



THE UNIVERSITY OF
CHICAGO

DEPARTMENT OF STATISTICS

Statistics Colloquium

Data Science Institute/Computer Science/Statistics Candidate Seminar

AARON SCHEIN

Postdoctoral Fellow
Columbia University

“Measurement and Experimentation in Complex Sociopolitical Processes”

THURSDAY, March 10, 2022, at 4:30 PM

In-Person: John Crerar Library, Room 390 and

Remote: [Live Stream](#) or Zoom

(Zoom information sent via email announcement)

Host: Victor Veitch

ABSTRACT

Complex social and political processes at many scales—from interpersonal networks of friends to international networks of countries—are a central theme of computational social science. Modern methods of data science that can contend with the complexity of data from such processes have the potential to break ground on long-standing questions of critical relevance to public policy. In this talk, I will present two lines of work on 1) estimating the causal effects of friend-to-friend mobilization in US elections, and 2) inferring complex latent structure in dyadic event data of country-to-country interactions. In the first part, I will discuss recent work using large-scale digital field experiments on the mobile app Outvote to estimate the causal effects of friend-to-friend texting on voter turnout in the 2018 and 2020 US elections. This work is among the first to rigorously assess the effectiveness of friend-to-friend “get out the vote” tactics, which political campaigns have increasingly embraced in recent elections. I will discuss the statistical challenges inherent to randomizing interactions between friends with a “light touch” design and will describe the methodology we developed to identify and precisely estimate causal effects. In the second part of this talk, I will discuss hierarchical Bayesian modeling of dyadic event data sets in international relations which contain millions of micro-records of the form “country i took action a to country j at time t ”. The models I will discuss blend elements of tensor decomposition and dynamical systems to capture complex temporal and network dependence structure in the data. Approximate posterior inference relies on new auxiliary variable augmentation schemes and theorems about the properties and relationships between different discrete distributions. At the end of the talk, I will outline the future of both lines of work, as well as their intersection, and sketch a broader vision for how data science can serve computational social science and vice versa.

Bio: Aaron is a postdoctoral fellow in the Data Science Institute at Columbia University, where he is co-advised by David Blei and Donald Green. His research develops machine learning and data science methods for computational social science. His recent work uses large-scale digital field experiments to assess the causal effects of friend-to-friend mobilization on voter turnout in US elections. Aaron did his PhD at UMass Amherst in Computer Science, where he was advised by Hanna Wallach. His doctoral research developed new hierarchical Bayesian models, tensor decomposition methods, and dynamical systems for analyzing massive data sets of country-to-country events. Aaron has interned at Google and Microsoft Research and worked in policy at the MITRE Corporation. He is also currently a senior technical advisor at Ocurate and a research affiliate at PredictWise. Prior to doing his PhD, he received a BA in Political Science and an MA in Computational Linguistics from UMass Amherst. He is on Twitter @AaronSchein.

For further information and inquiries about building access for persons with disabilities, please call 773.702.8333. If you wish to subscribe to our email list, please visit the following website: <https://lists.uchicago.edu/web/subscribe/statseminars>.