Women Politicians and Violence Against Women in Mexico Pre-Analysis Plan

Marco Alcocer, Rachel Skillman, Angie Torres-Beltran[1]

February 10, 2022

1 Pre-analysis plan

In this pre-analysis plan we pre-register a set of primary analyses for our paper tentatively titled "Women Politicians and Violence Against Women in Mexico." This is an observational study exploiting a Regression Discontinuity Design (RDD) to explore the outcome of local elections in Mexico for the 2018 election cycle.

While pre-registration of observational studies is becoming more common, not all researchers are engaging with this practice. We have decided to do so because we have a straightforward and intuitive research design that approximates the random assignment of treatment and control groups in experimental research and because our study relies on data which requires subjective coding decisions made by the researchers. By reporting our design, coding choices and expectations before seeing the results, we hope to mitigate concerns about these subjective coding decisions and potential "p-hacking", and therefore enhance the causal credibility of the study's findings.

We note that we are only pre-registering the paper's main hypotheses, which we commit to reporting in the paper. Any deviations from the pre-analysis plan will be noted and explained in the paper. Additionally, we plan on conducting additional tests to explore heterogeneous effects, and we commit to clearly communicating this in the paper. In this pre-analysis plan, we are focusing on the 2018 local elections, but given time and resource availability we might seek to extend the coverage to include additional elections between 2015 and 2017. If we do, the data collection will be identical to the collection of the 2018 data.

This pre-analysis plan outlines the project's motivation, the main hypotheses, the data collection procedures, data sources, and the research design to perform the statistical analysis. We attest that at the time this document was registered we have not had access to the main independent variable (we are still in the process of collecting the data) and thus have not conducted any of the statistical analyses

1.1 Motivation

The objective of this study is to investigate whether the presence of women politicians has an effect on the incidence and severity of violence against women (VAW). In short, we ask whether the presence of women politicians leads to better outcomes regarding VAW.

To answer this question, we focus on the Mexican case. Mexico offers an excellent and intriguing case for three reasons: (1) VAW is a major problem in Mexico, like in most countries, with 2 in 3 women experiencing some sort of violence against them (ENDIREH 2017; UN 2021); (2) gender-based policies, including VAW, are a salient political issue; and (3)

[1] Authors listed in alphabetical order.

Mexico has a notably high gender parity in politics with established gender quotas (Baldez 2004; Piscopo 2016; Htun and Weldon 2018), anti-violence legislation (Htun and Jensenius 2022; Htun and Weldon 2018), and is ranked fourth worldwide in women's representation (Inter-Parliamentary Union 2022). The high levels of women's representation in politics *and* high levels of VAW present a suitable case for disentangling the effects of women in power.

This topic is substantively important in political science because studying whether women politicians can affect VAW has important implications for not only democratic representation and participatory processes, but also for potentially reducing VAW. When women candidates become visible players in the political system, their presence and subsequent policy and practices are likely to empower and benefit other women (Desposato and Norrander 2008; Burns, Schlozman, and Verba 2001; Hansen 1997; Kock 1997; Verba, Burns, and Schlozman 1997). However, despite being such a fundamental public policy concern, little is known about the effectiveness of visible women candidates on other societal outcomes such as VAW.

1.2 Theory and hypotheses

Existing research and deductive theorizing about the Mexican case suggest several possible ways in which the election of women politicians could impact VAW and gender-based violence more generally. We focus on two main relationships: women politicians affecting (i) the actual prevalence of VAW and (ii) the probability that citizens report VAW.

***H1**: Municipalities that elect a woman politician will experience a reduction in overall VAW.*

Scholarship has shown that women's representation matters for a number of reasons and issues. Given that VAW is, by definition, an issue which victimizes women, and thus that women are more aware of, it is reasonable to believe that women politicians are more likely to create policies meant to address VAW and to reduce its prevalence.

At the same time, there are many other institutional and bureaucratic factors that likely condition the occurrence of gendered violence (e.g. resource availability, characteristics of the justice system and law enforcement agencies and civil society engagement). There are also political self-selection dynamics, socialization effects, and norms that may shape how women politicians behave. Because we are not able to account for all of these determinants, we may find that women politicians have a null effect on VAW. In the event of null results, we plan to explicitly theorize about what conclusions we can draw from the lack of effect. A key part of this theory-building exercise will involve interviews of women politicians in Mexico to assess what barriers might prevent them from impacting VAW.

***H2**: Municipalities that elect a woman politician will experience an increase in reports by individuals of crimes related to VAW.*

Beyond having or not having direct effects on the actual prevalence of VAW, women politicians may have an effect on how likely it is that VAW is reported by individuals. The effect could be driven by various mechanisms, including by actions that women politicians may take, such as being more vocal about VAW or directing authorities to prioritize VAW, or people's perception of women politicians, such as believing that women politicians take VAW more seriously.

***H3**: Municipalities that elect a woman politician will experience more reports of less severe forms of VAW than more severe forms.*

In addition to women politicians making it more likely that people report VAW incidents to authorities, there could be differential effects within this behavior if women politicians make it more likely that individuals report certain types of VAW more than others depending on the severity of that crime, e.g. sexual assault versus rape. This effect could be driven by the presence of women candidates driving individuals to report VAW, thus breaking the cycle of tolerating more extreme forms of VAW.

1.3 Identification strategy and right-hand-side variables

We exploit a natural experiment to infer the causal effect of women being elected to local office on VAW. We use a close-elections regression discontinuity design to estimate how electing a woman for mayor is related to the number and severity of instances of VAW in Mexican municipalities.

1.3.1 Unit on analysis

Our unit of analysis is the Mexican municipality.

1.3.2 Description of the voting system and municipal governments

Municipalities are headed by a municipal president (mayor) elected through plurality rule. Municipal elections occur every three years, and up until 2018 municipal presidents could not be reelected.

Mexico is a federal system where municipalities have constitutional powers over various duties, including public security, certain public spending, public works, social programs, and education, among others. By law, municipal governments are run by the municipal president along with the municipal council, though in practice there often exist few limits on the power of the municipal president. The municipal president, for example, has the power to appoint the heads of all municipal government agencies. Thus, municipal presidents have the power to set policy agendas and the means to pursue them.

1.3.3 Source of plausibly exogenous variation

We use a regression discontinuity design of close elections, as is now standard in the literature. Specifically, our research design exploits close races for the municipal president office. We restrict the estimation sample to the following cases:

1. A female candidate narrowly loses to a male candidate.
2. A male candidate narrowly loses to a female candidate.

The identification assumption is that, for this set of cases, whether or not a woman won or lost an election is as good as random. We find this highly plausible.

13.4 Election Data

Election data comes from Magar (2021).[2] This dataset contains the electoral results for all municipal elections in Mexico. We use municipal election data from 2018. The data includes the number of candidates, the votes that each candidate received, the total number of votes, and political party of the candidates, and the gender of the winning candidate.

While we plan on using the 2018 elections, given time and resources, we would like to extend this data to include all local elections between 2015 and 2017 as well.

1.3.5 Classifying candidates as women

While the data on elections includes the gender of the winning candidate that becomes municipal president, the gender of the candidates that do not win is not collected. This information is crucial for this study because the identification strategy relies on comparing municipalities where a woman candidate barely defeated a male candidate and municipalities where a male candidate barely defeated a woman candidate.

We thus plan on collecting the gender of the second place candidate for all local elections in 2018. At the time that this PAP is being registered, we are in the process of collecting this data.

Here we describe the data collection procedure for this information.

We have recruited two undergraduate research assistants (RAs) and provided them the political party and number of votes for each first and second place candidate for all municipal elections in 2018. This information also includes the gender of the winning candidate. To identify the gender of the second place candidate, the RAs have been instructed to search for the electoral results of each municipality in each state's electoral agency. This is because municipal election data is stored by each state's electoral agency, not the federal electoral agency. For each state, the RAs are searching for the list of candidates that includes their political party and electoral results. When the RAs identify the candidate that received

[2] Magar, Eric (2018) Recent Mexican election vote returns repository, https://github.com/emagar/elecRetrns

the second most electoral votes, they code whether this candidate is male or female based on their name.

In Mexico, the vast majority of names are easily attributable to a gender. Thus, the RAs code whether the name of the second place candidate is of a male or of a female. For names that are not gender specific, that the RAs could not code, or that the RAs are unsure about, they have been instructed to leave it blank and highlight it. The researchers, two of which are of Mexican origin, will then go through the names the RAs could not identify and make coding decisions based on the name, and if the name is still unclear, determine their gender based on background research on each one of these unknown candidates. In case this process still leaves names unclassified, we will use pre-existing R packages that predict the gender of an individual based on their name. Any name that cannot be clearly identified by this process will be left blank and excluded from the analysis.

This procedure has certain limitations, though we do not believe them to be prevalent or leading to bias. For transparency, we discuss these limitations here. This method bases the gender of individuals on their names, which may result in incorrect coding for a number of reasons. First, an individual may have a name that is traditionally associated with the other gender. Second, it may miscode individuals who use their birth name but who have transitioned genders. While we acknowledge that these may be issues present in our data, we do not believe that they are commonplace in this case nor that they introduce bias, for example, by being correlated to our outcomes of interest.

1.3.6 Bandwidth for narrow victory/loss

Regression discontinuity designs rely on close elections – in our case, on small margins of victory/loss for a woman candidate in mayoral races. Rather than selecting an arbitrary threshold to define "close elections," we rely on the growing literature which identifies the optimal bandwidth through a data-driven approach.

Specifically, we commit to estimating and reporting the two most prominent procedures to estimate optimal bandwidths: (1) Imbens and Kalyanaraman (2012)[3] and (3) Calonico, Cattaneo, and Farrell (2020).[4]

1.4 Outcome measures

VAW is an umbrella term that captures different forms of gender-based violence, from femicides (killing a woman due to their gender), to rape, to sexual harrassment, to emotional abuse, among others. In many cases, instances of VAW remain unreported, and therefore the prevalence of VAW and prevalence of reports of VAW are interrelated but may differ across

[3] Imbens G, Kalyanaraman K. 2012. Optimal bandwidth choice for the regression discontinuity estimator. Rev. Econ. Stud. 79(3): 933–59.
[4] Calonico, S., M. D. Cattaneo, and M. H. Farrell. 2020. Optimal Bandwidth Choice for Robust Bias Corrected Inference in Regression Discontinuity Designs. Econometrics Journal, 23(2): 192-210.

time and space for various reasons. We use various sources of data to measure these two interrelated but distinct concepts.

To measure actual instances of VAW we use:

1. Official femicide instances as defined by the Mexican government.
    a. In 2015 the Mexican government began collecting data on femicides. The government classifies intentional homicides as femicides if they are deemed to meet at least one of nine criteria (SESNSP, 3).[5]
2. The number of intentional homicides of women.
    a. We also plan to use the total number of intentional homicides of women to measure VAW. This is for robustness and because the femicide measure provided by the government is a conservative count. [6]
3. The number of intentional homicides of young women (15-45 years old).
    a. To more precisely measure intentional homicides of women that may be the result of gender-based violence, we create a new measure of intentional homicides of young women. The age range is chosen because sources have found that the majority of femicide victims in Mexico[7] and Latin America[8] fall within this age range.

The mortality data reporting intentional homicides of women comes from Mexico's National Institute of Statistics and Geography (Instituto Nacional de Estadística y Geografía or INEGI) and the femicide data from the crime registry of the Executive Secretariat of the National Public Security System (Secretariado Ejecutivo del Sistema Nacional de Seguridad Pública or SESNSP).

To measure the prevalence of reporting VAW incidents, we use the following:

1. Incidents of VAW crimes reported to authorities, which include:
    a. Sexual harrassment
    b. Sexual abuse
    c. Rape
    d. Domestic violence

---

[5] SESNSP. 2018. "Lineamientos para el registro y clasificación de los presuntos delitos de feminicidio para fines estadísticos"
[6] This is also becoming a standard practice of various human rights organizations in Mexico to account for the severe underreporting and lack of accurate data on femicides in Mexico.
[7] SEGOB, INMUJERES, ONU Mujeres. 2017. "La violencia feminicida en México, aproximaciones y tendencias 1985-2016."
[8] ECLAC. 2021. "At Least 4,091 Women Were Victims of Femicide in 2020 in Latin America and the Caribbean, Despite Greater Visibility and Social Condemnation." https://www.cepal.org/en/pressreleases/eclac-least-4091-women-were-victims-femicide-2020-latin-america-and-caribbean-despite

This data comes from the crime registry of the Executive Secretariat of the National Public Security System, which includes instances of reported crimes. One limitation is that these reported crimes are not disaggregated by the gender of the victim. However, since the vast majority of sexual crimes and domestic violence crimes are perpetrated against women, we believe that this data provides a valid measure of our proposed concept.

Any additional data sources we use for VAW-related outcomes that are not specified in this section will be clearly noted in the paper to clarify that they were not pre-registered.

All measures are available by municipality and year. We plan to use these outcomes by year of municipal administration (each year of the three-year terms) to see whether there are temporal effects as well as pooled (the three years of the administration combined) to see whether there are overall effects. We also plan on using both the rate of incidents per 100 thousand people and raw count.

1.5 Analysis

To estimate the regression discontinuity, we rely on procedures developed in the literature. Specifically, we utilize the rdrobust package in R for the main analysis (Calonico, Cattaneo, Farrell, and Titiunik 2017).[9] We estimate both linear and quadratic polynomial models.

Because the outcome measures are count variables, we also use a negative binomial regression to estimate the regression discontinuity for robustness.

1.6 Balance checks

As is standard in regression discontinuity designs, we will use pre-treatment variables to help validate the as-if random assumption. These tests are designed to assess balance between the treatment (woman candidate barely defeats male candidate) and control (male candidate barely defeats woman candidate) groups.

Specifically, we will use the 2010 census that provides information about each municipality, including population, economically active individuals, number of male/female residents, education levels, access to health care, marriage rates, religious affiliation, and number of homes, among others. We also use official data on each municipality's 2010 Gini and poverty levels from Mexico's National Council for the Evaluation of Social Development Policy.

We find it important to note, however, that due to chance, it should be expected that approximately 1 in 20 variables will lead to statistically significant differences with the significance level set at 0.05.

[9] Calonico, S., M. D. Cattaneo, M. H. Farrell, and R. Titiunik. 2017. rdrobust: Software for Regression Discontinuity Designs. Stata Journal, 17(2): 372-404.

1.7 Heterogeneous effects

We plan on exploring some important heterogeneous effects. Here we register three heterogeneous effects. Any test that we end up conducting that is not registered will be noted as exploratory to make it clear that it was not pre-registered.

First, we plan on testing whether effects differ depending on whether organized crime was present in a municipality. Organized crime is a major issue in Mexico that is sometimes tied to VAW, and especially to instances of femicides. Organized crime is also often associated with corruption in law enforcement agencies and local politics, perhaps making it more difficult to address VAW. We will use municipality-level data on the presence of organized crime in Mexico provided by the Mapping Criminal Organizations in Mexico project (Sobrino, Alcocer, Farfan Mendez, and Signoret 2022). This data covers up to 2018, so we will use the 2018 data to test whether municipalities with and without organized crime are affected differently by women politicians.

Second, we plan on testing whether there are differential effects based on pre-existing levels of violence. There may be reasons to believe that women politicians can address VAW differently in contexts of high versus low violence. This is especially salient in Mexico where some municipalities experience very high levels of violence while others do not. We will use data on intentional homicides to code municipalities as highly violent or not. While this threshold is arbitrary, for transparency, we plan on coding municipalities with homicide rates per capita greater than the third quartile of all municipalities in 2018 as highly violent municipalities.

Third, women's movements have become prevalent in Mexico, with many centering specifically on protesting against VAW. There is reason to believe that in municipalities where protests have taken place, politicians are more responsive to these issues—particularly women politicians—and therefore more likely to address them. To test for this, we use event data from the Armed Conflict Location & Event Data Project (ACLED) on protests, attempting to narrow in on protests dealing with women issues or VAW if possible, in order to code municipalities as having experienced these protests or not.

1.8 Standard errors

We will estimate heteroskedasticity-robust standard errors. In the case that we are able to collect data from years beyond 2018, we will cluster our standard errors to the municipality level.