

Research Statement

Gustavo Diaz

Department of Political Science, McMaster University

diazg2@mcmaster.ca

According to Freedom House, 2021 marks the 16th consecutive year with available data in which the number of countries where democracy deteriorated outnumbers the number of countries where democracy improved. This means roughly eight out of ten people in the world live in a country than can be considered “Partly Free” or “Not Free.” This trend emerges after several decades of continuous democratization.

At the same time, scholars, governments, and civil society organizations have redoubled their efforts to overcome the challenges to democracy around the world. Progress in research transparency and registration has lead to more coordinated efforts to produce credible evidence to determine what works, where, and why. The outstanding challenge in this area is that the empirical implications we wish to study are hard to observe directly.

My research focuses on the challenges to accountability and representation in the Global South. This is an area where strategic behavior from both citizens and politicians precludes the identification of successful interventions. My current agenda analyzes the unintended electoral consequences of investigating and exposing corruption, as well as the gendered patterns in citizens’ evaluation of politicians’ performance in office.

Given the focus on hard-to-observe phenomena, this substantive agenda also informs a methodological program on the application of tools from design-based causal inference and computational social science to navigate research design tradeoffs in experiments and quasi-experiments.

Challenges to accountability and representation in the Global South

The unintended electoral consequences of exposing corruption

The literature in political science and economics on electoral accountability highlights voters’ adverse selection problem: They prefer to have good over bad elected officials but they can only infer an incumbent’s type through observable outputs of their performance. Therefore, corrupt incumbents have incentives to hide illicit activities and appear good. To fight corruption, governments, civil society organizations, and scholars devote considerable resources to implementing and evaluating anti-corruption interventions that attempt to bridge the information gap between voters and politicians’ performance in office. However, the cumulative evidence suggests that these interventions rarely translate into voters sanctioning corrupt incumbents.

A recent explanation for the lack of electoral sanctions is that politicians undertake preemptive behavior to ward off the negative consequences of corruption. For example, previous research shows that political parties avoid nominating legislators investigated for corruption to reelection. On the other side of the coin, research also shows that local level elected officials switch away from parties investigated for corruption after securing reelection. In short, politicians avoid direct association with corruption.

I expand on this research in work funded by the Lemann Center for Brazilian studies at the University of Illinois, using data from a long-running anti-corruption program in Brazil that randomly selects municipalities to audit their use of federal funds, and combining text analysis and machine learning to create the most comprehensive dataset of corruption infractions to date. I make two contributions to the literature.

First, in a working paper, I argue that the main reason why politicians update their behavior in reaction to increased monitoring is because they expect increased scrutiny from their constituents. This contrasts from previous work suggesting that the primary incentive is top-down enforcement. I show that being selected for auditing leads mayors to shift public spending in an attempt to minimize the chances of uncovering irregularities. Mayors selected for auditing spend less and concentrate their spending in fewer areas when they are eligible for reelection and when the audits are announced close to an election year.

In another working paper, I argue that politicians are also responsive to revealed corruption even when they are not directly connected to it. This is because revealed corruption drives politicians in nearby localities to undertake preemptive behavior to ward off a negative reaction from their constituents. Whereas previous literature has shown evidence that partisans might react strategically to anticipated sanctions when their political party is linked to corruption, I argue that this behavior is even more widespread and that politicians fear being caught in an electoral anti-corruption wave even when there is no evidence of their own wrongdoing. This occurs because officeholders expect their constituencies to receive the news of corruption in nearby localities and react with increased scrutiny on their performance. I show that mayors in municipalities exposed to nearby corruption are more likely to seek reelection under a different party, and that they do so in a pattern that suggests an attempt to secure more resources for reelection.

Gendered evaluations of officeholders' performance

In work with Kelly Senters Piazza (US Air Force Academy), we expand this agenda toward the gendered patterns that underlie how voters evaluate politicians' performance in office in light of corruption. In a forthcoming book chapter in *Norms, Gender and Corruption: Understanding the Nexus*, we use my data on corruption infractions at the local level in Brazil to revisit previous findings on the gendered electoral consequences of corruption. Previous work in this topic uses public opinion data and survey experiments to show how revealing corruption increases support for female representation, yet also how female politicians face harsher sanctions for comparable levels of wrongdoing. We argue that these methods of inquiry do not allow us to translate their conclusions to realistic settings, which we overcome by focusing on the aforementioned anti-corruption program in Brazil. We show that revealing corruption increases support for female representation in the form of an increased proportion of female candidates running for mayor, yet this support does not translate into an increased proportion of women winning elections, which we attribute to incumbent mayors' incentives to counter the rise of female challengers. In another chapter forthcoming in the *Handbook on Gender and Corruption in Democracies*, we discuss the challenges and opportunities of using different data sources to study the relationship between gender and corruption, establishing a tradeoff between validity and generalizability.

This program extends to the study of gendered performance evaluations beyond corruption. In a piece in *World Development*, we highlight how exceptional events, such as corruption, conflict, and public health disasters, can promote female political representation by raising questions around the ability of male leadership to handle crises, as well as priming voters about issues around which women are perceived to have policy expertise. Examining public opinion data and recent media coverage around the world, we argue for the presence of both mechanisms in the current global pandemic, implying a potential surge in female political representation, which in turn can help alleviate the gender inequalities exacerbated by the pandemic.

In work under review with Virginia Oliveros (Tulane), Rebecca Weitz-Shapiro (Brown), and Matt Winters (Illinois), we use a survey experiment in Argentina to show gendered differential reactions to policy implementation. The literature in gender and politics identifies a double bind in performance evaluations, which implies that women officeholders are simultaneously perceived as better types and held to a higher standard than their male counterparts. Most of the literature attributes this differential treatment to different expectations of what

men and women should do in office. Our experiment highlights a situation where differential treatment exists even in a context of equally low expectations. We find that only men are receive credit for good policy implementation in the context of a food distribution program. We attribute this result to the voters' perception of male politicians as the default category, meaning that the mention of their gender does not provide any additional information on performance evaluation. On the flip side, mentioning the gender of a female politician leads voters to believe that good performance stems from factors beyond the incumbent's control.

Navigating research design tradeoffs

Improving precision before conducting an experiment

The last decade has seen considerable improvement in research transparency and registration. Recent advances in experimental design provide tools to diagnose the properties of a research design before data collection. For example, one can think about bias, power, or target sample size under different hypothetical data generation processes. My current agenda focuses on identifying practices and procedures that researchers can adopt at the pre-analysis stage to choose among alternative research designs. I emphasize improvements in statistical precision, the probability of detecting a non-zero effect (if it exists), since scholars studying the Global South often need to maximize limited resources.

In a manuscript conditionally accepted at the *Journal of Experimental Political Science*, I revisit double list experiments. Social scientists use list experiments in surveys when respondents may conceal their answer to sensitive questions. When their assumptions are met, list experiments reduce sensitivity biases from misreporting. However, they tend to produce estimates with high variance, which prevents researchers from improving upon direct questioning. Double list experiments promise to remedy this by implementing two parallel list experiments and then aggregating their results, which roughly halves the variance of the estimate for the prevalence of the sensitive trait. This implies an estimator that is more precise and still unbiased, but their implementation brings the question over whether the aggregation of the results of two parallel experiments yields a valid estimate. I propose a statistical test to assess the validity of estimates from double list experiments by leveraging variation in the order in which respondents receive the sensitive item, which provides researchers with a tool to apply this technique more widely.

In work under review with Erin Rossiter (Notre Dame), we discuss the circumstances under which adopting a research design feature aimed at improving precision can instead hurt precision through sample loss. For example, block randomization often improves precision, but it may require contacting participants multiple times to collect blocking covariates before treatment and then post-treatment outcomes. This creates space for either explicit (e.g. subjects dropping from the study between waves) or implicit sample loss (e.g. resource limitation forcing researchers to halve the sample size). Both forms of sample loss may offset the precision gains of blocking. Both standard and block randomized experiments have well-known unbiased estimators, but one must choose whether to block or not depending on the expected gains in precision. Through simulation and the reanalysis of a survey experiment on misinformation in social media, we illustrate how researchers can entertain this precision-retention tradeoff at the pre-analysis stage.

In work in progress with Jake Bowers (Illinois) and Christopher Grady (USAID), we discuss the circumstances under which researchers should prefer biased yet more precise estimators in the analysis of experimental data. Once again, an example of this trade-off comes from block-randomized experiments. Block-randomization entails grouping observations, conducting parallel experiments in each group, and then calculating a single treatment effect with a weighted average. Previous work in statistics suggests that block-size weights lead to an unbiased estimator of the average treatment effect in this setting. However, because experiments are

expensive to implement, we illustrate situations under which researchers may prefer to sacrifice unbiasedness in favor of more precise estimates, which would lead to the use of precision weights instead, the equivalent of using fixed effects in OLS regression. We extend this idea to other estimation techniques, such as M-estimators for outcomes variables with outliers. We illustrate these ideas through simulation and by revisiting the design of a messaging field experiment conducted by the Office of Evaluation Sciences in the United States.

In a chapter with Christopher Grady (USAID) and Jim Kuklinski (Illinois) in *The SAGE Handbook of Research Methods in Political Science and International Relations*, we discuss the trend of increasingly more complex research designs in survey experimentation. For example, short vignettes are replaced with factorial experiments including many attributes. This increase in complexity allows researchers to isolate causal effects from the role of potential confounders, but raises concerns over whether the tasks that respondents face in survey experiments tap into the intended construct outside of the survey framework. In other words, we draw attention to the trade-off between simple designs that risk confounding bias and complex designs that yield unbiased estimates but sacrifice ecological validity.

Geographic spillovers as a model selection task

I also maintain a research program on navigating research design trade-offs in the analysis of geographic spillovers in quasi-experiments. In causal inference, spillovers or interference happens when the treatment assignment of one unit affect the outcome of others. I focus on the cases in which the researcher believes that spillovers occur through geographical proximity. Current methods to study spillovers from a causal inference perspective assume the researcher observes all the relevant pathways through which treatments travel.

Researchers can justify this assumption through experiments designed to deliberately capture spillovers, but this level of control is otherwise rare. This is because researchers can credibly place units along a pathway, but theory alone does not guide how units connect to each other along that pathway. For example, if one believes that election monitoring efforts relocate election fraud to neighboring localities. How does one know what counts as a neighbor? One could commit to using distance or adjacency as a metric, but each alternative allows for multiple ways to operationalize spillovers. This implies an underlying tradeoff between narrow operationalizations that risk false positives and broad operationalizations that can introduce bias.

In a working paper recognized with the best poster award at the 2019 Latin American Political Methodology meeting. I propose to approach this problem as a model selection task. Researchers can express the theoretically relevant pathway as a distance metric that helps to identify all the plausible operationalizations for spillovers that follow from it. This allows researchers to be explicit about the gaps between theory and the assumptions required to justify model specification, and to use their preferred data-driven model selection criterion to choose the most appropriate model among alternatives. I illustrate the usefulness of this approach in the context of analyzing spillovers in voting registration irregularities in Ghana.