

Assistant Professor Search, Data Science

Dear Members of the Search Committee,

I write to express my interest in your call for an Assistant Professor in Data Science. I am an Assistant Professor of Instruction in the Department of Political Science at Northwestern University, where I teach courses on statistics, statistical computing, and computational social science and conduct research on applied statistics for the social sciences. Previously, I was a postdoctoral fellow in computational and quantitative methodologies at McMaster University. My work is published or forthcoming in leading outlets in the social and political sciences, including the *British Journal of Political Science*, *World Development*, and the *Journal of Experimental Political Science*.

My research agenda focuses on using statistics to improve research design before data collection. My current focus is on statistical precision. This is overlooked in the statistics, econometrics, and political methodology literature in favor of identifying unbiased estimators. Implicitly, this literature assumes that one can improve statistical precision by just increasing sample size. This is not feasible in many social science applications due to practical or ethical considerations.

Focusing on the design and analysis of surveys and experiments, this agenda seeks to shape applied research by focusing on cases where one can seemingly improve statistical precision without sacrificing unbiasedness. As I show in my work, this usually implies unforeseen costs in other dimensions.

For example, in “Balancing Precision and Retention in Experimental Design”, we discuss how implementing alternatives to the standard experimental design, such as block randomization, may attenuate expected precision gains via explicit or implicit sample loss, a concern that prevents researchers from applying these techniques widely. Through three replications and six reanalyses of previously published experiments in leading political science journals, we show how precision gains from alternative designs can withstand significant degrees of sample loss.

As another example, in a solo-authored publication in the *Journal of Experimental Political Science*, I discuss the unforeseen costs of implementing double list experiments. This is a variant of the list experiment that promises narrower confidence intervals but comes with under-explored questionnaire design complications in the form of carryover design effects, a special kind of question order effect. I introduce statistical tests to diagnose this effect, which in turn facilitate the implementation of a more efficient technique.

One of the core lessons from my research program on statistical precision is that combining different techniques helps overcome their respective limitations. For example, in work in progress, we combine list experiments with questions from the network scale up method (NSUM), a popular technique in the health sciences, to improve the estimation of sensitive attitudes and behaviors. On the one hand, list experiments suffer from low statistical precision. On the other hand, generalizing to a population of interest through NSUM requires assumptions that are untenable in social science applications. By using NSUM questions as auxiliary information to the list experiment, we improve precision without introducing cumbersome assumptions.

My research also influences substantive work in the social sciences. In a working paper, we follow on our efforts to incorporate NSUM into social science applications by documenting the prevalence of criminal governance strategies in Uruguay. This is the basis of an European Research Council grant application seeking to understand criminal governance in least-expected contexts from a comparative perspective, which will in turn serve as a platform for further methodological innovation.

As another example, in a manuscript forthcoming at the *British Journal of Political Science*, we use an information experiment in Argentina to study gendered differential reactions to policy implementation. Previous work suggests that women face higher scrutiny for their performance in office. However, in the context of the implementation of a food distribution program, we find that voters are only responsive to performance information among men officeholders and tend to ignore performance information when told that an officeholder is a woman. We attribute this result to voters' perception of men politicians as the default category, and the tendency to believe that women politicians' good performance stems from factors beyond the incumbent's control. This finding calls attention to an unforeseen informational imbalance on survey experimental work in the subject, which future studies should take into account.

My teaching focuses on making statistics accessible to diverse audiences through a combination of flexibility and accountability. At Northwestern, I am the central person teaching statistics courses in the department. I teach the first course in the PhD methods sequence, focusing on probability and statistical inference. I also lead the math camp for incoming political science and sociology students and run the year-long R workshop that introduces cutting-edge statistical programming practices. Later this year, I will teach an undergrad-level introductory course that helps political science majors become informed consumers of applied statistics, and a seminar on design-based causal inference and machine learning methods for evidence-informed decision making in academia, government, and industry.

Before joining Northwestern, I taught data analysis for public policy and public opinion at McMaster and evidence-based policy to address social and political challenges in developing democracies at Tulane. Both courses emphasized the use of applied statistics to generate credible evidence. Beyond the classroom, my previous role as the methods editorial assistant for the *American Political Science Review* gave me the opportunity to shape and influence the development of cutting-edge methods in the field, a goal that I continue to pursue through service and mentoring.

I am prepared to teach courses on probability and statistical inference, statistical computing, linear models, design-based causal inference, machine learning, and computational social science. You can find copies of current and future syllabi in my website.

I believe my expertise makes me an excellent fit at the University of Guelph. If you have any questions, you can contact me via email or phone.

Sincerely,

Gustavo Diaz
Department of Political Science
Northwestern University