

Yiqing Xu
77 Massachusetts Avenue, E53
Cambridge, MA 02139
xyq@mit.edu

March 4, 2015

Political Methodology Search Committee
c/o Professor Megumi Naoi
Department of Political Science,
University of California, San Diego
Social Sciences Building 301,
9500 Gilman Drive,
La Jolla, CA 92093-0521



Dear Members of the Search Committee,

I am writing to apply for the assistant professor position in Political Methodology listed on the website of University of California San Diego open recruitments. I am a fifth year PhD student in Political Science at Massachusetts Institute of Technology and am completing my dissertation, entitled “*Causal Inference with Time-Series Cross-Sectional Data with Applications to Chinese Political Economy*,” under the advisory of Professors Teppei Yamamoto and Jens Hainmueller. I expect to finish my PhD in June, 2016.

To date I have published two articles: one in *American Political Science Review (APSR)* and one in *Journal of Statistical Software*. In addition, two of my papers are currently under revision and resubmission in *American Journal of Political Science (AJPS)* and *Political Science Research and Methods (PSRM)*. I was also awarded the 2014 *John T. Williams Dissertation Prize* by the Society of Political Methodology. These articles, as well as my dissertation, reflect my interest and ability in developing new tools for causal inference and applying them to a wide range of important political science questions.

The core and methodological part of my dissertation is what I call the *generalized synthetic control* method. It unifies fixed effects models, including difference-in-differences (DID), and the synthetic control method under a single framework. Specifically, it improves causal inference with time-series cross-sectional (TSCS) data when treatment units are not randomly selected. Such cases are ample in political science. For example, researchers might be interested in the effect of a reform that took place in several US states, but those states could be fundamentally different from the rest of the country. Conventional two-way fixed effect models may not be useful when the average treated counterfactual and average control outcome do not follow parallel paths (in which case the “parallel trends” assumption fails). The generalized synthetic control method addresses this challenge. The basic idea is to take

into account unobserved time-varying confounders by decomposing the error structure into lower-dimensional factors and conditioning on these factors. This method is in the spirit of the original synthetic control method in the sense that, like synthetic control, it uses pre-treatment periods to learn the relationships between treated and control units, based on which it predicts counterfactuals for each treated unit.

Besides the reform example I mentioned above, this method is widely applicable in political science. For instance, researchers can use this method to estimate the effect a country's joining an international organization on its probability of having conflicts with other countries, or to examine the effect of foreign aid on economic growth. In both cases, we cannot readily assume the "parallel trends" assumption to be valid. This method has several attractive features. First, because it allows the treatment to be correlated with unobserved unit and time heterogeneities, it is more robust and often more efficient than conventional fixed-effect models. Second, as its name suggests, it generalizes the synthetic control method to the case of multiple treated units and variable treatment timing. With this method, users no longer need to find matches for each treated unit since the algorithm produces treated counterfactuals in a single run. Moreover, it addresses the inferential problem of the original synthetic control method and gives more interpretable uncertainty estimates. Finally, with a built-in cross-validation procedure, it avoids specification searches and thus is easy to implement.

In the future I would like to extend this project to broader and more comprehensive studies of causal inference with TSCS data. For practical and ethical reasons, many important questions in comparative politics, international relations, and political economy cannot be easily studied by experimental or quasi-experimental approaches. With more rigorous statistical tools and more careful research designs, researchers can significantly improve our understanding of these real-world questions based on observational TSCS data. Two projects closely related to my dissertation are currently under investigation. First, I plan to address the inferential issue of the DID type of estimators when the treatment group has only a few units—in my dissertation, I rely on a cross-validation scheme to simulate prediction errors for treated counterfactuals, a technique that can be used more widely and systematically. Second, I would like to investigate the possibility of using machine learning tools, including an ensemble of multiple algorithms, to forecast treated counterfactuals in a DID setup. Details need to be worked out to accommodate different data generating processes and to address the inferential problem.

To me, political methodology means more than developing new statistical methods. It also involves innovative applications of cutting-edge statistical and game-theoretic tools to core political science questions. For example, using TSCS analyses, I investigate the effect of democratization on state capacity in previously authoritarian countries and the effect of partisan composition on state policy liberalism in the United States. My main regional focus is China, a country that I am most familiar with and whose authoritarian institutions and features of governance require more thorough and rigorous studies.



MASSACHUSETTS INSTITUTE OF TECHNOLOGY

I have written a couple of papers around this substantive interest, including one published in *APSR* and two under revision and resubmission in *AJPS* and *PSRM*. In the *APSR* paper, I am interested in whether and how informal institutions can sustain the quality of governance in environments of weak democratic or bureaucratic institutions, specifically, the role of informal institutions on public investment in rural China. I address the problem of measuring informal institutions by using village leaders from large clans as a proxy. I then employ both TSCS and regression discontinuity research designs to establish causality. I find a large and positive effect of informal institutions on local public investment and show that they help local leaders overcome the collective action problem of financing public goods. The second paper examines the sources of authoritarian responsiveness based on a field experiment. It is the first one that systematically measures government responsiveness in authoritarian counties and pursues its determinants. Together with my coauthors, I submitted information requests, with slightly different languages of four treatment and control conditions, to 2,103 online forums set up by Chinese county governments. We find that threats of collective action and tattling to the upper-level government cause county governments to be considerably more responsive. The third paper adopts a game-theoretic approach to understand the puzzling fact that reforms and propaganda coexist in authoritarian regimes. It provides an explanation that the government's policy concessions lend credibility to the information it releases and make the latter more effective in reshaping citizens' belief in the regime's popularity. Together, these papers demonstrate my abilities to creatively collect and analyze both experimental and observational data, as well as construct formal models, in the study of often opaque authoritarian politics.

I have served as a teaching assistant of both Quantitative Methodology I and II at MIT for first year PhD students and an instructor for the math refresher for second years. I developed a major part of the problem sets for the main quantitative sequence and lecture notes for the math refresher. I enjoy teaching and care about helping my students learn, which is clearly illustrated by my teaching evaluations (e.g. overall score 6.5/7.0 for Quant II) and many anonymous student comments. My core teaching interests are in quantitative research methods, including research design, introductory and advanced causal inference, and topics ranging from TSCS analyses, machine learning, to text analysis. Because of my training and research in formal models, positive political economy, and authoritarian politics, I am also able to develop courses in these subfields.

Enclosed please find my curriculum vitae, writing samples, and teaching evaluations. Other published and working papers of mine are available on my website. Five letters from Professors Teppei Yamamoto, Jens Hainmueller, Danny H. Hidalgo, James M. Snyder, Jr., and Lily L. Tsai will arrive separately. Please do not hesitate to contact me if I can provide you with any additional information. I look forward to hearing from you.

Sincerely,

Yiqing Xu

