

Gustavo Diaz

*Kenneth Taylor Hall 527
1280 Main Street West
Hamilton, ON Canada*

✉ diazg2@mcmaster.ca

🌐 gustavodiaz.org

☎ +1 217 904 0581

September 20, 2023

Dear Members of the Search Committee,

I write to express my interest in your call for an Assistant Professor in Statistics and Machine Learning. I am a postdoctoral fellow in Advanced Statistical, Causal Inference, and Computational Methodologies in the Department of Political Science at McMaster University. I received my PhD from the Department of Political Science at the University of Illinois at Urbana-Champaign under the supervision of Jake Bowers, Matt Winters, Gisela Sin, and Avital Livny.

I study and teach the application tools from data science and design-based causal inference to develop standards to improve research design in the social sciences. I focus on applications in which human behavior prevents us from observing the phenomena we wish to understand directly, thus demanding creativity in data analysis.

I have experience teaching courses on quantitative methods and their application to policy. At McMaster, I teach data analysis for public policy and public opinion. At Tulane, I taught a seminar on evidence-informed policy to address social and political challenges in developing democracies. Both courses emphasize the value of statistics in making credible policy recommendations. At Illinois, I served as a teaching assistant for statistics courses at the undergraduate and PhD levels using a flipped classroom approach. I also served as a math camp instructor for incoming graduate students for three consecutive years. I also have experience teaching political science courses in online and hybrid formats.

My involvement beyond the classroom also complements my teaching and mentoring. My work as the methods editorial assistant for the *American Political Science Review* exposes me to cutting-edge statistical applications in the social sciences. As a PhD student at Illinois, I started a collaborative project in which graduate students introduced their peers to new methods and organized a reading group on data science.

These experiences have prepared me to teach to a diverse student body, to adapt to both online and in-person platforms, and to teach both the theory and application of statistics.

I am prepared to teach courses on statistical inference, statistical programming, causal inference, data science, machine learning, and applications to evidence-informed policy. You can find copies of current and sample syllabi in my website. As a first-generation scholar, my teaching philosophy emphasizes using the language of statistics as a tool to advance the

careers of students with different backgrounds and goals.

My research agenda focuses on developing tools that researchers can adopt to improve statistical precision before data collection. This is overlooked in the causal inference literature in favor of identifying unbiased estimators. Implicitly, this literature assumes that one can improve statistical precision by increasing sample size. This is not feasible in many social science applications due to resource limitations. Moreover, resource considerations aside, even the least intrusive study has an ethical mandate to identify a research design that maximizes benefits and minimizes harm at the lowest possible cost.

Focusing on randomized controlled trials, this agenda follows two strands. First, I examine cases where one can improve statistical precision without sacrificing unbiasedness, which implies unforeseen costs in other dimensions. For example, in work under review with Erin Rossiter (Notre Dame), we show how implementing alternatives to the standard experimental design, such as pre-post outcome measurement or block randomization, may attenuate the expected gains in precision via explicit or implicit sample loss. In a solo-authored article published at the *Journal of Experimental Political Science*, I introduce new tools to assess the validity of estimates in double list experiments. This is a variant of the list experiment that promises more precise results but comes with under-explored questionnaire design complications. In work in progress with Inés Fynn (PUC-Chile), Verónica Pérez Bentancur (Universidad de la República), and Lucía Tiscornia (University College Dublin), we show how to combine list experiments with network scale up questions to improve the precision of prevalence rate estimates for sensitive attitudes and behaviors at the cost one additional assumption.

The second strand focuses on cases where one can improve precision by sacrificing unbiasedness deliberately. For example, in work in progress with Jake Bowers (Illinois) and Christopher Grady (USAID), we discuss the circumstances under which researchers should prefer biased yet precise estimators to analyze experimental data, including applications to block-randomization and M-estimation. In a *SAGE Handbook* chapter with Christopher Grady and Jim Kuklinski (Illinois), we discuss the merits and challenges of increasingly complex survey experimental designs that improve precision at the cost of external validity bias.

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

Sincerely,

Gustavo Diaz
Postdoctoral Fellow
Department of Political Science
McMaster University