussion quality diffuses in the digital public square. In Proceedings of WWW. IW3C2.	May 1	Zivvy
Centola, D., Becker, J., Brackbill, D., & Baronchelli, A. (2018). Experimental evidence for tipping points in social convention. Science, 360(6393), 1116-1119.	May 1	Manon
Kalla, J. L., & Broockman, D. E. (2018). The minimal persuasive effects of campaign contact in general elections: Evidence from 49 field experiments. American Political Science Review, 112(1), 148-166.	May 1	Yifei
Breza, E., & Chandrasekhar, A. G. (2019). Social Networks, Reputation, and Commitment: Evidence From a Savings Monitors Experiment. Econometrica, 87(1), 175-216.	May 8	Alex
Crépon, B., Duflo, E., Gurgand, M., Rathelot, R., & Zamora, P. (2013). Do labor market policies have displacement effects? Evidence from a clustered randomized experiment. The Quarterly Journal of Economics, 128(2), 531-580.	May 8	Zanele

Doctor, J. N., Nguyen, A., Lev, R., Lucas, J., Knight, T., Zhao, H., & Menchine, M. (2018). Opioid prescribing decreases after learning of a patient's fatal overdose. Science, 361(6402), 588-590.	May 8	Marie
Gerber, A. S., Green, D. P., & Shachar, R. (2003). Voting may be habit-forming: Evidence from a randomized field experiment. American Journal of Political Science, 47(3), 540-550.	May 15	Jerry
Toussaert, S. (2018). Eliciting Temptation and Self-Control Through Menu Choices: A Lab Experiment. Econometrica, 86(3), 859-889.	May 15	Shuyi
Sparkman, G., & Walton, G. M. (2017). Dynamic norms promote sustainable behavior, even if it is counternormative. Psychological Science, 28(11), 1663-1674.	May 15	Matt

Requirements

This section provides an overview of the main requirements of the course. More details about the requirements will be provided later.

Class participation (20%)

Attendance is required. Everyone's learning can benefit from informed discussion. After the first few sessions, students will have extra responsibility for particular experimental papers, for which they will help guide the discussion. This will apply to approximately 6 to 8 of the class sessions. The number of people in the class will affect the nature of this requirement.

Quiz (15%)

A short in-class quiz on some of the foundational topics.
Quiz: Wednesday, February 27

Assignment (20%)

There will be one assignment focused on implementing some of the foundational topics. In particular, students will implement simulations for planning an experiment, Fisherian randomization inference, and basic counterfactual policy evaluation estimators.
Assignment due: Monday, March 18 (midnight Eastern)
Contest submissions accepted until Wednesday, March 20 (midnight Eastern)

Paper (45%)

One requirement is a term paper. Students will write a short proposal for the term paper.

Paper proposal due: Monday, April 1st
Paper due: Friday, May 10

One recommended format for the term paper is a proposed experiment. The paper would detail the theoretical motivation, practical considerations, the design, and the proposed analysis, along with more methodological motivations for the specifics of this (e.g., power analysis). Ideally, you actually want to run an experiment in this area. Alternatively, it might be that you want to conduct a related observational study, in which case thinking about the ideal, close-to-feasible experiment is quite useful (and you would perhaps have data to do good prospective power calculations with).

Other formats include a methodological paper or reanalysis of existing experimental data (e.g., using counterfactual policy evaluation).

Further guidelines and policies

I intend to follow all relevant policies. Let me know if you think I'm missing something.

MIT Sloan Values (https://mysloan.mit.edu/offices/deans/values/Pages/default.aspx)