Gustavo Diaz, Northwestern Political Science, Scott Hall ste 224, 601 University Place, Evanston IL 60208

Tenure Track Faculty Search, Statistics and Data Science

Dear Members of the Search Committee,

I write to express my interest in your call for faculty positions in Statistics and 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. My work is published or forthcoming in leading social science outlets, including *Political Analysis*, 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 social science 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 the social sciences applications due to practical or ethical considerations.

Focusing on survey and experimental data, 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", published in *Political Analysis*, I discuss how implementing alternatives to the standard experimental design, such as block randomization or repeated measures, 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, I 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 collaborative work in progress, I 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 interdisciplinary work in comparative politics and political behavior. In a working paper, I follow on my team's efforts to incorporate NSUM into social science applications by documenting the prevalence of criminal governance strategies in Uruguay. This is the basis of a recently awarded *European Research Council Starting Grant* 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 in survey research.

As another example, in an article recently published in the *British Journal of Political Science*, I 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, this paper 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. This stems from 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 methods courses in the department. I teach the first course in the PhD methods sequence, focusing on Probability and Statistical Inference, the required undergrad-level Introduction to Empirical Methods, and an undergraduate research seminar on how experimentation and machine learning are used in academia, government, and industry to inform decision-making. I also lead the Math Camp for incoming political science and sociology students and run a year-long Statistical Computing workshop that introduces cutting-edge statistical programming practices. This year, I am adding a graduate seminar on Machine Learning to my portfolio.

I believe my experience teaching statistics courses in a PhD program is comparable to that required to offer courses at all levels in an undergraduate program in statistics and data science. Beyond the classroom, I routinely work with and advise undergraduate students on topics related to applied statistics. Currently, I serve as a mentor for the Farrell Fellowship program at Northwestern, which paired me with a rising sophomore student to collaborate on various projects on experimental design. I also advise two senior theses. One in Political Science with emphasis on experimental data, and another in the program in Mathematical Methods for the Social Sciences.

I am prepared to teach courses on probability and statistics, statistical inferences, design-based causal inference, machine learning, statistical programming, and data visualization. You can find copies of current and future syllabi in my website.

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

Sincerely,

Gustavo Diaz
Assistant Professor of Instruction
Northwestern University