IDB WORKING PAPER SERIES Nº IDB-WP-1379

When Women Win:

Can Female Representation Decrease Gender-Based Violence?

Veronica Frisancho
Evi Pappa
Chiara Santantonio

Inter-American Development Bank
Department of Research and Chief Economist

October 2022

When Women Win:

Can Female Representation Decrease Gender-Based Violence

Veronica Frisancho*
Evi Pappa**
Chiara Santantonio***

* Inter-American Development Bank
** Universidad Carlos III de Madrid and CEPR
*** Luiss University, email: csantantonio@luiss.it

October 2022

Cataloging-in-Publication data provided by the
Inter-American Development Bank
Felipe Herrera Library

Frisancho Robles, Verónica C.
When women win: can female representation decrease gender-based violence? / Veronica Frisancho, Evi Pappa, Chiara Santantonio.
p. cm. — (IDB Working Paper Series ; 1379)
Includes bibliographic references.
1.Violence in women-Economic aspects-United States. 2. Women-Crimes against-Economic aspects-United States. 3. Regression analysis. I. Pappa, Evi. II. Santantonio, Chiara. III. Inter-American Development Bank. Department of Research and Chief Economist. IV. Title. V. Series.
IDB-WP-1379

http://www.iadb.org

Abstract

Every day, three women are murdered in the United States by a current or former partner. Yet policy action to prevent gender-based violence has been limited. Previous studies have highlighted the effect of female political representation on crimes against women in the developing world. This paper investigates whether the election of a female politician reduces the incidence of gender-based violence in the United States. Using a regression discontinuity design on mixed-gender races, we find that the election of a female House Representative leads to a short-lived decline in the prevalence of femicides in her electoral district. The drop in femicides is mainly driven by a deterrence effect that results from higher police responsiveness and effort in solving gender-related crimes.

Keywords: Gender-related violence, Close elections, Police effort, Regression discontinuity

JEL Codes: D72, J12, J16.

1 Introduction

Domestic violence has become a silent epidemic in the United States. One in four American women have experienced severe physical violence by their partner. Twenty percent of women in the United States will be raped in their lifetime, and half of these victims are raped by an intimate partner (NCADV, 2014). In its most extreme form, gender-violence also kills women: every day, three women are murdered by a current or former partner (Matawaran, 2021). Seventy percent of all femicides in high-income countries occur in the United States (Hemenway et al., 2002). Even after adjusting for population size, the domestic rate of femicides is double compared to, for example, that of France.

Despite the magnitude of the problem, Americans often think of gender-based violence as an issue concerning poorer countries or segregated and marginalized communities at home. This is in part explained by the country's inaction in both generating awareness of and addressing gender-based violence. Contrary to the progress made in Latin America and Europe, the United States has not even passed legislation to specifically prosecute femicide. There is not a single law calling femicide a crime.

While there is a pressing agenda to change the country's approach to gender-based violence, there is also an urgent need to work on the prevention angle. Previous studies, mostly focusing on developing countries, have tried to identify protective factors such as promoting women's economic empowerment (Hidrobo et al., 2016, Amaral, 2017), fostering women's trust in the police (Amaral et al., 2021), and limiting the use of substances (Luca et al., 2015). A much less explored factor is the role of female representation, either as directly influencing policy or as a way to change the perception of female leadership and of women's social status in general. This paper investigates whether the election of a female Representative to the U.S. House reduces the incidence of gender-based violence in the United States. We show that the election of a female Representative leads to a short-lived decline in the prevalence of femicides in her electoral district. This drop in femicides is mainly driven by a deterrence effect that results from increased responsiveness and effort by the police for solving gender-related crimes.

We rely on a regression discontinuity design on close mixed-gender races to identify the effect of female political leaders on gender-based violence. We combine official data on reported crime from the Federal Bureau of Investigation (FBI) with data on elections for the U.S. House of Representatives for the period 1970-2017. We focus on two distinct felonies, femicides and rapes, which vary in their degree of misreporting. While there is a large set of studies that argue that there is non-random measurement error in the self-reporting of violence (United Nations, 2014, Palermo et al., 2014, Agüero and Frisancho, 2022), we argue that our measure of domestic violence based on femicides is less biased since it is a fatal crime that is hard to misreport. Looking at the impact of female representation on both femicides and rapes allows us to better disentangle the effects on reporting versus actual changes in the prevalence of violence.

According to our main estimates, the number of femicides in a district decreases by about 10 percent the year a female Representative is elected, but we find no effect on rapes. We show that the absence of significant effects on reported rapes can be attributed to the combination of increased reporting and lower incidence of that crime, factors that offset each other. Looking, instead, at the entire duration of the term, i.e., the two years in office, we see that the effect on femicides vanishes. However, there is a significant increase of around 25 percent in the number of rape victims. We argue that this pattern is consistent with higher reporting rather than an increase in the actual prevalence of violence, a finding which is in line with previous studies (Iyer et al., 2012).

We explore two potential mechanisms that may be driving our results: police effort and attitudes towards reporting and the police. During the first year in office, we see a decrease of 7.5 percent in the number of female victims of unsolved cases of murder in the police records, while we fail to find such an effect when the victim is male. Furthermore, over the entire term, the number of clearances in the case of rapes goes up by around 50 percent. The analysis of arrests data also seems to suggest increased police effectiveness in solving violent crimes perpetrated by men.

Relying on survey data from the American National Election Study (ANES), we also show that after the election of a female Representative women exhibit greater propensity to complain about harassment and discrimination as well as more favorable attitudes towards the police. Both of these results are consistent with the hypothesis of women becoming more likely to report gender-based crimes. In turn, we rule out that the election of a female Representative leads to any changes in men's views about reporting or attitudes towards the police.

We perform a series of robustness tests to corroborate our findings. First, our results are very stable to alternative definitions of the outcome variables and the analysis sample. Second, we rule out the concern that the decrease in gender-based violence is driven by candidates' party. Even though elected women disproportionately come from the Democratic rather than the Republican party, we show that electing a Democrat over a Republican does not have a significant impact on gender-based violence.

Previous studies have shown that the election of a female leader is linked to more favorable outcomes for women. On one hand, female political leaders can change the perception of female leadership and of women in general (Beaman et al., 2009), fostering changes in gender-biased social norms. Beaman et al. (2012), Priyanka (2020) show that female leaders can become role models and expand girls' educational and career aspirations. Female leaders' preferences may also lead to policies that are more liberal, favor redistribution, and benefit children (see, e.g., Brollo and Troiano (2016), Bruce et al. (2022), Clots-Figueras (2012), Bhalotra et al. (2022), Bhalotra and Clots-Figueras (2014)). They also tend to invest in public goods which are closer to women's concerns (Chattopadhyay and Duflo, 2004), favor "women-friendly" laws (Clots-Figueras, 2011), introduce more bills related to women's issues (Gerrity et al., 2007) and have more active roles on such issues (Lippmann, 2022). Our paper builds on these studies and asks whether women in power can also play a role in reducing gender-based violence.

This paper also speaks to the economic literature that seeks to identify protective factors to reduce violence against women. While previous studies have focused on cash-transfers and vouchers (Hidrobo et al., 2016), laws that limit alcohol consumption (Luca et al., 2015), increased property rights (Amaral, 2017), or the establishment of women police stations (Amaral et al., 2021), less attention has been devoted to the role of female representation. There are two notable exceptions: Iyer et al. (2012) and Delaporte and Pino (2022). The former exploit the introduction of gender quotas in local councils in India to show that the appointment of women significantly increases the prevalence of crime against women. By exploiting the variation in the likelihood of reporting biases across different crimes, they argue that this effect is driven by a higher likelihood to report rather than greater prevalence rates. Delaporte and Pino (2022) provides a more recent contribution for Brazil, focusing on the effect of female mayors on violence against women. This paper is more closely related to ours as it also relies on a regression discontinuity design for close mixed-races. Using this design and administrative records on all episodes regarding confirmed or suspected gender violence between 2005 and 2016, the authors find that electing female mayors leads to a 63% reduction in the prevalence rate of domestic violence.

Our study provides at least four main contributions. First, building on Iyer et al. (2012), we leverage data on femicides and rape records to disentangle between changes in reporting and changes in the actual prevalence of violence against women. Second, our research design poses an advantage over the use of the staggered enactment of quotas, as these reforms might actually alter the very nature of political competition (Beaman et al., 2009, Clots-Figueras, 2012). Third, we provide evidence on potential channels of impact including police responsiveness as well as attitudes towards reporting and the police among men and women. In contrast to Delaporte and Pino (2022), our results favor a role model effect over a policy effect. Most importantly, and unlike the existing literature, we extend the question to the United States, one of the wealthiest countries in the world with levels of gender-violence well above the average among rich countries.

The remainder of the paper is as follows: Section 2 provides a framework to explain how female representation might affect violence against women, Section 3 describes the data used for this analysis, Section 4 presents the empirical strategy, Section 5 discusses the results, while Section 6 presents robustness tests for the main results. Finally, Section 7 concludes.

2 Female Representation and Violence Against Women

For any crime we can identify three key choices made by three corresponding agents (Donohue and Levitt, 2001, Iyer et al., 2012). First, the offender's choice of whether or not to commit the crime. Second, the victim's choice of whether or not to report the crime to the police. Third, the police's choice of whether to dismiss or impose charges. In this simple setting, the election of a woman may alter agents' incentives to commit a crime, to report it, or to pursue an investigation in different ways.

On the one hand, female representation could directly affect the probability of a gender-based crime being committed by improving males' views on females' role in the society, hence lowering the incidence of gender-based violence. For example, Beaman et al. (2009) show that long-term exposure to a woman in office improves perceptions about female leaders and tends to decrease gender biases in India. However, the opposite might happen if men resent the election of a female, yielding a backlash effect that may translate into greater levels of violence against women as a way to express their frustration or to intimidate female leaders and voters (Iyer et al., 2012).

A visible female political leader can also empower women to come forward and report violence. For instance, Kuipers (2020) finds a significant negative relationship between women's presence in local councils and the share of women who approve of a husband assaulting his wife in Indonesia. In this case, greater levels of reporting may artificially suggest an increase in the prevalence of crime.

Female representatives may also influence police behavior. If the likelihood of reporting increases, officers dealing with more gender-based crimes can become more likely to pursue them. In addition, an elected female leader can directly exert pressure on their behavior with respect to gender-related violence either by going to the press or approaching higher-level officials (Iyer et al., 2012). Higher police responsiveness could also have feedback effects on women's likelihood to report and men's likelihood to commit gender-based crimes. Greater responsiveness may, in fact, lead to higher levels of confidence in the police, making women more willing to report. At the same time, both increased reporting and increased policing should deter males from committing the crime in the first place.

The final net effect of the election of a female representative on the prevalence and reporting of gender-based violence will depend on the interaction of all these forces in a given context. Notice, in particular, that the influence of female politicians on the perpetrators, victims, and the police may lead to diverging changes in reporting and actual prevalence of violence. For instance, women may become more likely to report and police more responsive and effective when dealing with gender-based crimes, which can result in an increase in the prevalence rate of certain crimes that went under-reported before the election. While the actual prevalence may be decreasing, the administrative records may still show an increase whenever the uptick in reporting offsets the decrease in prevalence rates. This is less of an issue in the case of crimes that are less likely to suffer from misreporting biases, such as femicides. In this case, the change in prevalence rates can be attributed to actual changes in the crime rate rather than reporting.

3 Data

3.1 Election Data

Data on elections for the United States House of Representatives during the period 1970-2016 come from CQ Press. We obtain the sex of all candidates for the Republican (R) and the Democratic (D) party from PoliticalParity (2015), the Center for American Women and Politics (CAWP), and Pettigrew et al. (2014).[1]

Races for the U.S. House occur at the congressional district (CD) level, each district electing one Representative to Congress. Every two years, i.e., in even years, all 435 seats in the House are up for re-election, comprising a total of 10,440 races occurring between 1970 and 2016. Looking at races in which the top two candidates were either R or D,[2] we divide our sample into the following categories: man running against a man, woman against woman, and man against woman. Table 1 shows a break-down of elections by type of race, defined by the sexes of the top candidates and the winning party. Mixed-gender races are further divided by the sex of the winning candidate. Overall, mixed-gender races represent 18.2 percent of the whole sample of races (1,895), and appear to be more frequently won by male candidates (1,081 vs 814 wins).

Interestingly, while R and D do not substantially differ in the total number of mixed-gender races won (956 D vs 939 R), they do not seem to end up electing the same candidates: out of all the elections in which a woman defeated a man, the number of Democratic winners is more than twice that of Republicans, with D-Congresswomen representing 68 percent of women elected in this type of races (559 vs 255).[3] Instead, the opposite holds for mixed-gender elections won by

[1]These datasets do not include the sex of candidates from other parties. Because of this restriction and because of the strong bipolarity of the American party system, we focus on the sex of Republican and Democratic candidates only. On average, in all elections in which the first and second places correspond to R and/or D, the share of votes for these two candidates is above 98 percent.

[2]These include one-candidate races. We also considered as Democrats those whom CQ Press listed as Democrat Farmer-Labor (for Minnesota), or as Democrat-Open Primary. Similarly, we considered as Republicans those listed as Independent Republican (for Minnesota), Republican Write-in, and Republican-Open Primary. Races in which candidates from the same party occupy both first and second places are also included.

[3]These numbers also include mixed-gender races in which candidates belong to the same party. If we only consider mixed-gender races between a Democrat and a Republican, results are virtually unchanged: out of 1,880 races, 43 percent are won by a woman (805). Out of these, 550 are won by a Democratic woman, and 255 by a Republican woman. A similar proportion is also found when considering all races: out of a total of 1,063 women appointed to the U.S. House, irrespective of the type of election, 69 percent belonged to the Democratic party (736) and 31 percent to

Table 1 Races by the Sex of Top Candidates and Winning

	Winning party					
	D		R		Total	
Race	No.	%	No.	%	No.	%
Man vs man	3,721	65.6	3,260	68.6	6,981	66.9
Man vs woman, M	397	7.0	684	14.4	1,081	10.4
Man vs woman, W	559	9.9	255	5.4	814	7.8
Unopposed	894	15.8	499	10.5	1,393	13.4
Woman vs woman	103	1.8	56	1.2	159	1.5
Total	5,674	100.0	4,754	100.0	10,428	100.0

Notes: Races span from 1970 to 2016. Number or share of elections won by type of race (defined by sex of top candidates) and winning party. During this period, 12 races are won by an Independent, to which no gender was assigned. In this respect, note that *Unopposed* includes cases in which a D or R has defeated a third-party opponent, in addition to actual one-candidate races. Instances in which candidates from the same party occupy both first and second place are also considered.

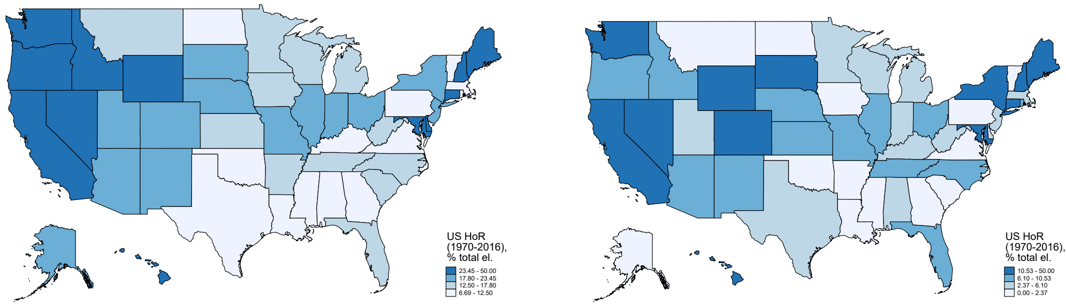
men, which are more likely to elect a male Republican.

Given the correlation between political affiliation and geography in the U.S. (e.g., southern states leaning towards the Republican party), Figure 1 reports the geographical distribution of mixed-gender races. For each state, Panel (a) shows the overall number of mixed-gender races for the House held between 1970 and 2016 as a percentage of the total number of elections held during the same period, aggregated at the state level as districts' boundaries and assignment to a given state may change over time. Panel (b) depicts the share of mixed-gender races that were won by women. As we can see, coastal states tend to have more mixed-gender races. Interestingly, some of the more central states have a higher share of mixed-gender races won by females. This would suggest that, even though these states are less likely to hold mixed-gender races, it is more likely that the female candidate wins whenever they occur.

We match the elections data with information on the CD in which the races were held. Specifically, we use the files provided by Foster-Molina (2017) to retrieve district characteristics for the period 1972-2016 (93rd to 115th U.S. Congress). Since the original data span from 1972 to 2013 (93rd to 113th U.S. Congress), we expanded the dataset by inputting data from the 113th Congress to the 114th and 115th Congresses. The district characteristics we consider are: the percentage of old age citizens; the percentage of black or Hispanic population; the percentage of people with a high school degree and with a bachelor degree; the unemployment rate; the median family income; and the Gini index of the district.[4]

the Republican party (327).

[4]The cutoff to determine "old age" share in the district is defined in Foster-Molina (2017) as 60 or 62 depending on the age brackets used by the Census for that decade.



(a) Total **(b) Mixed-gender races won by women**

Notes: Data on elections from 1970-2016, source: CQ Press. Panel A: share of mixed-gender races, by state. Panel B: share of mixed-gender races won by women, by state. Both shares are computed with respect to the total number of elections occurring in a given state.

Figure 1 Mixed-Gender Races, % of Total Number of Elections

3.2 Crime data

Data on crime come from the Uniform Crime Report of the FBI, henceforth denoted as UCR, a voluntary program under which law enforcement agencies (LEA) provide high-frequency reports on crime statistics corresponding to their jurisdictions. Despite being voluntary and, possibly, not covering the universe of all active LEAs in the United States, the sample of agencies included in the UCR is substantial. Recent statistics from the U.S. Department of Justice, for example, highlight that agencies in the UCR provide service to approximately 94% of the U.S. population (see also Appendix A). In particular, we consider the Offenses Known and Clearances by Arrest (OKCA) files for the period 1970-2017 and the Supplementary Homicide Report (SHR) files for the period 1976-2017, by Kaplan (2019*b*) and Kaplan (2019*a*).

The OKCA provide data at the month-agency level for several felonies. These statistics capture information such as the number of victims and whether the report was cleared, i.e., considered closed or solved for crime reporting purposes. These records do not include information on the relationship between the victim and the perpetrator. From these files, we retrieve the number of rape victims (including both forcible and attempted rapes) between 1970 and 2012. Throughout this period, that crime was explicitly defined as "carnal knowledge of a female forcibly against her will" by a male (FBI, 2013).[5] Other felonies like assault, robbery, and motor vehicle theft are also included in these files.

The SHR is an incident-based dataset providing detailed information on single cases of homicide, such as the month in which it occurred and the sex of both victim(s) and offender(s), whenever known. Since not all killings of women constitute a gender-based crime, we restrict our attention to murders and non-negligent manslaughters in which all victims are females and all offenders are males. This yields a dataset of 97,705 so-defined femicides occurring between 1976 and 2017. We

[5]This definition was expanded in 2013, after which i) the term forcible was dropped, and ii) victims and perpetrators of both genders were considered. See the FBI website at ucr.fbi.gov for more details.

further focus on cases in which the victim-offender relationship was one of the following: ex-wife, wife, common-law wife, girlfriend, daughter, mother, sister, stepdaughter, or stepmother. This restriction yields 55,263 femicides.[6] A brief description of this data is provided in Appendix B.

The UCR poses at least three challenges: i) data might not be representative, ii) there are possibly missing observations and other anomalies in the reporting, and iii) assumptions need to be made regarding how to appropriately measure crime at the congressional district level. Appendix A discusses in detail how we deal with these three issues.

In order to have a more precise measurement of crime at the congressional district level, we restrict our attention to local law enforcement agencies, such as a city police department or a Sheriff's office. Moreover, since there is great variation in the frequency of reporting, we aggregate the data at the year-district level. In order to do so, we first augment these files with information on the geographical location of the agencies. Using the addresses listed in the OKCA, we recover their spatial coordinates and determine the district in which an agency was located in each year, using district maps from Lewis et al. (2013) and the Census Bureau. We then compute the number of victims of femicide and of rape reported by every agency a in district d during year t, and normalize them by the population size of the district in year t (per 100,000 residents), retrieved from Adler (2003) and the Census Bureau. A similar procedure is followed to measure other crimes (e.g., assaults, burglaries) and clearances in a given district. Crimes against individuals record the number of victims, while crimes against property (e.g., burglaries or motor vehicle thefts) record the number of occurrences. In addition, we also consider the number of victims of femicides and rapes in terms of female population[7].

Crime data are linked to the race happening in district d in year t as follows: we merge all crime rates reported in year $t+1$, as well as those reported over the full elected term (i.e., the sum of crimes reported between $t+1$ and $t+2$). We deemed this appropriate as winning candidates of a November election are officially sworn into Congress, and thus start their term, only the following January. Focusing on the first year after election and on the cumulative sum for the two years after, our measures will in effect count crimes for the first year in office, and for the overall term.

A final note pertains to the issue of congressional districts boundaries, which can be redrawn from one session of Congress to another. This implies that a district in year $t+2$, despite maintaining the same name (e.g., Texas 13th CD) might cover a different area than in year t. While using yearly maps helps us to keep track of the different districts where agencies are located, the change

[6]It is important to note that the SHR contains information on the relationship between what is registered as *first* victim and the offender(s), where up to 11 offenders can be listed. However, the vast majority of cases involve only one victim. See also Appendix B.

[7]Female of voting age population from Lublin (1997) and from the Census Bureau is used. For femicides, we also considered the number of cases, normalized with respect to total population. Results are substantially unchanged (available upon request).

in their shapes might still generate inconsistencies when aggregating for longer periods. Adopting a more conservative approach, we will use the geographical boundaries from Lewis et al. (2013), rather than its denomination, to define districts. In other words, we will consider two areas to be the same CD only if they have the same boundaries. If a district changes boundaries between elections, we are still able to merge the corresponding crime data for the following year, $t+1$, to our elections data, but we are unable to compute the overall level of crime registered, since new boundaries have emerged in $t+2$.[8]

We refer the reader to Appendix A for a more detailed explanation of the aggregation of our crime data, and of the geo-localization procedure followed.

4 Empirical Strategy

4.1 Regression Discontinuity Design

We estimate the causal effect of female representation using a regression discontinuity (RD) design in a sample of close mixed-gender races. Our identification strategy relies on the assumption that, in close races between a man and a woman, it is mostly random factors that tilt the scale in favor of one of the candidates. Thus, a district in which a man wins against a woman with a narrow margin of victory becomes a good counterfactual for districts in which the opposite result was obtained. In our setting, close mixed-gender races are defined as those in which the top two places are occupied by a male (R or D) and a female (R or D), irrespective of the total number of candidates.

We first compute the female margin of votes, i.e., the difference between the votes received by the female and the male candidate as share of total votes, for each race taking place in district d, state s, and year t:

$$MV_{d,s,t} = (\frac{F_{d,s,t} - M_{d,s,t}}{Votes_{d,s,t}}) \times 100$$

We define our assignment to treatment as $W_{d,s,t} = 1[MV_{d,s,t} > 0]$, an indicator variable which takes the value of one if the winner of the (mixed-gender) race is female. The effect of having a woman in power can, then, be estimated with the local polynomial of p-th order:

$$y_{d,s,t} = \gamma_s + \lambda_t + \sum_{k=0}^{p} \eta_k MV_{d,s,t}^k + W_{d,s,t} \sum_{k=0}^{p} \beta_k MV_{d,s,t}^k + \varepsilon_{d,s,t} \tag{1}$$

in the sample of races in which the female margin of votes lies within the interval $[-h, +h]$. In the

[8] To have a sense of the loss in terms of observations, out of the 1,502 districts undergoing mixed-gender elections between 1970 and 2010, 398 report changing boundaries in the following election and missing the overall count for rape victims. Out of the 1,679 districts undergoing mixed-gender elections between 1976 and 2014, 372 report having changed boundaries in the following election and missing the overall count for femicide victims.

equation above, $y_{d,s,t}$ is our measure of gender-based crimes for district d in year t, while γ and λ denote state and year fixed effects. Our coefficient of interest is β, which identifies the effect of electing a female representative around the threshold, i.e., a *local* effect.

Following Imbens and Lemieux (2008) and similar to Broockman (2014), Brollo and Troiano (2016), our main and preferred specification is a non-parametric local linear regression. Specifically, as suggested by Cattaneo et al. (2020), we estimate a polynomial of order 1, choosing the bandwidth h optimally according to the procedure developed by Calonico et al. (2014*b*) and Calonico et al. (2014*a*). As a baseline, we also present the results of an OLS specification estimated in the sample of all mixed-gender races, where W is the main explanatory variable. Since one district can contribute with multiple mixed-gender elections over the years, standard errors are always clustered at the district level.

The use of close races to isolate the effect of an elected politician characteristic is widespread in the literature, but, unless additional strong assumptions are invoked, these estimates may be asymptotically biased (Marshall, 2022). In our setting, isolating the effect of gender would require that the gender of the candidate does not affect vote shares or that other predetermined potential confounders (e.g., competence, or social preferences) do not affect the outcome of interest. While it is hard to argue that gender does not affect the margins of victory, we argue that voters in the United States do not particularly care about gender-based violence. The Pew Research Center cyclically conduct a study to identify the top voting issues in the U.S. congressional elections. Their instrument does not even include as an alternative domestic violence or gender issues in general. Even after the increased visibility of domestic violence during the COVID-19 pandemic, other recent polls fail to include these topics when inquiring about what they consider to be the most important issues in their vote for Congress. Data from ANES (2019*b*) also seem to support this claim. When asked to identify the most important problem facing the United States, overall, only 13 percent of the respondents surveyed between 1970 and 2000 mention ”Public Order,” a category which includes crime, drugs, women's rights, and gun control, among other issues.

We acknowledge that the assumption that compensating differentials induced by variation in gender do not affect the outcome of interest may fail. Thus, we interpret our estimates of β as a compound treatment effect that incorporates the effects of gender as well as those of all compensating differentials induced or altered by conditioning on close races (Marshall, 2022). We describe β as the causal effect of electing a female representative as opposed as the causal effect of gender. Section 6 indeed shows that female candidates in close races greatly differ from their male counterparts in terms of their party affiliation. We thus test if a similar RD design to evaluate the effect of party affiliation explains the results on femicides and rapes (see Section 6, and Table C8). Reassuringly, this placebo test yields null impacts on both violence indicators.

Table 2 Balance Checks: Districts' Characteristics

Panel A: district characteristics	%Old	%Black	%Hisp	%HighSchool	%BA	%Unemp	Median Inc	Gini	Population
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Coeff	-0.364	-0.844	0.238	-0.347	0.237	0.077	-0.558	-0.003	-0.079
p-value	0.542	0.521	0.845	0.808	0.799	0.788	0.9	0.569	0.302
h	25	20	22	27	27	21	29	26	24
N	769	658	715	896	878	685	866	875	808

Panel B: macro-region	Midwest	Northeast	South	West
	(1)	(2)	(3)	(4)
Coeff	-0.125	0.035	-0.011	0.094
p-value	0.133	0.529	0.898	0.438
h	28	23	28	25
N	916	790	938	839

Notes: Results from a local linear polynomial specification. MSE optimal bandwidths and a triangular kernel are used. Robust p-values are reported. N denotes the total number of observations. Time fixed effects are included, standard errors clustered at the district level. Population is expressed in 100,000 residents, median income in thousands dollars.

4.2 Validity Tests

In the RD design just described, identification rests on the assumption that, around the threshold $MV = 0$, randomness determines the election result. One way to support this assumption is to show that all relevant factors which might affect the probability of electing either one of the candidates are not significantly different around that threshold, making two districts represented by Congresspersons of different genders comparable. To test for this assumption, we use data from Foster-Molina (2017) to look for the presence of discontinuities for a wide range of district characteristics, using the specification in equation (1) for each attribute. Specifically, we consider the percentage of people defined as old age in the district ($\%Old$), the percentage of black ($\%Black$) and Hispanic ($\%Hisp$) population, the percentage of people with a high school degree ($\%HS$) and with a bachelor's degree ($\%BA$), and the percentage of unemployed ($\%Unem$), as well as the median family income ($MedInc$), the Gini index ($Gini$) and the population size ($Population$).

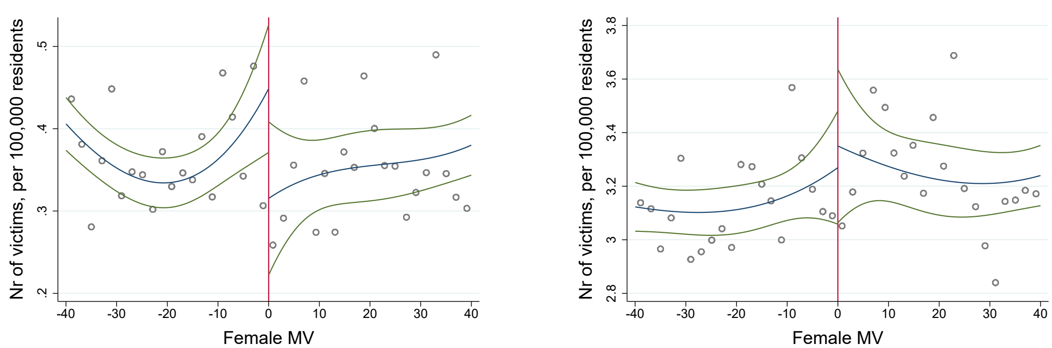
Results are reported in Panel A of Table 2. As we can see, none of the coefficients are statistically significant, supporting the comparability of CDs around the cutoff. To control for possible differences related to the geographical location of the district, we also check for differences for the macro-region in which the race took place. Results are reported in Panel B of Table 2. Again, the probability of electing either one of the candidates is not significantly different across different regions.

Another important assumption on which the design hinges is that candidates are not able to control the probability of falling on either side of the threshold, which in our setting translates into being able to manipulate the outcome of an election. In other words, the validity of our identification strategy relies on the continuity of the female margin of votes around this threshold. Hence, we first check for possible spikes in the distribution of MV by plotting its frequency. Then, more formally, we implement a McCrary test (McCrary, 2008). Both plots and test are reported in Figure D1 and Figure D2 in Appendix D. We also perform an additional discontinuity test as suggested in Cattaneo et al. (2018*a*) and Cattaneo et al. (2019), and report it in Figure D3. In all, we fail to reject the null hypothesis of the continuity of that distribution.

5 Results

5.1 Violence against women

Figure 2 provides a graphical analysis of the local impact of electing a female representative on femicides and rapes after one year in office, plotting average values – averaged within bins of 2% of the female margin of votes – against MV, together with fitted values from a third order polynomial in MV (Brollo and Troiano, 2016). As we can see, female political representation seems to decrease the number of victims of femicides (per 100,000 residents), but it does not seem to affect the prevalence of rape.



(a) Femicides **(b) Rapes**

Legislator's gender and gender violence: Measures on the y-axis expressed as number of victims per 100,000 residents. *Female MV* is the margin of votes of the female candidate in district d in year t. $MV > 0$ when the winning candidate is female and $MV < 0$ when male. Scatter points represents values averaged within bins of 2% of MV. The blue line represents fitted values from a third-order polynomial in the margin of victory MV, fitted separately on each side of the threshold, while the green lines are the 95% confidence interval of that polynomial. Values refer to the first year of the term.

Figure 2 Gender-Based Crime and Margin of Victory

This initial evidence is confirmed by the results in Panel A of Table 3, which looks at femicides in Columns (1)-(4) and rapes in Columns (5)-(8). Both measures are normalized by population and expressed in logs. In Columns (1) and (5) we show simple correlations as a baseline. To estimate these coefficients, we restrict the sample to mixed-gender races and include as regressors the indicator variable W for female winners, time fixed effects, state fixed effects, and a set of district-specific covariates.[9] This set includes macro-area dummies ($West$, $Northeast$, $Midwest$, and an omitted category $South$), a dummy for whether the winner of the election belonged to the Democratic Party (*D*), the percentage of black (*PctBlack*) and Hispanic (*PctHisp*) population in the district, the unemployment rate (*PctUnemp*), and the percentage of agencies reporting data to the FBI.[10] In the following columns, we adopt a robust non-parametric approach. As explained in Section 4.1, we consider a first order polynomial in MV over the sample of close mixed-gender races, optimally selecting the bandwidth h following Calonico et al. (2014*b*) and Calonico et al. (2014*a*).

The results are reported in Columns (2) and (6), which include both time and state fixed effect as well as the set of district-specific covariates just mentioned. As a robustness check, Columns (3) and (7) replicate the same analysis, manually setting the bandwidth to be 1.5 as large as the optimal. Similarly, Columns (4) and (8) set the bandwidth to half of the optimal.

The results show that, during the first year in office, the prevalence of femicides in districts electing a Congresswoman is significantly lower with respect to those in which a man won. According to our preferred specification in Column (2), femicides decrease by 9.5 percent in the short-run. We obtain negative results of a similar magnitude when we manually modify the bandwidths (also significant in Column (3)). At the same time, reported rapes do not seem to be affected by the election of a female candidate in the first year of her term. The fact that we find a decrease in femicides, our measure of gender-based violence that is less subject to misreporting, signals an actual decrease in the level of violence. In turn, the lack of an impact on rapes may be due to the interplay between reporting and deterrence effects. One hypothesis is that rapes go down, but at the same time women become more likely to report, muting the recorded impact on rapes. In Section 5.3 we provide evidence in support of the latter.

Panel B of Table 3 repeats the analysis, but focuses on the cumulative effect on the number of victims of femicides and of rapes during the overall term. The results show that the impact on femicides does not persist. Even though the sign of the coefficients still suggests a decrease, the coefficients are not significant in any of the specifications. In turn, the prevalence of rapes significantly increases by about 25 percent in our preferred specification (Column (6)) during the overall term of a Congresswoman, relative to that of a Congressman. While the size of this coeffi-

[9] While we do not have data on district characteristics for 1970, we only see 21 mixed-gender elections in our sample for that year. To have a better sense of the loss in terms of number of observations, only in seven of those elections is the female margin of victory below 15 percent (in absolute value).

[10] The measures for reporting agencies vary depending on the outcome considered: for rapes, the percentage of agencies in the district reporting non-missing OKCA data is used; for femicides, the percentage of agencies in the district reporting non-missing SHR data is used. Shares are computed with respect to the number of agencies listed in the UCR. For the first-year measures, we consider agencies reporting during the first year in office. For the overall period, we use the average of the yearly shares for the first and second year of the term.

Table 3 The Effect of Female Representation on Femicides and Rapes

Panel A: First year								
	Femicides				**Rapes**			
	OLS	RDD			OLS	RDD		
Bandwidth:		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coeff	-0.030	-0.100	-0.118	-0.088	0.034	0.035	0.011	-0.126
p-value	0.149	0.062	0.063	0.155	0.594	0.769	0.728	0.507
h	100	19	29	10	100	14	20	7
N	1515	518	783	231	1358	308	486	159
Avg Outcome	0.393				3.212			
Panel B: Overall term								
	Femicides				**Rapes**			
	OLS	RDD			OLS	RDD		
Bandwidth:		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coeff	-0.022	-0.011	-0.038	-0.018	0.077	0.238	0.077	0.465
p-value	0.462	0.979	0.982	0.636	0.298	0.093	0.22	0.099
h	100	15	22	7	100	22	33	11
N	1069	265	426	129	973	378	575	170
Avg Outcome	0.653				3.887			

Notes: Results from a local linear polynomial specification. Dependent variables expressed as logs of number of victims per 100,000 residents. Values refer to the first year of the term (Panel A) and to the overall term (Panel B). All regressions include a set of controls as well as time and state fixed effects. Controls include an indicator variable for each geographical macro-area, an indicator for Democratic winner D, percentage of black population, percentage of Hispanic population, unemployment rate, and percentage of agencies reporting non-missing data (OKCA or SHR). Standard errors clustered at the district level. For the RDD specification, MSE-optimal bandwidths and a triangular kernel are used, and robust p-values are reported. *Avg Outcome* is computed across all districts undergoing mixed-gender races.

cient varies considerably across specifications, the sign remains robust, confirming an increase in recorded rapes.

The magnitude of the effect on rapes is comparable to Iyer et al. (2012), who find an increase in reported rapes of 23 percent after the implementation of female quotas in local councils in India. Our results can also be compared to the effects of opening woman-run police stations in India: Amaral et al. (2021) estimate that the introduction of female-run stations is associated with a 29 percent increase in police reports of crimes against women.

All in all, lack of persistence of the effect on femicides suggests that the impact on the actual level of violence is short-lived. The trend observed in rapes during the first year and the overall term, in turn, shows that the muted impact early on is probably the result of opposing effects coming from deterrence and reporting. During the full mandate, the reporting effect persists. Importantly, we do not find a wider change in the level of crime in the district, either in the short or medium run (see Table C2), underscoring that the persistence of reporting effects is only found for gender-based crimes.

These results can be driven both by a *policy effect* and a *role model effect*. On one hand, a female Representative might influence the level of gender-based crime in a given district by enacting specific pieces of legislation. Considering how policy preferences differ across men and women, we could expect, for example, a female legislator to be more likely than a male legislator to propose and pass bills that deter violence against women. On the other hand, increased reporting and a contraction of femicides can also come from the mere exposure to such a leader: female politicians can shape aspirations, attitudes, and opinions, in particular for individuals with shared backgrounds and traits (Priyanka, 2020). Their election can thus foster changes in gender-biased social norms, and their public persona may become a role model for other women in the reference district. Women who are aware of having a female local representative in a position of power may feel more empowered and become more likely to both abandon abusive relationships and to come forward when facing domestic violence or other types of gender-based abuse. Female politicians can also influence the effort exerted by the police in dealing with gender-based crimes either directly (e.g., by pressuring them in local or national media) or indirectly (e.g., gender-biased stereotypes and beliefs held by a local police chief may change when he/she becomes aware of a local female in a position of power).

We argue that the short-run effect identified on actual violence and the sustained effect on reporting is unlikely to be due to a particular policy change introduced by the female legislator regarding either specific or general crimes. First, we do not find an effect on overall crime (see also the discussion in Section 6). Second, any such policy or reform is likely to take time to be implemented and yield results. If it were policies driving the decrease in gender-based crimes, we would expect to see a delayed effect and capture it during the overall term. Third, the legislation enacted by Congress has an effect at the federal level. While we cannot completely rule out the role of policy action as a channel, it is definitely harder to envision members of Congress having such a local impact at the district level through legislation. Hence, we argue that the short-lived reduction in violence and increase in reporting mostly comes from female leaders' visibility and, thus, their indirect influence at the district level, i.e., *a role model effect*. Data from ANES (2019*b*) suggest that U.S. citizens are indeed aware of the district political developments and of the identity of their candidates and Representatives. On average, 58 percent of all respondents says they care "very much/pretty much" about the outcome of the election for the U.S. House. Furthermore, 52 percent of them report having voted in the elections for the U.S. House in their districts. This would suggest that, since they cast their ballots, they have at least basic knowledge of the candidates and are aware of the presence of women on the ballots. Indeed, 37 percent of the respondents in districts undergoing mixed-gender races are able to correctly identify at least one of the candidates by name.

5.2 Police effort: clearances

Police are usually front-line agents when reporting and dealing with violent crimes, including gender-based ones. Elected Congresswomen may enact policy changes affecting their funding or their span of action, but they can also exert indirect pressure on their performance. As public figures, they may denounce or praise the district's police work in certain arenas and affect their effort level for specific crimes. Iyer et al. (2012) document, for instance, that in India the number of arrests for crimes against women significantly increases after the implementation of a law requiring the reservation of specific political offices for women, but they do not find such an increase for other felonies.

Columns (1)-(7) in Panel A of Table 4 below report the effect of electing a woman on the rate of clearances of different types of crimes, using our preferred specification, i.e., Column (2) of Table 3. The outcome variables are the (log) number of clearances per 100,000 district residents registered during the first year in office.[11] In terms of crimes against women, since the OKCA data does not provide information on the sex of either the victim or the offender, we focus on clearances for rapes, in Column (2). The results show that, during a Congresswoman's first year in office, police effort when pursuing rapes or other crimes remains unchanged. Using the information provided by the SHR, we also obtain the number of victims of unsolved murders as an additional

[11] This measure includes clearances both by arrests and by exceptional means. One clearance by arrest is counted for each offense meeting these three criteria: at least one person must have been i) arrested, ii) charged, and iii) turned over to the court for prosecution. Instead, exceptional means clearances refer to those situations where elements beyond law enforcement's control prevent the agency from arresting and formally charging the offender, e.g., death of the offender (see FBI at ucr.fbi.org). Also notice that, in the case of crimes against individuals, every victim counts as a separate offense (see FBI (2013)).

indicator for police effort. Specifically, we consider a case as unsolved whenever the gender of the offender(s) is unknown – since we can assume that this information is present only when the perpetrator is apprehended (see also Fox (2004)). If officers become more responsive to violence against women, we should see a decrease in killings of females going unsolved, but not in that of males. Hence, in Columns (8) and (9) we investigate the impact of the election of a female politician on the (log) number of females and males whose killings remain unsolved, normalized by population in both cases.

The results show that, in the short run, the number of female victims of unsolved murders significantly decreases by 7.5 percent. There is no such effect for cases involving male victims. Naturally, we can only interpret this effect as increased police effort when addressing gender-based crimes if the drop in the clearance rate of female murders is not entirely driven by the drop in the number of cases identified in Table 3. However, we find no evidence of a significant change in the total number of women killed or, interestingly, in that of women killed by men overall (Table C1).

While we do not find that the police are increasingly effective in dealing with rapes in the short run, Panel B of Table 4 suggests that higher police effort is exerted to deal with these crimes during the full term. According to Column (2), there is an increase of about 50 percent in rape cases cleared, suggesting that as reported rapes overall go up, so does the effort exerted by the officers in pursuing those cases. Noticeably, the number of clearances for murders also seems to go up, as can be seen in Column (1), while no significant change is observed for clearances regarding other crimes.

To corroborate these findings, in Table C3 in the Appendix, we look at arrests, that is the (log) number of people arrested for a given felony, normalized by population.[12] The results follow the same pattern: no effect in the short run, but a significant increase in the number of people arrested for murder and rape during the full term of about 30 percent. More importantly, when we break down the data by gender, we see that these results are driven by males (Table C4). We find no change in the number of females being arrested (results available upon request). The fact that we find an increase in the number of males arrested for both rapes and murders suggests particular police attention towards more violent crimes perpetrated by men, more broadly.

Note that our overall effect on rapes is around 26 percent, which would suggest that officers increase their effort at a rate similar to that at which the reporting of rapes increases. Our results are also aligned with Iyer et al. (2012), who find an increase of nearly 30 percent in the number of arrests for crimes against women after women's reservations are implemented.

[12] Data come from a separate file: Arrests by Age, Sex and Race (ASR), provided by Kaplan (2021). This file provides the number or people arrested, broken down by either age, sex, or race, at the month-agency level. To aggregate this data, the same procedure as for the OKCA and SHR is followed. In this case, lacking alternative more precise methods to identify missing observations, values are set to missing if every reported element is a zero. Hence, caution should be taken in interpreting these results, which might be an upper bound of the true effect.

Table 4 The Effect of Female Representation on Police Effort

Panel A: First year

	Clearances							Murders, unsolved	
	Murder (1)	Rape (2)	Aggr Ass (3)	Robbery (4)	Burglary (5)	Larceny (6)	Mv Theft (7)	Women (8)	Men (9)
Coeff	0.04	0.041	-0.026	-0.012	-0.058	-0.056	0.117	-0.078	-0.031
p-value	0.599	0.855	0.952	0.953	0.745	0.671	0.431	0.06	0.775
h	11	14	15	14	17	18	14	14	21
N	292	327	428	382	503	556	405	217	515
Avg Outcome	1.302	2.446	4.558	3.204	4.407	5.887	3.609	0.327	0.724

Panel B: Overall term

	Clearances							Murders, unsolved	
	Murder (1)	Rape (2)	Aggr Ass (3)	Robbery (4)	Burglary (5)	Larceny (6)	Mv Theft (7)	Women (8)	Men (9)
Coeff	0.24	0.435	0.207	0.195	0.112	0.192	0.207	0.123	0.071
p-value	0.07	0.004	0.267	0.254	0.386	0.105	0.13	0.347	0.75
h	14	15	16	17	15	14	23	15	16
N	284	244	339	376	325	285	517	162	252
Avg Outcome	1.797	3.094	5.262	3.841	5.107	6.607	4.303	0.55	1.076

Notes: Results from a local linear polynomial specification. MSE-optimal bandwidths and a triangular kernel are used. Dependent variables expressed as logs of number of clearances per 100,000 residents. Values refer to the first year (Panel A) and the overall term (Panel B). All regressions include a set of controls as well as time and state fixed effects. Controls include an indicator variable for each geographical macro-area, an indicator for Democratic winner D, percentage of black population, percentage of Hispanic population, unemployment rate, and percentage of agencies reporting non-missing data (OKCA or SHR). Standard errors clustered at the district level. Robust p-values are reported. *Murders unsolved, women* is computed by counting female victims of murder whose offenders were (all) unknown. We consider an offender as such if the sex was *unknown* (see Fox (2004)). *Murders unsolved, men* counts male victims of unsolved murders. *Avg Outcome* is computed across all districts undergoing mixed-gender races.

Overall, we show that the election of a female representative leads to greater police responsiveness in terms of their solving of gender-based crimes during the first year in office, particularly female homicides. This effect may have a *deterrence effect* on males, explaining the short-run decrease in femicides. Moreover, we also identify greater effectiveness in dealing with reported rapes during the full term, which would support the hypothesis of more reporting by women.

5.3 Females and males' attitudes

The nature of the impact of female leadership on femicides and rapes suggests that the channel for the reported dynamics is not policy intervention, but rather an indirect influence on gender-biased social norms. Both men and women may change their perceptions about gender equality when a female from the district becomes a visible political leader. On one hand, men may be less likely to commit gender-based crimes as a result of more positive views about women in society. On the other hand, women may become more empowered and thus be more likely to stand up to violence and to increase their rate of reporting when they are victimized.

To shed light on these mechanisms, we retrieve data from the Time Series Cumulative Data Files of the American National Election Study (ANES) series for the period 1970-2016 (ANES, 2019*b*), and from individual studies for the period 2004-2016 (ANES, 2005, 2015, 2016, 2019*a*). ANES is a nationally representative study conducted in the United States approximately every election year, around the time of the elections. Surveys can take place both before and after an election is held. In our analysis, we only use the information whenever it is collected post-election. The final dataset allow us to estimate effects on attitudes in the very short run as the outcome variables are measured, at the earliest, on the day of the election and, at the latest, around 90 days after the election.

There are several limitations with regards to these series. First, they are not representative at the district level. Yet, if sampling is orthogonal to specific characteristics of the survey respondents, the results of our regressions should continue to be unbiased. Second, the survey instrument is not consistent over time, which generates gaps in the series for certain questions. Nevertheless, ANES remains a valuable source of information, as it enquires about a vast array of issues, not exclusively political.

In particular, in 2004 and 2008, survey respondents were asked to give their opinion about the statement "women who complain about harassment cause more problems than they solve." Similarly, data from 2012 and 2016 provide us with individual responses on the statement "women who complain about discrimination cause more problems than they solve." Both opinions are collected using Likert-type questions. Although these questions do not specifically relate harassment or discrimination to the use of violence against women, they are still a good indicator of how women perceive reporting. Based on these data, we define two indicator variables: *Harassment* and *Complain*. *Harassment* is equal to one if the individual disagrees with the statement on harassment. *Complain* is equal to one if the individual disagrees with either of the two statements, allowing us to expand our data to more recent years. To estimate the impact of female leadership on these views, we rely on local randomization methods (Cattaneo et al., 2016, 2018*b*), which we deem more appropriate to deal with the limited sample size and the masspoints we identify in the ANES data (Cattaneo et al., 2017). Columns (1) to (4) in Table 5, present the results by sex of the respondent and for two different specifications: no transformation and a first order transformation.

Table 5 The Effect of Female Representation on Views and Attitudes towards Reporting and the Police

Panel A: Females						
	Harassment		Complain		Police	
	(1)	(2)	(3)	(4)	(5)	(6)
Coeff	0.247	0.406	0.06	0.136	0.068	0.464
p-value	0.149	0.007	0.488	0.027	0.76	0
Order	0	1	0	1	0	1
lWindow	9.439	9.439	8.11	8.11	3.997	3.997
rWindow	22.309	22.309	6.557	6.557	4.209	4.209
N-	21	21	64	64	16	16
N+	31	31	65	65	35	35
Panel B: Males						
	Harassment		Complain		Police	
	(1)	(2)	(3)	(4)	(5)	(6)
Coeff	0.147	0.213	0.074	-0.075	-0.004	-0.004
p-value	0.418	0.162	0.373	0.239	1	0.811
Order	0	1	0	1	0	1
lWindow	3.783	3.783	4.093	4.093	13.682	13.682
rWindow	26.478	26.478	3.765	3.765	13.917	13.917
N-	23	23	56	56	77	77
N+	32	32	55	55	94	94

Notes: Dependent variables are indicator variables of the responses, as described in Section 5.3. Results from a local randomization (rdrandinf). Individual characteristics used for balance-test: black, Hispanic, age, Democratic. Order refers to the order of the transformation used. $N+$ refers to the number of observations above the threshold MV=0, $N-$ to those below. lWindow and rWindow are the left-and-right limit of the interval considered. Values for the difference in mean, and p-values are reported.

Our estimates show that the election of a female representative has a local positive effect on women's views about reporting. Columns (2) and (4) in Panel A in Table 5 provide, in fact, supporting evidence of women becoming more likely to hold positive views about complaining and denouncing harassment. Even though the effects are not significant when there is no transformation (Columns (1) and (3)), the estimated coefficients go in the same direction. In turn, Panel B shows that the election of a Congresswoman has no impact on men's views about women who decide to report harassment and discrimination.

The ANES Cumulative Data Files also allow us to measure the effect of elected female leaders on views about the police, asking respondents to rate their feelings towards police on a scale from 0 to 100, where higher values are associated with more favorable and warmer feelings toward them. This question is asked after the elections in the following years: 1970, 1972, 1974, 1976, 1992 and 2016. We construct an indicator variable that is equal to one if the respondent rated her feelings above 50. Column (6) in Panel A shows that women do view the police more favorably after a female leader is elected. Once more, Panel B shows that there is no such effect among men[13].

Overall, the improvement in the attitudes of women towards reporting harassment and discrimination and in the way they perceive police provides some evidence in support of increased reporting by women. At the same time, the absence of an effect on men goes against the hypothesis that the documented short-run decrease in violence is due to more favorable views about women on their part. Instead, it reinforces the idea of a *deterrence effect* exerted by more police effort and more reporting.

5.4 Other Channels

We have already ruled out that the effect on gender-based violence is coming from an overall decrease in crime (see Table C2). However, our results may be partially driven by elections (or appointments) for other offices linked to crime. For instance, both a female judge and a female Representative could be elected at the same time in the same district. In this case, the effect we detect could be a combination of more women in the Congress and in the judicial system. Similar considerations might apply when looking at elections for offices in the law-enforcement system, such as those for sheriffs.

Studying the selection process of state trial court judges, U.S. attorneys (i.e., federal prosecutors), district attorneys (DAs, i.e., state prosecutors), and sheriffs, however, leads us to rule out that other elections confound the effects of electing a female representative. First, the elections of state trial court judges, U.S. attorneys and DAs do not perfectly coincide with that of a Congressperson, in terms of the length of their terms and/or their geographical jurisdictions. Typically, their terms are longer than those for the U.S. House, and the area over which they can exert their authorities is different (larger) than that of a CD. Furthermore, the holders of such offices are not always elected but appointed by either the state's Governor or the U.S. President. These appointments

[13] These results are robust to an alternative specification of the dependent variable as standardized values of the answers (results available upon request)

may be correlated with voters' preferences, but within a larger jurisdiction such as the state or the nation. Second, even though sheriffs preside over a geographical area that likely overlaps with or is contained by a given CD (i.e., a county, see also Appendix A.3), and even though they are elected, in the overwhelming majority of cases their terms extend past that of an elected Representative. Third, we note that many DA and Sheriff candidates run unopposed and are overwhelmingly (white) men. This would indicate that elections are hardly ever competitive and mostly dominated by males, even when competition is present.

We refer the reader to Appendix E for a more detailed discussion of the selection process for these offices.

6 Robustness

In order to validate our findings on gender-related violence, we perform several robustness checks. First, we investigate how our results change when using alternative definitions of the outcome variables and the analysis sample. On one hand, we try an alternative normalization for femicides and rapes and divide these counts by the size of the female (voting) population, rather than by total population (see Table C5). On the other hand, we repeat our main regressions by expressing the number of victims by total population in levels, instead of considering a logarithmic transformation (see Table C6). In all of these cases, the results point to a decrease in femicides in the short run and to an increase in rapes over the whole term.

Another important test pertains to the sample of agencies considered which, in our main analysis, was limited to local law enforcement agencies only. We expand our dataset to include the whole sample of LEA (Table C7). When we do so, our main findings are qualitatively unchanged. We argue that the inclusion of non-local agencies (e.g., State Police) renders our data more noisy, because of the increased measurement error, which could explain the loss of significance in some of the specifications (see also Appendix A.4).

As mentioned at the end of Section 4.1, it is possible that by conditioning on politicians who win close elections, we are identifying the effect of the specific characteristic of interest as well as all compensating differentials (i.e., other candidate characteristics that lead to close elections between candidates who differ in their gender). For instance, previous studies suggest that women in politics tend to be more educated (Baltrunaite et al., 2014), better performers (Black and Erickson (2003), Bauer (2020)) and more effective in securing funding for their constituencies (Anzia and Berry, 2011) than males. The intuition behind these findings is that candidates can sort into races according to a multiplicity of factors. In particular, if women shy away from competition or expect to face discrimination on part of voters, only more qualified and more able women would run for election. Similarly, if voters are biased against a female candidate, only more qualified and more able women will be able to win an election (Anzia and Berry, 2011). Anticipating discrimination, parties might also strategically select candidates based on gender, e.g., by putting males in the more contestable districts (Le Barbanchon and Sauvagnat, 2021), by putting females in poorer positions on the ballot (Esteve-Volart and Bagues, 2012), or by choosing women only when they are sufficiently strong (Cella and Manzoni, 2019). At the same time, the belief of being held to a higher standard could make elected women harder workers.[14]

Unfortunately, we have limited information on candidates' characteristics, but we can compare party affiliation by sex in our sample of close mixed-gender elections. In general, female candidates in such races are more likely to belong to the Democratic party than male candidates. If we look at mixed-gender races where the margin of votes was below 5 percent (in absolute value), for example, we find that 68 percent of women candidates are Democrats, while this figure is only 32 percent among male candidates. Since party affiliation is likely to be an important compensating differential, we conduct a placebo test to examine if the effects on femicides and rapes are coming from electing a Democrat, rather than from electing a woman. We implement an RD design on "mixed-party races," i.e., all elections in which the top two candidates are either R or D. In this case, $MV_{d,s,t}$ is defined as the difference in the share of (total) votes between the Democratic and the Republican candidate. Similarly, the treatment variable is defined as $D_{d,s,t} = 1(MV_{d,s,t} > 0)$, indicating the victory of the Democratic candidate. Besides these differences, the econometric specifications used are equivalent to those in Table 3. Table C8 shows that there are no significant impacts of electing a Democrat on either short-run or medium-run effects.

7 Conclusion

Thousands of women and girls are facing insecurity and violence in the United States. We show that the election of a female politician reduces the incidence of gender-based violence in the United States. The drop is mainly due to a higher responsiveness and effort of the police in solving gender-related crimes after a female politician is elected. We argue that the police response induces a deterrence effect that decreases femicides by 10 percent in the first year of the politician's mandate. The election of a female Representative also encourages women to report abuse, increasing reported rapes significantly by 26 percent during her two years in office.

This paper builds on previous studies that have tried to measure the impact of electing a female political leader on gender-based crimes. Previous studies have focused on developing countries and have relied on the staggered enactment of quotas to identify the effect of female political leadership

[14]Higher ability and effort could help explain women's higher performance in the legislative process, e.g., more bills sponsored (Volden et al., 2013), and more speeches made on the House floor (Pearson and Dancey, 2011). See also Gagliarducci and Paserman (2021) for a discussion on *selection*.

on the prevalence of violence against women. We take advantage of close elections between male and female candidates and rely on rich FBI records to measure femicides and rapes. To the best of our knowledge, this is the first study that provides evidence on the importance and effectiveness of female representation in the United States in addressing violence against women. All in all, this paper is an important step in investigating the role of female political representation in developed economies in closing the gap on gender equality.

References

Adler, E. Scott. 2003. "Congressional District Data File, [92-105 Congress]." University of Colorado, Boulder, CO.
9

Agüero, Jorge, and Veronica Frisancho. 2022. "Measuring Violence against Women with Experimental Methods." *Economic Development and Cultural Change*.
2

Amaral, Sofia. 2017. "Do Improved Property Rights Decrease Violence Against Women in India?" *Working Paper*. Available at SSRN: https://ssrn.com/abstract=2504579 or https://dx.doi.org/10.2139/ssrn.2504579.
2, 4

Amaral, Sofia, Sonia Bhalotra, and Nishith Prakash. 2021. "Gender, Crime and Punishment: Evidence from Women Police Stations in India." *CESifo Working Paper No. 9002*. Available at SSRN: https://ssrn.com/abstract=3827615 or http://dx.doi.org/10.2139/ssrn.3827615.
2, 4, 16

ANES. 2005. "The American National Election Studies, 2004 Time Series Study."
20

ANES. 2015. "The American National Election Studies, 2008 Time Series Study."
20

ANES. 2016. "The American National Election Studies, 2012 Time Series Study."
20

ANES. 2019*a*. "The American National Election Studies, 2012 Time Series Study."
20

ANES. 2019*b*. "The American National Election Studies, Time Series Cumulative Data File (1948-2016)."
11, 17, 20

Anzia, Sarah F., and Christopher R. Berry. 2011. "The Jackie (and Jill) Robinson Effect: Why Do Congresswomen Outperform Congressmen?" *American Journal of Political Science*, 55(3): 478–493.
23, 24

Baltrunaite, Audinga, Piera Bello, Alessandra Casarico, and Paola Profeta. 2014. "Gender quotas and the quality of politicians." *Journal of Public Economics*, 118: 62 – 74.
23

Bauer, Nichole M. 2020. *The Qualifications Gap: Why Women Must Be Better than Men to Win Political Office.* Cambridge University Press.
23

Beaman, Lori, Esther Duflo, Rohini Pande, and Petia Topalova. 2012. "Female Leadership Raises Aspirations and Educational Attainment for Girls: A Policy Experiment in India." *Science*, 335(6068): 582–586.
3

Beaman, Lori, Raghabendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova. 2009. "Powerful Women: Does Exposure Reduce Bias?" *The Quarterly Journal of Economics*, 124(4): 1497–1540.

3, 4, 5

Bhalotra, Sonia, and Irma Clots-Figueras. 2014. "Health and the Political Agency of Women." *American Economic Journal: Economic Policy*, 6(2): 164–97.
3

Bhalotra, Sonia R, Damian Clarke, Joseph F Gomes, and Atheendar Venkataramani. 2022. "Maternal Mortality and Women's Political Power." National Bureau of Economic Research Working Paper 30103.
3

Black, J.H, and L Erickson. 2003. "Women candidates and voter bias: do women politicians need to be better?" *Electoral Studies*, 22(1): 81 – 100.
23
Brennan Center for Justice

Brennan Center for Justice. 2022. "Judicial Selection: An Interactive Map." Dataset. Available online from `https://brennancenter.org/judicial-selection-map` (last visited July 2022).
51, 52

Brollo, Fernanda, and Ugo Troiano. 2016. "What happens when a woman wins an election? Evidence from close races in Brazil." *Journal of Development Economics*, 122(C): 28–45.
3, 11, 13

Broockman, David E. 2014. "Do female politicians empower women to vote or run for office? A regression discontinuity approach." *Electoral Studies*, 34: 190 – 204.
11

Bruce, Raphael, Alexsandros Cavgias, Luis Meloni, and Mário Remígio. 2022. "Under pressure: Women's leadership during the COVID-19 crisis." *Journal of Development Economics*, 154: 102761.
3

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014*a*. "Robust data-driven inference in the regression-discontinuity design." *Stata Journal*, 14(4): 909–946(38).
11, 14

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014*b*. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica*, 82(6): 2295–2326.
11, 14

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2018*a*. "Manipulation Testing Based on Density Discontinuity." *The Stata Journal*, 18(1): 234–261.
13

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2019. "Simple Local Polynomial Density Estimators." *Journal of the American Statistical Association*, 0(0): 1–7.
13

Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik. 2018*b*. *A Practical Introduction to Regression Discontinuity Designs: Volume II. Prepared for: Elements in Quantitative and Computational Methods for the Social Sciences.*
21

Cattaneo, Matias D., Nicolás Idrobo, and Rocío Titiunik. 2020. *A Practical Introduction to Regression Discontinuity Designs: Foundations. Elements in Quantitative and Computational*

Methods for the Social Sciences, Cambridge University Press.
11

Cattaneo, Matias D., Rocío Titiunik, and Gonzalo Vazquez-Bare. 2017. "Comparing Inference Approaches for RD Designs: A Reexamination of the Effect of Head Start on Child Mortality." *Journal of Policy Analysis and Management*, 36(3): 643–681.
21

Cattaneo, Matias, Rocio Titiunik, and Gonzalo Vazquez-Bare. 2016. "Inference in regression discontinuity designs under local randomization." *Stata Journal*, 16(2): 331–367.
21

Cella, Michela, and Elena Manzoni. 2019. "Gender bias and women political performance." University of Milano-Bicocca, Department of Economics Working Papers 414.
24

Chattopadhyay, Raghabendra, and Esther Duflo. 2004. "Women as Policy Makers: Evidence from a Randomized Policy Experiment in India." *Econometrica*, 72(5): 1409–1443.
4

Clots-Figueras, Irma. 2011. "Women in politics: Evidence from the Indian States." *Journal of Public Economics*, 95(7): 664 – 690.
4

Clots-Figueras, Irma. 2012. "Are Female Leaders Good for Education? Evidence from India." *American Economic Journal: Applied Economics*, 4(1): 212–244.
3, 4

Coppolo, George Chief Attorney. 2003. "States That Elect Their Chief Prosecutors." Connecticut General Assembly. Report 2003-R-0231. Available online from `https://www.cga.ct.gov/2003/rpt/2003-R-0231.htm`.
52

Delaporte, Magdalena, and Francisco Pino. 2022. "Female Political Representation and Violence against Women: Evidence from Brazil." IZA Discussion Paper No. 15365.
4

Donohue, John J, III, and Steven D Levitt. 2001. "The Impact of Race on Policing and Arrests." *Journal of Law and Economics*, 44(2): 367–394.
5

Esteve-Volart, Berta, and Manuel Bagues. 2012. "Are women pawns in the political game? Evidence from elections to the Spanish Senate." *Journal of Public Economics*, 96(3): 387 – 399.
24

FBI, Criminal Justice Information Services Division. 2013. "Summary Reporting System (SRS) User Manual." U.S. Department of Justice - FBI.
8, 17, 34, 35, 42

Foster-Molina, Ella. 2017. "Historical Congressional Legislation and District Demographics 1972-2014."
7, 12

Fox, James Alan. 2004. "Missing Data Problems in the SHR: Imputing Offender and Relationship Characteristics." *Homicide Studies*, 8(3): 214–254.
18, 19, 33, 34

Gagliarducci, Stefano, and M Daniele Paserman. 2021. "Gender Differences in Cooperative En-

vironments? Evidence from The U.S. Congress." *The Economic Journal*, 132(641): 218–257.
24

Gerrity, Jessica C., Tracy Osborn, and Jeanette Morehouse Mendez. 2007. "Women and Representation: A Different View of the District?" *Politics & Gender*, 3(2): 179–200.
4

Hemenway, David, Tomoko Shinoda-Tagawa, and Matthew Miller. 2002. "Firearm availability and female homicide victimization rates among 25 populous high-income countries." *Journal of the American Medical Women's Association*, 57: 100–104.
2

Hidrobo, Melissa, Amber Peterman, and Lori Heise. 2016. "The Effect of Cash, Vouchers, and Food Transfers on Intimate Partner Violence: Evidence from a Randomized Experiment in Northern Ecuador." *American Economic Journal: Applied Economics*, 8(3): 284–303.
2, 4

Imbens, Guido, and Thomas Lemieux. 2008. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics*, 142(2): 615–635.
11

Iyer, Lakshmi, Anandi Mani, Prachi Mishra, and Petia Topalova. 2012. "The Power of Political Voice: Women's Political Representation and Crime in India." *American Economic Journal: Applied Economics*, 4(4): 165–193.
3, 4, 5, 16, 17, 18

Kaplan, Jacob. 2019*a*. "Jacob Kaplan's Concatenated Files: Uniform Crime Reporting (UCR) Program Data: Supplementary Homicide Reports, 1976-2017." Inter-university Consortium for Political and Social Research [distributor].
8

Kaplan, Jacob. 2019*b*. "Uniform Crime Reporting (UCR) Program Data: Raw Data." Inter-university Consortium for Political and Social Research [distributor].
8

Kaplan, Jacob. 2021. "Jacob Kaplan's Concatenated Files: Uniform Crime Reporting (UCR) Program Data: Arrests by Age, Sex, and Race, 1974-2019." Inter-university Consortium for Political and Social Research [distributor].
18

Koteen, Bernard, Elise Baranouski, Joan Ruttenberg, and Carolyn Stafford Stein. 2014. "The Fast Track to a U.S. Attorney's Office." Harvard Law School, report. Available online from `https://hls.harvard.edu/content/uploads/2008/06/fast-track-final.pdf`.
52

Kuipers, Nicholas. 2020. "The Effect of Electing Female Candidates on Attitudes toward Intimate Partner Violence." *The Journal of Politics*, 0(0): 000–000.
5

Le Barbanchon, Thomas, and Julien Sauvagnat. 2021. "Electoral Competition, Voter Bias, and Women in Politics." *Journal of the European Economic Association*, 20(1): 352–394.
24

Lewis, Jeffrey B., Brandon DeVine, Lincoln Pitcher, and Kenneth C. Martis. 2013. "Digital Boundary Definitions of United States Congressional Districts, 1789-2012." Retrieved

from http://cdmaps.polisci.ucla.edu.
9, 10, 37

Lim, Claire S. H., Jr. Snyder, James M., and David Strömberg. 2015. "The Judge, the Politician, and the Press: Newspaper Coverage and Criminal Sentencing across Electoral Systems." *American Economic Journal: Applied Economics*, 7(4): 103–35.
51, 52

Lippmann, Quentin. 2022. "Gender and lawmaking in times of quotas." *Journal of Public Economics*, 207: 104610.
4

Loftin, Colin, David McDowall, and Matthew D. Fetzer. 2008. "A Comparison of SHR and Vital Statistics Homicide Estimates for U.S. Cities." *Journal of Contemporary Criminal Justice*, 24(1): 4–17.
35

Lublin, David. 1997. "Congressional District Demographic and Political Data, 1972-1994."
9

Luca, Dara Lee, Emily Owens, and Gunjan Sharma. 2015. "Can Alcohol Prohibition Reduce Violence against Women?" *American Economic Review*, 105(5): 625–29.
2, 4

Maltz, Michael D. 1999. "Bridging Gaps in Police Crime Data." Report No. NCJ-1176365. Bureau of Justice Statistics, Office of Justice Programs, U.S. Department of Justice, Washington, DC, September, 1999.
33

Maltz, Michael D., and Joseph Targonski. 2002. "A Note on the Use of County-Level UCR Data." *Journal of Quantitative Criminology*, 18: 297–318.
33

Maltz, Michael D., and Joseph Targonski. 2004. "Making UCR Data Useful and Accessible." , (NCJ 205171).
34

Marshall, John. 2022. "Can Close Election Regression Discontinuity Designs Identify Effects of Winning Politician Characteristics?" *American journal of Political Science*.
11

Matawaran, Caroline. 2021. "A Global Look at Femicide."
2

McCrary, Justin. 2008. "Manipulation of the running variable in the regression discontinuity design: A density test." *Journal of Econometrics*, 142(2): 698 – 714. The regression discontinuity design: Theory and applications.
13

NCADV, National Coalition Against Domestic Violence. 2014. "Statistics."
2

Palermo, Tia, Jennifer Bleck, and Amber Peterman. 2014. "Tip of the Iceberg: Reporting and Gender-based Violence in Developing Countries." *American Journal of Epidemiology*, 179(5): 602–612.
2

Pearson, Kathryn, and Logan Dancey. 2011. "Elevating Women's Voices in Congress: Speech Participation in the House of Representatives." *Political Research Quarterly*, 64(4): 910–

923.
24
Pettigrew, Stephen, Karen Owen, and Emily Wanless. 2014. "U.S. House Primary Election Results (1956-2010)."
6
PoliticalParity. 2015. Dataset for the report *Where Women Win: Closing the Gap in Congress*, by Shauna Shames. Available at `https://www.politicalparity.org/research/where-women-win/`.
6
Priyanka, Sadia. 2020. "Do female politicians matter for female labor market outcomes? Evidence from state legislative elections in India." *Labour Economics*, 64: 101822. European Association of Labour Economists, 31st annual conference, Uppsala Sweden, 19-21 September 2019.
3, 16
Reflective Democracy
Reflective Democracy. 2019. "Tipping the Scales: Challengers Take on the Old Boys Club of Elected Prosecutors." Report. Available online from `https://wholeads.us/research/tipping-the-scales-elected-prosecutors/`.
52
Reflective Democracy
Reflective Democracy. 2020. "Confronting the Demographics of Power: America'S Sheriffs." Report. Available online from `https://wholeads.us/wp-content/uploads/2020/06/reflectivedemocracy-americassheriffs-06.04.2020.pdf`.
53
Riedel, Marc, and Wendy C. Regoeczi. 2004. "Missing Data in Homicide Research." *Homicide Studies*, 8(3): 163–192.
33
Siczewicz, Peter. 2011. "U.S. Historical Counties." Dataset. Emily Kelley, digital comp. Atlas of Historical County Boundaries, ed. by John H. Long. Chicago: The Newberry Library, 2011. Available online from `http://publications.newberry.org/ahcbp`.
36
Targonski, Joseph R. 2011. "A Comparison of Imputation Methodologies in the Offenses Known Uniform Crime Reports." PhD diss. University of Illinois at Chicago.
34
United Nations. 2014. *Guidelines for Producing Statistics on Violence Against Women: Statistical Surveys.* United Nations. Statistical Office.
2
U.S. Dept. of Justice, Federal Bureau of Investigation. 2000. "Uniform Crime Reporting Program Data: [United States], 1975-1997 [Supplementary Homicide Reports, 1982] [Computer file]. Compiled by the U.S. Dept. of Justice, Federal Bureau of Investigation. 2nd ICPSR ed." Ann Arbor, MI: Inter-university Consortium for Political and Social Research [producer and distributor].
33
U.S. Dept. of Justice, Federal Bureau of Investigation. 2005. "DATA, Uniform Crime Reporting Program Data [United States]: 1975-1997." Inter-university Consortium for Political and

Social Research [distributor].
35

U.S. Dept. of Justice, Office of Justice, Programs Bureau of Justice Statistics. 1992. "Directory of Law Enforcement Agencies, 1986: [United States]." 1992-02-16. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor].
33

Volden, Craig, Alan E. Wiseman, and Dana E. Wittmer. 2013. "When Are Women More Effective Lawmakers Than Men?" *American Journal of Political Science*, 57(2): 326–341.
24

Zeigermann, Lars. 2018. "OPENCAGEGEO: Stata module for forward and reverse geocoding using the OpenCage Geocoder API."
37

Appendix A Notes on the FBI data

A novelty of our work relative to the existing literature is that we use rich and detailed FBI data. However, this came at a cost in terms of time and effort to collect this data and make them usable for academic purposes. In this Section we discuss some of the main challenges encountered in the use of the UCR data, and how we dealt with them including: (i) representativeness; (ii) identification of missing observations and frequency of the reporting, which all call for a serious reflection on (iii) how to correctly measure crime at the congressional district level. Then, we provide more details on the procedure used to geo-localise our law enforcement agencies.

Appendix A.1 Representativeness

With respect to representativeness, the problem is threefold. On the one hand, the UCR records incidents reported to the police, hence remaining inevitably a subset of actual crime in the United States. This is common when dealing with this type of data, and unfortunately there is no easy way to account for the so-called *dark figure* of crime. However, it should be noted that fatal crimes, such as those included in the SHR, should not be subject to under-reporting as much as other types of felonies, as also argued for example by Fox (2004).

On the other hand, representativeness is undermined by the fact that participation in the program is voluntary, making the list of LEA contributing to the UCR *de facto* a subset of U.S. law enforcement agencies, which might arguably be non-random. In this respect, as already mentioned in Section 3.2, the substantial level of actual participation and the extensive population covered should make this matter less concerning. For example, according to the U.S. Dept. of Justice (2000), as of 1976 over 16,000 law enforcement agencies representing a population in excess of 220 million participated in it. For comparison, the first available census of LEA of 1986 (U.S. Dept. of Justice, 1992) counted 16,707 agencies in the whole United States in that year. Moreover, more recent statistics from the DoJ highlight that agencies in the UCR provide service to approximately 94% of the US population. The DoJ also recognises that coverage is volatile across time and data items, e.g., "the proportion of the U.S. population covered by agencies reporting arrest information ranged from 63% to 94% between 1980 and 2008, with 2008 coverage of 76%. Generally, the coverage of reported crimes (offenses) data exceeds that of arrests (see *ojjdp.gov*).

At the same time, the representativeness of participating agencies does not necessarily imply that the UCR is representative of crimes in the country. As highlighted in Maltz (1999), the fact that the FBI cannot mandate agencies to provide data on time or at all also affects the quality of the data. Amongst the reasons reducing their ability to report consistently, Maltz and Targonski (2002) point at natural disasters, budgetary restrictions, personnel changes, inadequate training, and conversion to new computer or crime reporting system, as well as having very few crimes to report. Despite this, Riedel and Regoeczi (2004) note that, as of 1999, 25 states legally require reporting by their law enforcement agencies.

The SHR file presents the additional issue of being *requested but not required* to all agencies registering some homicides in a given month, i.e., a non-zero value in the corresponding field in the OKCA files. In other words, some agencies might register a murder but they might not provide additional information while including it in the monthly aggregate in the OKCA, effectively making the SHR a subsample of reported homicides. In our dataset, we find that the SHR information is complete for around 90 percent of the cases in this respect.[15] This is consistent with the results of a similar analysis by Fox (2004). Comparing the victim counts from the SHR with the national FBI estimates for the period 1976-2001, he finds that "the SHR file appears to be slightly more than 90% complete, with annual rates of completeness ranging from a high of 96.7% in 1983 to a low of 83.7% in 1998".

In conclusion, it is worth pointing out how this program remains, to our knowledge, the most complete and reliable source of information for crime in the U.S.

Appendix A.2 Frequency of reporting and missing observations

As mentioned, even when participating to the program, agencies decide *when* and *how* to send their reports to the FBI. Although it should be done on a monthly basis, some LEAs record quarterly data, others report yearly or bi-yearly data, while some others only report monthly data for some months. This poses several challenges with respect to the aggregation of our crime data, especially those coming from OKCA (not incident-based). First, arguably the lack of reporting is not always motivated by the lack of crime.[16] Second, these instances are not clearly flagged in the files with gaps filled in with zeros.

This all calls for a careful investigation to correctly distinguish missing observations, real zeros, and aggregated values. A great deal of effort to do this has been put by Maltz and Targonski (2004) and Targonski (2011), who thoroughly investigated the OKCA reports from 1977 to 2000. While we do not attempt to replicate their meticulous analysis, we rely on their methodology to identify missing values at the month-agency level in our OKCA data. More specifically, true missings are found using the field *date-of-last-update*, which the criminologists explain is filled in only for those months in which the agency submits a report to the FBI, and hence "can serve as a proxy for whether or not [a given month's] crime data are missing" (Targonski, 2011). Following this, we set all variables as missing for agency a in month m if, in such month, (i) date-of-last-update is either missing or 000000, and (ii) the sum of all crimes reported is 0.[17]

SHR needs to be treated differently, as it should be filled in only by those agencies which report having had any homicide in the OKCA. In this case we proceed as follows to identify zeros and real missings. First, we use the same SHR and compare the number for total murders and that of femicides, at the month-agency level: if a reported a non-zero value of murders but no value for femicides for a given month m, then femicides were set to 0 in that month. Then, we compared these data with the OKCA: if zero murders were reported in the OKCA and a was not included in the SHR - that is, it was missing according to our calculation based on the SHR data - then both homicides and femicides were set to 0. Otherwise, observations remained missing. The same procedure is followed for other types of killings, e.g., men killed by men or total number of women killed.

[15] Out of a total of 10,434,084 agency-month observations in the OKCA files, we have non-missing values on the number of murders registered by an agency for 7,506,789 observations, and 6,865,154 also have non-missing information in the SHR files. In 6,836,734 cases the two reports coincide, implying that a corresponding supplementary report is filed for every murder registered and recorded by an agency in the OKCA.

[16] The same guidelines of the FBI specify that "if no offenses have occurred during the month, the reporting agency submits the Return A with zeros in the Grand Total row" (FBI (2013))

[17] According to their methodology, a given month is missing if date-of-last-update is missing and if all fields in the OKCA file are 0. They also note how there are cases where there was no "date updated," but the agency actually had reported crime data (Maltz and Targonski, 2004).

With respect to the aggregation of values time-wise, closer inspection revealed a great deal of variation both across and within agencies, making it challenging to combine them appropriately. Rather than unpack data to build monthly estimates, we deemed it more reliable to sum data yearly, as this appears to be the highest frequency of reporting.

Finally, it is important to remark that in doing so we are not augmenting our measures in any way, and we do not attempt to fill in the gaps. Hence, the level of crime which we calculate is likely lower relative to the FBI estimates that follow a different procedure. For example, when aggregating at the county-year level, the Bureau adopted the following strategy (until 1994): for agencies reporting between 3 and 11 months of data - say x months - the number of crimes reported was multiplied by the relative weight $12/x$; for those reporting less than 3 months, crime data from similar agencies were imputed.

Appendix A.3 Measuring crime at the CD level

The main concern for our analysis is the use of spatial aggregation to retrieve CD estimates of crime, as there might not always be perfect coincidence between the location of the agency and that in which the crime occurred. If this were the case, the number of incidents recorded by a given LEA would not necessarily be indicative of the crimes reported in a given area. In this respect, the Summary Reporting System User Manual of the FBI clearly states that "jurisdictional guidelines provide for the most local reporting. Whenever possible, the local LEA of the geographical area in which the crime occurred should report the data" (FBI, 2013). Loftin et al. (2008) also point out that, for the SHR, "victim records are associated with the place where the underlying assault occurred, regardless of where the death occurred or where the victim or offender resided". This should apply in cases of multiple agencies investigating the same crime too. Notice that, in such cases, the FBI guidelines further specify that "a crime that occurs in one jurisdiction should not be counted by another city or county even though more than one agency participated in the investigation or the arrest of the subject of that crime" (Notes to U.S. Dept. of Justice (2005)).

A possible threat to the validity of our methodology comes from agencies *covered by* other ones, a term used to indicate LEA using the services of another one that also reports data to the FBI. As long as the two are located in the same congressional district, however, our measures are not affected by such relationships. With regard to this issue, we note that, generally, a city uses the police services of the sheriff of the county (or of another city, usually nearby). Moreover, the FBI (2013) states that "as a rule, cities having heir own police department report their own crime data".

Most importantly, agencies have different - and at times overlapping - jurisdictions. In fact, other than more local agencies (e.g., a police station covering a city), we find others being responsible for an entire county (e.g., Sheriff), and some with state-wide jurisdictions. In order to avoid double counting, the FBI guidelines say that (i) police count crimes that occur within the city limits, (ii) county and state LEA count crimes that happen in the county outside the city limits (Notes to U.S. Dept. of Justice (2005)). Nevertheless, simply summing the number of incidents recorded by all LEA located in a given CD might still mis-estimate the number of crimes actually occurring in such area when we also include non-local agencies in our sample. In order to avoid inputting values of crime occurring in another part of the state to a given CD, we exclude from our main analysis all of those which we could identify as state-wide (e.g., State Patrol), or which had a special jurisdiction (e.g., Park Police). We include, instead, county-wide agencies (Sheriff's offices) because, in the majority of cases, a county is included in one CD only. Overlapping the maps for county-boundaries for the period 1970-2015 (Siczewicz (2011) and Census Bureau) with that of the corresponding CD-boundaries we find, in fact, that on average only around 10 percent of counties are divided into more than 1 district.[18] Consistently with the criteria according to which population must be equally divided across the different districts assigned to a given state, we also note that counties divided into multiple CD are often the most populated ones, e.g., Los Angeles County is divided into 17/18 districts. This constitutes the dataset of what we refer to as, for simplicity, *local agencies*, and the one we use in our main analysis.

A final concern relates to the possibility that our panel of agencies might not be balanced. In other words, that more agencies might be starting to report their data to the FBI after the election of a woman. If reporting is orthogonal to the agency being located in a CD which undergoes a mixed-gender race and which elects a candidate of a specific gender, our estimates of the effect of women representation on gender-related violence should not be affected by this. In Table A1 we test this hypothesis by estimating the effect on the election of a woman on the number of agencies reporting non-missing data, using our preferred specification. We consider the number of agencies reporting SHR data in Column (1), and its logs in Column (2). In Columns (3) and (4), we replicate the exercise considering those reporting OKCA data. In all cases, we take the average between the first and second year in office. We detect no significant effect on reporting, making us more confident that our panel of agencies is indeed balanced.

Table A1 Number of agencies reporting non-missing data

	SHR		OKCA	
	Levels	Logs	Levels	Logs
	(1)	(2)	(3)	(4)
Coeff	0.596	-0.083	0.451	-0.101
p-value	0.825	0.523	0.846	0.422
h	18	24	17	21
N	391	557	383	495
Avg Outcome	27.719	2.975	27.461	2.963

Notes: Results from a local linear polynomial specification. MSE-optimal bandwidths and a triangular kernel are used. Dependent variables: number of agencies reporting non-missing data (2-year average), expressed either in levels or logs. All regressions include a set of controls as well as time and state fixed effects. Controls include an indicator variable for each geographical macro-area, an indicator for Democratic winner D, percentage of black population, percentage of Hispanic population, and unemployment rate. Standard errors clustered at the district level. Robust p-values are reported. *Avg Outcome* is computed across all districts undergoing mixed-gender races.

[18] Specifically, 16,269 out of 144,182 observations are divided in more than 1 district (yearly share: mean 0.112 (11%), min 0.042, max 0.144). From these computations, we exclude counties overlapping with a congressional district for less than 10 square km, to get rid of possible mistakes due to discrepancies in the shapefiles.

Appendix A.4 FBI data geography

The imputation of the right congressional district to a given agency is, of course, another crucial issue for our study. Agencies might change their location, and redistricting might occur over the years, changing the CD of reference. While the FBI provides additional datasets with some geographical information - namely Crosswalk files, Census of Local and the State Law Enforcement agencies, and Directories of Law Enforcement Agencies - these rarely include data on the CD associated to a given agency, or its geographical coordinates. More importantly, these data only partially cover the period we are considering, which would force us to make important approximations. For example, the first Census takes place in 1986 only, and it is then taken every 4 years. The first Crosswalk file only dates back to 2000.

For this reason, we use the OKCA files to recover the geographical location for each LEA included in the files. Specifically, we use their reported address to retrieve their spatial coordinates through the *Opencage* package by Zeigermann (2018). Overlapping our yearly LEA-maps with those for the CD boundaries by Lewis et al. (2013) for the 92nd-114th Congressional Sessions, and by Census Bureau for the 115th Session of Congress, we are able to retrieve the CD in which a given agency was located in a given year. The exact association made between year and relative Congress is shown in Table A2. The intuition behind it is that members of Congress c starting in year t are elected in the November of year $t-1$ (when the boundaries of the districts are those as of the c-th Congress).

The performance of the package is, inevitably, just as good as the addresses provided, whose quality varies across the years. For example, we find spelling mistakes, or agencies reporting only the name of a city (or, at times, just a county). In this last case, the program will return the coordinates of the centre of the city (county). In order to minimise the occurrences of such instances, we perform several checks and corrections to our original dataset of addresses. We also drop from the analysis those observation which could not be correctly located (e.g., an agency of the state of New York whose address listed New Jersey, or whose address was impossible to find). Nevertheless, this remains a small sample, mostly of non-local agencies.

As a check of our strategy, we compare our results with the Crosswalk files for 2012, which include information related to the agencies' Congressional district.[19] We merge this information with our geo-data for 2012, obtaining a dataset of 22,030 agencies participating in the UCR, of which 3,087 are being listed under more than one congressional district. Comparing the CD which we associated with (any of) those recorded in the Crosswalk, we find that in 1,393 cases (6.32 percent) we have associated our agency to a different district. If we further exclude state-wide agencies, i.e., those not included in the main analysis, we obtain a dataset of 18,565 observations, and in only 430 cases (2.32 percent) is the congressional district we have associated not listed in the Crosswalk. The fact that most of the mistakes are found in non-local agencies – which can be explained by the fact that at times they are all located in a given city (e.g., the state capital), even when responsible for a different area – reinforces the validity of the restriction adopted in our main

[19] The data claim to be reporting congressional district as of 2010, however, we believe this to be a simple mistake in the FBI data, which actually refer to the CD in 2012. For instance, amongst the values listed in their congressional districts, we also find the 34th CD of Texas. However, this district saw its first election in 2012, as it was created as a result of the 2010 Census (see the website of the Congressman representing the district *here*, as well as the Wikipedia page of the district *here*). Therefore, This discrepancy might just be due to a discrepancy with respect to the time references adopted.

analysis.

Table A2 Year and Congress-CD

Year	Congress	Duration of Session	Election
1970	92nd	Jan 3, 1971 to Jan 3, 1973	Nov 3, 1970
1971	92nd		
1972	93rd	Jan 3, 1973 to Jan 3, 1975	Nov 7, 1972
1973	93rd		
1974	94th	Jan 3, 1975 to Jan 3, 1977	Nov 5, 1974
1975	94rd		
1976	95th	Jan 3, 1977 to Jan 3, 1979	Nov 2, 1976
1977	95th		
1978	96th	Jan 3, 1979 to Jan 3, 1981	Nov 7, 1978
1979	96th		
1980	97th	Jan 3, 1981 to Jan 3, 1983	Nov 4, 1980
1981	97th		
1982	98th	Jan 3, 1983 to Jan 3, 1985	Nov 2, 1982
1983	98th		
1984	99th	Jan 3, 1985 to Jan 3, 1987	Nov 6, 1984
1985	99th		
1986	100th	Jan 3, 1987 to Jan 22, 1989	Nov 4, 1986
1987	100th		
1988	101th	Jan 3, 1989 to Jan 28, 1991	Nov 8, 1988
1989	101th		
1990	102nd	Jan 3, 1991 to Jan 9, 1993	Nov 6, 1990
1991	102nd		
1992	103rd	Jan 3, 1993 to Jan 3, 1995	Nov 3, 1992
1993	103rd		

Year	Congress	Duration of Session	Election
1994	104th	Jan 3, 1995 to Jan 3, 1997	Nov 4,1994
1995	104th		
1996	105th	Jan 3, 1997 to Jan 3, 1999	Nov 5, 1996
1997	105th		
1998	106th	Jan 3, 1999 to Jan 3, 2001	Nov 3, 1998
1999	106th		
2000	107th	Jan 3, 2001 to Jan 3, 2003	Nov 7, 2000
2001	107th		
2002	108th	Jan 3, 2003 to Jan 3, 2005	Nov 5,2002
2003	108th		
2004	109th	Jan 3, 2005 to Jan 3, 2007	Nov 2, 2004
2005	109th		
2006	110th	Jan 3, 2007 to Jan 3, 2009	Nov 7, 2006
2007	110th		
2008	111th	Jan 3, 2009 to Jan 3, 2011	Nov 4, 2008
2009	111th		
2010	112th	Jan 3, 2011 to Jan 3, 2013	Nov 2, 2010
2011	112th		
2012	113th	Jan 3, 2013 to Jan 3, 2015	Nov 6, 2012
2013	113th		
2014	114th	Jan 3, 2015 to Jan, 3 2017	Nov 4, 2014
2015	114th		
2016	115th	Jan, 3 2017 to Jan 3, 2019	Nov 8, 2016
2017	115th		

Source: *Wikipedia*. The table indicates the years associated to a given Session of Congress. For each Session of Congress, it additionally specifies the election of reference, as well as the start and end date of the Session.

Appendix B Femicides dataset

As mentioned in Section 3.2, we restrict the SHR data to a sample of 97,705 cases of murder and non-negligent manslaughters, occurring between 1976 and 2017, in which all victims are females and all offenders are males, i.e., killings of women by men. Using additional information regarding the victim-offender relationship, we are able to further restrict our attention to cases in which the victim and the offender had a personal relationship, that is when the victim was one of followings: a common-law wife, a daughter, an ex-wife, a girlfriend, a mother, a sister, a stepsister, a stepmother, or wife, i.e., femicides. Importantly, the SHR provides information only on the relationship between what is recorded as *first* victim and any of the offenders (up to 11). Since the vast majority of cases in our sample involved either one victim or one offender only, we deem this to be a good approximation of the overall victim-offender relationship. This can also be seen in Table B1, which provides a brief description of the broader dataset including all killings of women by men: the proportion of cases involving either one victim only or one offender only always exceeds 94 percent of the total.

Table B1 also shows that femicides are a big proportion of the total cases of killings of women by men by tabulating the relationship between the first victim and first offender.[20] For example, in around a quarter of these killings the victim is the wife. Overall, the femicides in our sample represent around 55 percent of the total cases of women killed by men (55,263 out of 97,705).[21]

[20] Our measure of femicides includes cases in which *any* of the offenders had such a relationship with the victim, e.g., the one listed as second offender. Hence, it includes cases not shown by this tabulation.

[21] In order to better understand the particular characteristics of femicides involving multiple offenders, we also tabulated the number of offenders by distinguishing whether amongst them we found any with no *domestic* ties to the victim (e.g., 2 offenders: a boyfriend and an acquaintance). Specifically, we considered an offender to be *non domestic* if he was listed as an acquaintance, an employee, a friend, a neighbour, other - known to victim, a stranger, or unknown. These instances represent only around 1 percent of the sample. In the vast majority of cases, someone outside of the domestic circle was also involved.

Table B1 Killings of women by men

Nr of victims	No.	%
1	95,313	97.6
2	2,081	2.1
3	249	0.3
4	45	0.0
>4	17	0.0
Total	97,705	100.0

Nr of offenders	No.	%
1	92,197	94.4
2	3,966	4.1
3	1,052	1.1
4	356	0.4
>4	134	0.1
Total	97,705	100.0

Victim-offender relationship	No.	%
acquaintance	15,668	16.0
common-law wife	3,065	3.1
daughter	2,811	2.9
ex-wife	2,002	2.0
friend	2,477	2.5
girlfriend	17,711	18.1
in-law	670	0.7
mother	3,201	3.3
neighbor	1,361	1.4
other	10,880	11.1
other-family	2,014	2.1
sister	805	0.8
stepdaughter	588	0.6
stepmother	150	0.2
stranger	9,483	9.7
wife	24,819	25.4
Total	97,705	100.0

Notes: Total SHR: 705,434 cases, of which *murder and non-negligent manslaughter*: 698,686. Femicides: Max number of victims is 8, max number of offenders is 11. *Relationship* refers to that between first victim and first offender. The label *other* includes relationships such as employee, other-known to victim, unknown.

Appendix C Additional tables

Table C1 Alternative normalisations and measures

	W killed by M				**W Killed**			
	Pop		Fem Pop		Pop		Fem Pop	
	t1	t2	t1	t2	t1	t2	t1	t2
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coeff	-0.081	0.038	-0.127	0.07	-0.059	0.112	-0.046	0.203
p-value	0.274	0.616	0.305	0.535	0.545	0.334	0.847	0.168
h	18	14	18	15	19	15	16	14
N	494	270	484	274	530	280	447	269
Avg Outcome	0.569	0.898	1.046	1.53	0.721	1.094	1.265	1.782

Notes: Results from a local linear polynomial specification. MSE-optimal bandwidths and a triangular kernel are used. Dependent variables expressed as log of number of victims either per 100,000 residents (*Pop*) or per 100,000 females in voting age (*Fem Pop*). Values refer to either the first year of the term (t1) or the overall term (t2). All regressions include a set of controls as well as time and state fixed effects. Controls include an indicator variable for each geographical macro-area, an indicator for Democratic winner D, percentage of black population, percentage of Hispanic population, unemployment rate, and percentage of agencies reporting non-missing data (SHR). Standard errors clustered at the district level. Robust p-values are reported. *Avg Outcome* is computed across all districts undergoing mixed-gender races.

Table C2 Other crimes

Panel A: First year								
	Violent Crimes				**Non-Violent Crimes**			
	Murders	M killing M	Violent Crimes	Aggr ass	Robbery	Burglary	Larceny	Mv Theft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coeff	0.004	-0.002	-0.007	-0.053	-0.044	-0.059	0.03	0.15
p-value	0.824	0.828	0.865	0.82	0.932	0.85	0.75	0.312
h	13	20	14	23	14	12	13	13
N	368	552	314	731	397	351	354	358
Avg Outcome	1.589	1.044	5.438	5.175	4.376	6.429	7.579	5.47
Panel B: Overall term								
	Violent Crimes				**Non-Violent Crimes**			
	Murders	M killing M	Violent Crimes	Aggr ass	Robbery	Burglary	Larceny	Mv Theft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coeff	0.227	0.172	0.149	0.183	0.132	0.084	0.108	0.231
p-value	0.135	0.125	0.283	0.247	0.429	0.435	0.274	0.157
h	14	12	21	16	15	19	21	19
N	295	230	369	353	319	440	482	433
Avg Outcome	2.117	1.47	6.126	5.857	5.026	7.122	8.278	6.153

Notes: Results from a local linear polynomial specification. MSE-optimal bandwidths and a triangular kernel are used. Dependent variables expressed as log of number of victims/occurrences per 100,000 residents. Values refer to either the first year of the term (Panel A) or the overall term (Panel B). All regressions include a set of controls as well as time and state fixed effects. Controls include an indicator variable for each geographical macro-area, an indicator for Democratic winner D, percentage of black population, percentage of Hispanic population, unemployment rate, and percentage of agencies reporting non-missing data (OKCA or SHR). Standard errors clustered at the district level. Robust p-values are reported. *M killing M* refers to male-victims of murder killed by another male(s). *Aggravated Assault* is an unlawful attack by one person upon another for the purpose of inflicting severe or aggravated bodily injury. *Violent crimes* are defined by the FBI as those offenses involving force or threat of force. For our analysis, murder and non-negligent manslaughter, rape, and aggravated assault are included. The official FBI definition also includes robbery. As this felony is counted according to the number of occurrences rather than number of victims, this category is excluded for consistency. See FBI (2013), FBI website at ucr.fbi.gov/violentcrime and ucr.fbi.gov/assault. *Avg Outcome* is computed across all districts undergoing mixed-gender races.

Table C3 Arrests - Total

Panel A: First year	Murder	Rape	Aggr Ass	Robbery	Burglary	Larceny	Mv Theft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coeff	-0.035	0.081	-0.07	0.049	0.046	0.03	0.068
p-value	0.983	0.425	0.745	0.659	0.618	0.619	0.627
h	14	15	20	12	16	15	15
N	366	382	556	300	440	381	408
Avg Outcome	1.437	2.08	4.519	3.316	4.506	1.653	3.375

Panel B: Overall term	Murder	Rape	Aggr Ass	Robbery	Burglary	Larceny	Mv Theft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coeff	0.23	0.285	0.088	0.233	0.196	0.105	0.109
p-value	0.095	0.036	0.473	0.221	0.168	0.3	0.414
h	14	13	15	12	13	13	14
N	264	240	299	230	239	239	267
Avg Outcome	1.939	2.676	5.195	3.942	5.166	2.223	4.038

Note: Results from a local linear polynomial specification. MSE-optimal bandwidths and a triangular kernel are used. Dependent variables expressed as log of number of people arrested per 100,000 residents. Values refer to either the first year of office (Panel A) or the overall term (Panel B). All regressions include a set of controls as well as time and state fixed effects. Controls include an indicator variable for each geographical macro-area, an indicator for Democratic winner D, percentage of black population, percentage of Hispanic population, unemployment rate, and percentage of agencies reporting non-missing data (ASR). Standard errors clustered at the district level. Robust p-values are reported. *Avg Outcome* is computed across all districts undergoing mixed-gender races.

Table C4 Arrests - Males

Panel A: First year	Murder	Rape	Aggr Ass	Robbery	Burglary	Larceny	Mv Theft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coeff	-0.025	0.065	-0.073	0.056	0.056	-0.001	0.057
p-value	0.921	0.501	0.731	0.633	0.561	0.811	0.679
h	14	15	20	12	14	14	15
N	364	394	550	304	377	344	383
Avg Outcome	1.35	2.069	4.326	3.212	4.379	1.529	3.218
Panel B: Overall term	Murder	Rape	Aggr Ass	Robbery	Burglary	Larceny	Mv Theft
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Coeff	0.235	0.279	0.077	0.256	0.193	0.076	0.094
p-value	0.089	0.04	0.504	0.181	0.176	0.388	0.487
h	14	13	15	12	13	13	13
N	273	240	299	230	239	246	260
Avg Outcome	1.837	2.663	4.998	3.831	5.033	2.082	3.873

Note: Results from a local linear polynomial specification. MSE-optimal bandwidths and a triangular kernel are used. Dependent variables expressed as log of number of males arrested per 100,000 residents. Values refer to either the first year of office (Panel A) or the overall term (Panel B). All regressions include a set of controls as well as time and state fixed effects. Controls include an indicator variable for each geographical macro-area, an indicator for Democratic winner D, percentage of black population, percentage of Hispanic population, unemployment rate, and percentage of agencies reporting non-missing data (ASR). Standard errors clustered at the district level. Robust p-values are reported. *Avg Outcome* is computed across all districts undergoing mixed-gender races.

Appendix C.1 Robustness tables

In this subsection, we provide details on the robustness tests introduced in Section 6.

In Tables C5 and C6 we evaluate how our main results change when using alternative normalisations and econometric specifications. More specifically, in Table C5 we replicate the analysis in Table 3 normalising our dependent variables by female (voting) population rather than by total population, then expressing it in logs. In Table C6 we consider the number of victims (divided by total population) in levels rather than its logarithmic transformation. In both cases, our main results are maintained: we find a significant decrease in femicides in the short-run, and an increase in rapes over the whole term (Column (2) in Panel A, and Column (6) in Panel B). As we might have expected, the coefficients in Table C5 become slightly larger with respect to those in Table 3.

In Table C7 we consider an alternative sample of law enforcement agencies which, in our main analysis, was limited to local agencies only. We now expand it to all agencies included in the UCR files, replicating once again the analysis in Table 3. As we can see, when we do so our main findings are qualitatively unchanged. As mentioned, we argue that the loss of significance in some of the specifications can be explained by the fact that the inclusion of non-local agencies (e.g., State Police) renders our data more noisy, because of the increased measurement error (see also Appendix A.4).

Finally, in Table C8 we test whether our results are driven by another factor, different than gender, which correlates with it: party affiliation, and in particular affiliation to the Democratic party. We do so by implementing a regression discontinuity design on close-races between a Democratic and a Republican candidate. Adapting the strategy described in Section 4, we now consider all races where the first and second place are either (i) R and D, or (ii) D and R, and compute $MV_{d,s,t} = \frac{(D_{d,s,t} - R_{d,s,t})}{Votes_{d,s,t}}$, that is the difference in the share of (total) votes between the Democratic and the Republican candidate. In this case, our treatment variable is $D_{d,s,t} = 1(MV_{d,s,t} > 0)$, an indicator for when the Democratic candidate wins. Besides this, the specifications are equivalent to the ones used in Table 3, as are both the normalisation and the transformation used (i.e., total population, and logs). The results in both Panels show that the political affiliation of a Congressperson does not have an effect on either femicides or rapes, neither in the short- or the medium-run. As we might have expected, the OLS results indicate the presence of some correlation, but in none of the other specifications are the coefficients significant. This finding makes us more confident that our results are not driven by the party affiliation of the winning candidate.

Table C5 Alternative normalisation: female population

Panel A: First year								
	Femicides				**Rapes**			
	OLS	RDD			OLS	RDD		
Bandwidth:		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$		$1h_{MSE}$	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coeff	-0.052	-0.164	-0.187	-0.138	0.034	0.03	0.008	-0.126
p-value	0.118	0.081	0.088	0.313	0.608	0.804	0.765	0.572
h	100	19	29	10	100	14	20	7
N	1510	513	780	227	1353	305	483	157
Avg Outcome	0.77				4.173			

Panel B: Overall term								
	Femicides				**Rapes**			
	OLS	RDD			OLS	RDD		
Bandwidth:		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coeff	-0.040	0	-0.048	0.024	0.078	0.33	0.133	0.488
p-value	0.347	0.920	0.865	0.981	0.295	0.032	0.069	0.122
h	100	18	27	9	100	18	27	9
N	1066	337	517	153	970	297	458	133
Avg Outcome	1.191				4.859			

Note: dependent variables expressed as logs of number of victims per 100,000 female in voting age population. Values refer to either the first year of office (Panel A) or the overall term (Panel B). All regressions include a set of controls as well as time and state fixed effects. Controls include an indicator variable for each geographical macro-area, an indicator for Democratic winner D, percentage of black population, percentage of Hispanic population, unemployment rate, and percentage of agencies reporting non-missing data (OKCA or SHR). Standard errors clustered at the state level. For the RDD specification, MSE-optimal bandwidths and a triangular kernel are used, and robust p-values are reported. *Avg Outcome* is computed across all districts undergoing mixed-gender races.

Table C6 Levels

Panel A: First year								
	Femicides				**Rapes**			
	OLS	RDD			OLS	RDD		
Bandwidth:		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coeff	-0.058	-0.188	-0.221	-0.089	2.421	1.686	0.618	9.887
p-value	0.256	0.059	0.066	0.105	0.431	0.636	0.448	0.274
h	100	20	30	10	100	24	36	12
N	1515	535	812	239	1358	578	877	272
Avg Outcome	0.586				37.193			

Panel B: Overall term								
	Femicides				**Rapes**			
	OLS	RDD			OLS	RDD		
Bandwidth:		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coeff	-0.027	-0.127	-0.123	-0.258	5.432	27.828	23.099	22.648
p-value	0.802	0.63	0.802	0.391	0.455	0.02	0.035	0.126
h	100	13	20	7	100	12	17	6
N	1069	230	375	119	973	183	291	94
Avg Outcome	1.129				71.991			

Notes: dependent variables expressed as number of victims per 100,000 residents. Values refer to either the first year of office (Panel A) or the overall term (Panel B). All regressions include a set of controls as well as time and state fixed effects. Controls include an indicator variable for each geographical macro-area, an indicator for Democratic winner D, percentage of black population, percentage of Hispanic population, unemployment rate, and percentage of agencies reporting non-missing data (OKCA or SHR). Standard errors clustered at the district level. For the RDD specification, MSE-optimal bandwidths and a triangular kernel are used, and robust p-values are reported. *Avg Outcome* is computed across all districts undergoing mixed-gender races.

Table C7 Whole sample

Panel A: First year

	Femicides				Rapes			
	OLS	RDD			OLS	RDD		
Bandwidth:		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coeff	-0.022	-0.104	-0.115	-0.119	0.007	-0.065	-0.041	-0.308
p-value	0.281	0.036	0.034	0.04	0.916	0.765	0.984	0.092
h	100	20	30	10	100	14	20	7
N	1528	537	811	239	1374	314	494	163
Avg Outcome	0.406				3.236			

Panel B: Overall term

	Femicides				Rapes			
	OLS	RDD			OLS	RDD		
Bandwidth:		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coeff	-0.015	-0.04	-0.056	-0.095	0.046	0.17	0.018	0.373
p-value	0.606	0.674	0.68	0.297	0.544	0.19	0.233	0.164
h	100	16	24	8	100	20	30	10
N	1078	294	466	139	993	355	537	162
Avg Outcome	0.677				3.918			

Note: dependent variables expressed as logs of number of victims per 100,000 residents. Values refer to either the first year of office (Panel A) or the overall term (Panel B), and consider the whole sample of agencies (local and non-local). All regressions include a set of controls as well as time and state fixed effects. Controls include an indicator variable for each geographical macro-area, an indicator for Democratic winner D, percentage of black population, percentage of Hispanic population, unemployment rate, and percentage of agencies reporting non-missing data (OKCA or SHR). Standard errors clustered at the district level. For the RDD specification, MSE-optimal bandwidths and a triangular kernel are used, and robust p-values are reported. *Avg Outcome* is computed across all districts undergoing mixed-gender races.

Table C8 Republican vs Democrats - RDD

Panel A: First year

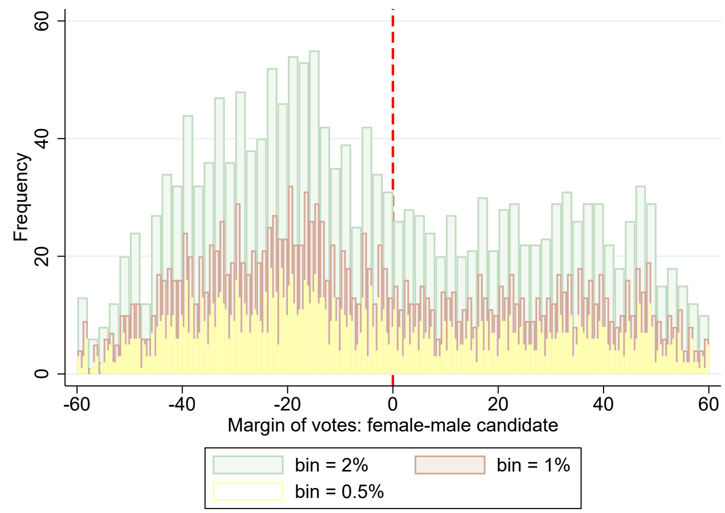
	Femicides				**Rapes**			
	OLS	RDD			OLS	RDD		
Bandwidth:		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coeff	0.028	-0.018	-0.014	-0.038	0.095	0.001	-0.001	-0.069
p-value	0.007	0.52	0.556	0.415	0.004	0.943	0.914	.232
h	100	19	29	10	100	16	24	8
N	6559	2071	3280	983	6901	1825	2785	888
AvgOutcome	0.405				3.121			

Panel B: Overall term

	Femicides				**Rapes**			
	OLS	RDD			OLS	RDD		
Bandwidth:		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$		h_{MSE}	$\frac{3}{2}h_{MSE}$	$\frac{1}{2}h_{MSE}$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Coeff	0.042	0.011	0.011	-0.016	.105	-0.039	-0.025	-0.076
p-value	0.004	0.783	0.757	0.688	0.005	0.591	0.657	0.317
h	100	25	37	12	100	17	25	8
N	4573	1947	3116	916	4984	1390	2131	684
Avg Outcome	0.679				3.815			

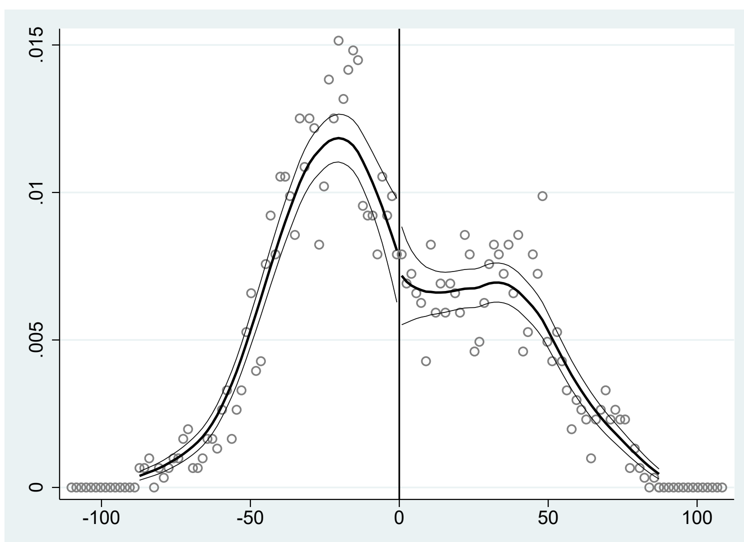
Notes: dependent variables expressed as logs of number of victims per 100,000 residents. Values refer to either the first year of the term (Panel A) or the overall term (Panel B). All regressions include a set of controls as well as time and state fixed effects. Controls include an indicator variable for each geographical macro-area, an indicator for female winner W, percentage of black population, percentage of Hispanic population, unemployment rate, and percentage of agencies reporting non-missing data (OKCA or SHR). Standard errors clustered at the district level. For the RDD specification, MSE-optimal bandwidths and a triangular kernel are used, and robust p-values are reported. *Avg Outcome* is computed across all districts undergoing "mixed-party" races.

Appendix D Additional graphs



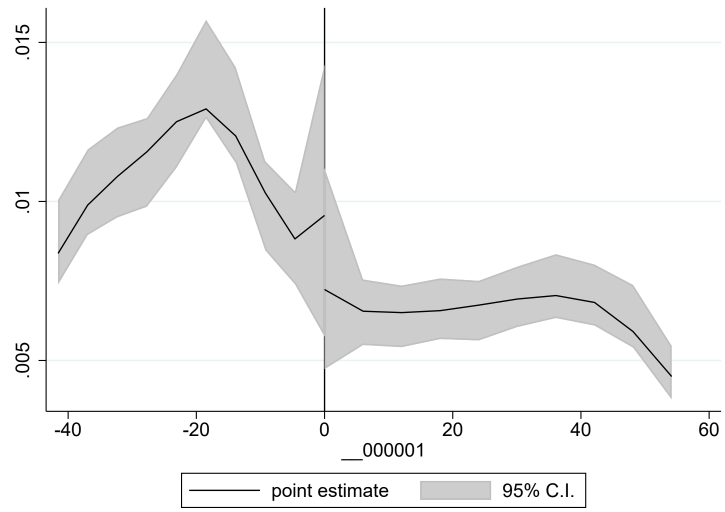
Frequency plot. The bars indicate the (absolute) frequency of the races within a given bin, computed according to three different bandwidths: 2 percent, 1 percent and 0.5 percent.

Figure D1 Frequency of female margin of votes, MV



McCrary plot: Discontinuity estimate, point estimate: -0.0707 (SE 0.177).

Figure D2 McCrary density plot



Rddensity plot: p-value for bias-corrected density test: 0.436.

Figure D3 RD-density plot

Appendix E Other offices: judges and more

As mentioned in Section 5.4, in order to investigate the possibility of conflicting elections (or appointments) confounding the effect of female political representation we documented, we examined four offices of particular importance in the justice and in the law-enforcement systems: state trial court judges[22], U.S. attorneys, district attorneys[23], and sheriffs. We focus, in particular, on three features of such offices, namely (i) the selection process, e.g., appointment or election, (ii) the geographical area over which these authorities exert their jurisdiction, (iii) the length of their terms, which we detail below. Compared to the office for U.S. Representative, we find substantial differences, especially with respect to the timing in office and in the area served, which limits our concerns.

More specifically, consider the case of state trial court judges. These are the first level in a state's court system, below the appellate courts and a state's supreme court. According to data from the Brennan Center for Justice (2022) and to the analysis in Lim et al. (2015), there is vast heterogeneity in the way these are selected across states: in some they are appointed by the Governor, in some others they are elected in partisan elections after being nominated by political parties and, most commonly, in some they are elected in non-partisan elections (i.e., without party identification on the ballot). Importantly, this variation refers not only to the selection for their first full term in office, but also to the the way judges are, eventually, re-instated to their offices. Only looking at how they are selected for their first full term, we find that judges are elected in 28 states, while gubernatorial appointments are used in 18. In a few cases variation can also be found *within* the same state.[24] Where judges are appointed (around a third of the states) our concerns are limited to the extent that many more and many different factors, not only "local", enter into consideration when a Governor or a state Legislature needs to make such a decision. But whatever the method of selection, it is important to note that judges likely have broader geographical jurisdiction than a U.S. Representative, and that they are always in charge for longer periods. Lim et al. (2015), in particular, note that some small states in New England (e.g., Maine, New Hampshire) have just one judicial district covering the whole state; that, in Southern and Midwestern states, judicial districts tend to cover several (three or four) counties; that in the Pacific region (e.g., California) and Mid-Atlantic region (e.g., New Jersey), judicial districts tend to cover just one or two counties. Given the possibility that one of more counties falls within the boundaries of a given Congressional district (see also Appendix A.3), this would open the possibility that – in some very specific cases – the boundaries of a judicial district might coincide with those of a CD, or that more than one judicial districts could be exactly included in a given CD. This concern might be particularly acute in those states where the geographical jurisdiction seems smaller, that is in the Pacific and Mid-Atlantic regions, which also elected judges. In this respect, we note that these represent a smaller set of states, since these two regions include 8 states, out of which 5 elect judges (Brennan

[22]State trial courts are also called district courts, circuit courts, or superior courts (Lim et al., 2015).

[23]Depending on the state, district attorneys are also called: state attorneys, prosecuting attorneys, county attorneys, or public prosecutors. We refer to *District Attorneys* for any of these five denominations. See the American Civil Liberties Union (ACLU) and the Legal Information Institute (LII) at Cornell University.

[24]According to the Brennan Center for Justice (2022), 19 states use non-partisan elections, 9 states use partisan elections, 18 states use gubernatorial appointments. The remaining states use either a legislative appointment, that is a vote of the state Legislature (2 states), or other mixed-system of selection (3 states, e.g., Arizona, where judges can be selected either by gubernatorial appointment or by non-partisan elections, depending on the jurisdiction). Data include the District of Columbia. See also Lim et al. (2015).

Center for Justice, 2022).[25]

The duration of a judicial term remains, however, the stronger case against the possibility of a potential confounding effect. Irrespective of the way judges are selected, in fact, they are typically in charge for 6 years, and always for periods longer than 2 years. Hence, it is often the case that a Representative is elected when a judge has already been serving his or her term for some years. We also note that in 3 states judges are life-tenured (New Hampshire, Rhode Island and Massachusetts, where they are selected by the Governor). Moreover, that 8 states use retention election, that is: when a judge is up for an additional term, he or she stands for an yes-or-no vote, and no other candidates appear on the ballot. Lim et al. (2015) note that, in this type of races, incumbent judges win more than 99 percent of the time, making the office of a judge basically a life-time appointment in states with retention elections. Of these 8, in 2 states judges run for elections for their first term (Brennan Center for Justice, 2022).

In addition, it is important to remark how each of the features just discussed varies across states, which creates a mosaic of selection processes (e.g., two states selecting a judge through elections for his first term, but differing in the way this is later re-appointed) and geographical jurisdictions, making it less likely for our results to be driven by any particular small sample of states in *some* even years, that is when the timing of the two elections also coincides.

U.S. Attorneys are, instead, always appointed by the U.S. President and confirmed by the Senate. They serve terms of 4 years (or at the President's discretion, see 28 U.S. Code § 541), which means that the term of a given U.S. attorney overlaps with approximately 2 Sessions of Congress. Moreover, their jurisdiction is over a federal judicial districts, typically wider than a Congressional District. To have a sense, in the U.S. there are 93 federal judicial districts, but 435 Congressional districts (Koteen et al., 2014). Hence, similar arguments as those just made when discussing judges apply, and we can more confidently exclude the possibility the elections for this office conflict with those of a U.S. Representative.

In the case of DAs, these also serve for 4 years, they are mostly elected, and they might serve either a judicial district or county. Hence, in this instance too a Representative is likely elected at a time when a DA has been in office for some time already. Nevertheless, there is the possibility that, in *some years*, the elections for a DA and for a Representative might "conflict", in the sub-sample of states where state-prosecutors are elected in and serving a county whose boundaries coincide with (or are included in) those of a CD. To have a sense, as of 2001, in 3 states DAs were appointed (Alaska, Connecticut, New Jersey) and in 47 elected, out of which in 26 the DAs had jurisdictions over a county (Coppolo, 2003). In this respect, three more things need mention. First, it seems that not all DAs are up for re-election at the same time. For example, the American Civil Liberties Union (ACLU) of Oregon notes that in the state "in any given election, at least a third of the District Attorneys are up for re-election". Second, that most DAs run unopposed. During the 2018 *prosecutorial elections*, 80 percent (836) of candidates run for office without an opponent. Third, that white males represents the vast majority of *prosecutors* – around 70 percent of them, as of 2019 (Reflective Democracy, 2019).[26] This would indicate that, typically, elections are hardly

[25] Mid-Atlantic: New Jersey, New York, and Pennsylvania. Pacific: Alaska, California, Hawaii, Oregon, and Washington. Judges are appointed by the Governor in Alaska, Hawaii, New Jersey, and elected in the the remaining states. To note: in Alaska and Pennsylvania, judges face retention elections.

[26] The definition of *Prosecutors* used in Reflective Democracy (2019) is broader than the one we adopted, in that it includes not only District Attorneys, State's Attorneys, Prosecuting Attorneys, County Attorneys or Public Prosecutors, but also e.g., Solicitor Generals, and Attorney Generals. However, our original categories cover around 80

ever competitive and mostly dominated by males. Furthermore that, even when competition is present, it is mostly amongst males.

Similar arguments can be applied in the case of sheriffs, who have jurisdictions over a county and are mostly elected. More specifically, sheriffs are elected in 46 out of all 48 states in which such office is present.[27] The length of their term varies, but the overwhelming majority of sheriffs stays in office for 4 years: as of 2012, in two states only they were elected to a 2-year term (Arizona and New Hampshire), in one to a 3-year term (New Jersey), in one to a 6-year term (Massachusetts).[28] Therefore, once again, despite the more likely geographical overlap of elections, our concerns of a confounding effect are limited by the fact that a sheriff's term almost always extend that of an elected U.S. Representative. Most importantly, recent statistics show how, also in the case of sheriffs, the majority of them (60 percent) runs unopposed, and that they are overwhelmingly men. In 35 states, fewer than 5 percent of sheriffs are women, out of which 16 have no women Sheriffs. Overall, (white) men represents around 90 percent of Sheriffs in the U.S. (Reflective Democracy, 2020).

All in all, these elements seem to go against the interpretation of our results being confounded by the fact that the election of female a Representative coincides with broader female representation in the district, not just in the political but also in the judicial and law-enforcement sphere, making us more confident of the fact that what we capture is, in fact, coming from female *political* representation.

percent of the sample of all elected prosecutors which were in office in 2019.

[27]Alaska does not have a Sheriff's Office, Connecticut abolished the office in early 2000s. Sheriffs are appointed in Rhode Island and Hawaii.

[28]Data from the office of the Sheriffs (archived from the original on 19 October 2014, with Wayback Machine). See also the *Wikipedia* page