When Women Run Against Men: Evidence from Political Platforms

Marli Fernandes[*]

Università Degli Studi di Padova

October 2, 2024

Abstract

Do female and male candidates differ in their political campaigns? And why do they differ? Using individual political platforms from legislative elections in France, I combine computational text analysis with a regression discontinuity design setup in the two-round French legislative elections to understand political campaign differences between women and men. I find that women give more salience to topics such as security and foreign policy than males. This result is stronger in places where there is more substantial voter gender bias: in districts that have never elected a woman or where the gender wage gap is higher. I causally show that when women run against men, as opposed to running against a woman, they strategically give more prevalence to male-stereotypical topics. However, once elected, women provide more coverage to female stereotyped issues, health and education compared to male colleagues. In contrast, when male politicians run against women, they adapt their platforms more marginally.

Keywords: Elections, gender, text as data

JEL classification: D72, J16, P0

1 Introduction

The share of women in politics has increased remarkably over the past few decades in almost every country in the world. While in 1988, women were only 8.9% of the candidates for US primaries, in 2024 they constituted 26.2% (the highest share was in 2020, 29.1%).[1] In France, the share increased from 11.5% to 41.1% in the same period. These numbers show a remarkable improvement but demonstrate that although women constitute half of the population, they are still a minority in politics. How different are women from men during political campaigns? And why do they differ? Answering these two questions is important for two reasons. First, it will permit us to understand what the increased visibility of women implies for policy promotion. Second, identifying why women differ from men during political campaigns is key to understanding the gender gap in politics and designing policies that aim to reduce it.

This paper studies whether there are gender differences in political campaigns and why they emerge. Several research studies have shown gender differences in preferences (Croson and Gneezy (2009)); this implies that political campaigns are informative to voters since these differences inform voters about how the candidates genuinely differ. On the other hand, it is documented that women respond less favorably to competition than men (Niederle and Vesterlund (2011)). In addition, being elected is incredibly challenging for women since they face voter bias (De Paola, Scoppa, and Lombardo (2010), Le Barbanchon and Sauvagnat (2022), Eyméoud and Vertier (2023)). If women believe that the voters discriminate against their characteristics, women need to be careful about what information they should convey to voters or what policies to promote. This last situation implies that female politicians cannot signal to voters their true type and convey correct information about them.

Studying gender differences in political campaigns and why they emerge is methodologically challenging because a female presence in a race is endogenous to electoral districts' characteristics. In addition, data on political discourse during campaigns is often lacking. In most countries, there is no record of campaigns run by individual politicians, except France. Because the majority of the data only covers election winners in their political office, research is mostly concentrated on how women perform once elected.

I focus on the legislative elections in France, where the same pool of voters is called to participate in France's first and second rounds. With an average of nine candidates in the first round and two in the second, politicians face the challenge of

---

[1]Source: Center for American Women and Politics.

appealing to a broader electorate. To overcome endogeneity issues, I exploit two-stage elections to estimate the impact of gender in campaign strategies. I use a regression discontinuity design to isolate quasi-random variation in the candidates' gender. The research design exploits that in a two-round system, where a female politician, instead of a man, barely makes it to the final round, and those who just miss the qualification threshold, the electoral district and opponents' characteristics are arguably comparable in observed and unobserved characteristics.

I use a comprehensive candidates' political platforms data set that provides the record of the campaign messages sent by every competing politician during the legislative elections in France. Individual candidates issue their own campaign platforms, which the State prints and sends to all registered voters a few days before the election. Politicians use these platforms to inform their constituents about themselves, their program and appeal to vote.

Employing a semi-supervised machine learning technique (seeded Latent Dirichlet Allocation), I estimate the share of each topic in a political platform. Compared to men, women give the same coverage to all policy topics except for security and foreign policy, a male-stereotyped topic, which provides more coverage by 6.6 percentage points. I show that differences in campaign financing and convergence to the opponent or electorate preferences cannot explain the results. In fact, I provide evidence that this difference is an adaptive behavior to voter bias. In districts that have never elected a woman before, women give more coverage to security and foreign policy than males by 8.4 percentage points. In districts where the gender wage gap is above average, the difference is 13 percentage points, while no statistically significant difference exists in districts where the gender gap is below average. I provide further evidence that women adapt to voter bias against them. When women marginally compete with a man instead of a woman, they campaign more on this topic. On the other hand, men do not change the coverage of their topics when they compete against a woman rather than a man.

Finally, I test whether politicians commit to their announced policy platform after the election. I focus on the behavior of politicians in legislative debates, where there is strong party discipline, and in the written questions sent to the members of the government, which members of the parliament (MPs) usually send as representatives of their constituency and party discipline are less significant. A simple summary of statistics shows that, although women give more coverage than males to security and foreign policy issues during campaigns, this difference disappears after elections in both legislative activities. Focusing on mixed-gender races, I compare barely elected female candidates with males elected by a small margin. I causally prove that women are as active as men in legislative debates,

a participation that requires charismatic and good rhetorical abilities. Moreover, in legislative debates and written questions, women speak more about health & education than men by a significant margin. This last result has several implications. First, it reinforces the idea that women strategically adapt stereotyped male traits during their campaigns to address voters' gender bias. Second, it proves that voters cannot infer women's correct type during campaigns, potentially leading to adverse selection. Third, it shows that issues favored by women do not get more attention in political campaigns by the most legitimate politicians.

The paper is organized as follows. In Section 2 I overview the related literature. In Section 3 I describe the institutional context. Section 4 presents the data, Section 5 the methodology, Section 6 the results and Section 7 concludes.

2   Literature review

The paper contributes to several strands of the literature. First and foremost, the paper contributes to the literature that studies gender differences in the behavior of politicians and what factors contribute to these differences. Several studies have shown that women prefer to work on different policies, specifically on health and education (Chattopadhyay and Duflo (2004), Bhalotra and Clots-Figueras (2014), Funk and Gathmann (2015), Lippmann (2021)). In contrast, others do not find any significant differences (Ferreira and Gyourko (2014), Bagues and Campa (2021), Casarico, Lattanzio, and Profeta (2022)). This literature is concentrated on elected politicians as it is considerably easier to obtain data. My paper provides, to my knowledge, for the first time, a causal description of gender differences during political campaigns. Studying behavior during political campaigns is equally essential; in models of electoral competition, politicians are held accountable and must commit to their announced policy platform since their campaign promises reflect future policies (Downs (1957), Besley and Coate (1997)).

Several studies have also shown that voters tend to be biased against women (De Paola et al. (2010), Le Barbanchon and Sauvagnat (2022), Eyméoud and Vertier (2023)). In addition, voters are gender-biased when they evaluate politicians' abilities. According to the role congruity theory, it is harder for women to become leaders due to two forms of prejudice: (1) women are perceived less favorably than males as potential occupants of leadership roles, (2) women are less perceived as having the necessary abilities of a leader role (Eagly and Karau (2002)). In accordance with this theory, Lawless (2004) shows that female politicians are evaluated less favorably than males in dealing with male-stereotyped issues, such as national security and military crises. On the other hand, Herrnson, Lay, and Stokes (2003)

shows that women gain a strategic advantage when they campaign as a "woman" on issues such as health and education. This paper contributes to this literature since it studies how candidates adapt their campaigns to voter bias.

Second, the paper is related to the impact of politicians' identity. The Median Voter Theorem (Downs (1957)) assumes the parties' unique objective is winning elections. Consequently, if two parties have the same information about voters' preferences, they will converge to the viewpoint of the median voter, ignoring politicians' identity. Le Pennec (2023) and Di Tella, Kotty, Le Pennec, and Pons (2023) empirically demonstrate the convergence mechanism underlying the median voter theorem. However, candidates only converge to a certain extent. In citizen-candidate models, parties do not only care about winning elections but also about implementing their preferred policies (Alesina (1988), Osborne and Slivinski (1996), Besley and Coate (1997)). The findings of this paper extend this literature in understanding the influence of alternative dimensions of identity, in this case, gender. It contributes to clarifying whether any of the models can predict politicians' behavior in identity situations.

Third, the paper also contributes to the extensive literature on gender and competition. For a review, see Niederle and Vesterlund (2011). A series of laboratory studies documents that, conditional on performance, women are often more reluctant to compete than men (e.g., Niederle and Vesterlund (2007), Markowsky and Beblo (2022)). This pattern has been confirmed in the case of elections (Kanthak and Woon (2015), Barber, Butler, Preece, et al. (2016)). These differences might be a reflection of social learning (Booth and Nolen (2012)) or culture (Gneezy, Leonard, and List (2009)) rather than inherent gender traits. Research has also shown that these differences depend on their own gender and on the gender of people with whom they interact. In a two-person bargaining game, competition and retaliation are higher when the bargaining partners have the same gender (Sutter, Bosman, Kocher, and van Winden (2009)). Finding opposing results, Datta Gupta, Poulsen, and Villeval (2013) conclude that individuals compete less with same-sex opponents in tournaments. Indeed, literature on how individuals compete, considering the opponents' gender, is limited and needs further exploitation. This paper contributes to this literature by studying, to my knowledge, for the first time, how politicians adapt their races when they compete with the opposing sex.

3 Institutional context

3.1 French parliamentary elections

The paper focuses on parliamentary elections, which elect all the members of the National Assembly, the lower house of Parliament.

These elections are held under direct universal suffrage with a two-round plurality voting rule. Candidates can directly win in the first round if they obtain a number of votes greater than 50 percent of the votes and 25 percent of the registered citizens. However, in 88.74% of the elections, no candidate is elected in the first round, leading to the regular and predictable occurrence of a second round held one week later. The second round is decided by simple plurality: the candidate with the largest vote share wins the election.

Candidates that obtain at least 12.5 percent of the vote share of the eligible voters are qualified for the second round. However, if only one of the candidates (or none) meets the threshold, the two candidates with the largest vote share can proceed to the second round. This last situation corresponds to 28.19% of all elections.

Selection of political candidates It is possible to run for one of the 577 French constituencies without being affiliated to a party as long as the aspiring MP meets all the necessary conditions to enter the race, in particular being at least 18 years old, having the right to vote, and not being ineligible (because of a court decision or a function incompatible with the mandate of an MP, such as being mayor). However, most candidates run under a party label, making them much more visible to voters.

According to the 2000 gender parity law, parties must present an equal fraction of male and female candidates across the electoral districts. If the difference between female and male candidates exceeds 4% (48% females and 52% males, or the reverse), non-compliance with the gender parity rules results in a financial penalty. The financial penalty is computed as follows: "public funding provided to political parties based on the number of votes they receive in the first round of elections is reduced by a percentage equivalent to one-half of the difference between the total number of candidates of each sex, out of the total number of candidates" (Le Barbanchon and Sauvagnat (2022)). Notice that in France, voters vote for two candidates: the leading candidate, who, if elected, will become an MP, and the substitute. The last one replaces the top candidate only if he leaves office; potential reasons are being nominated to a function incompatible with the mandate of an MP (e.g., minister, secretary of state, elected as a mayor) or death/health issues. In most cases, the substitute never becomes an MP during the legislative term. After the approval of the law, in the 2002 legislative elections, women were 38.8% of

the leading candidates, compared to 23% in the 1997 legislative elections.

Ideological classification I use the official party labels provided by the Ministry of the Interior to classify candidates (including independent ones) and following Jolly et al. (2022) I classify candidates into six partisan families: far-left, left, liberal, right, far-right, and other. The last category refers to politicians who do not fall into any of these traditional ideological categories or do not classify themselves into any ideology.

In the rest of the paper, I refer to political orientation as the broader categories "left" (far-left and left) and "right" (liberal, right and far-right), unless specified otherwise. Essentially, I classify "left" parties as the ones that score less than 5 in the left-right dimension of Jolly et al. (2022) and "right" as the ones that score more than 6.[2] Parties classification is available in the Appendix Section A.

3.2 Political platforms

During the legislative campaign, individual candidates can emit one political platform (*trans.* profession de foi) before each election round.[3] The appendix provides one example in the Figure B.1. What is the traditional content of a political platform? The manifesto permits a candidate to present his program and ideas that he plans to commit to when elected. Since they represent an electoral district, candidates might run a campaign by focusing on national policies and/or addressing local issues. They can also run a more partisan program or personalize their campaign, focus on preferred policy topics, or, as expected, appeal to vote and criticize the opposition.

Candidates are responsible for printing these platforms, and the state can reimburse their costs if they gather at least 5% of the votes in one of the rounds (Electoral law, articles R39 and L216). An official local propaganda committee is responsible for mailing the manifestos to voters at least four days before the first round and three days before the second round (if it happens).[4]

According to the Ipsos - CEVIPOF 2022 Presidential electoral survey, the primary sources of candidates' information are: 38% of the electors follow the television, 15% the internet (many candidates share their platforms online), 13% newspapers, 12% the manifestos received in the mailbox, 6% use other sources around

---

[2]According to Jolly et al. (2022), all liberal parties in the sample have a left-right score between 6 and 7, in a spectrum between 0 and 10.

[3]Throughout the paper, I use the words "political platform" and "manifesto" interchangeably; in this paper, they are synonymous.

[4]These platforms are only mailed if they obey these rules: they must have a maximum size of 210x297 millimeters, weigh between 60 and 80 grams per square meter (Electoral law, article R29) (Le Pennec (2023)).

them, and 16% do not inform. Relatively to the 1988, 1993, and 1997 manifestos, the internet was yet to be widely available and used. Furthermore, per election, there are 577 constituencies and an average of 4079 candidates, making it impossible for television to give coverage to all candidates. Consequently, platforms are likely more important in legislative than presidential elections.

4 Data

4.1 Electoral data

Each dataset records the number of registered voters, abstentions, cast votes, valid and invalid votes, and the votes for each candidate in each electoral district. The electoral data for French elections comes from the Ministry of Interior.

4.2 Political platforms

Candidate manifestos for the 1988 and 1993 elections were digitized by the Archelec project (Gaultier-Voituriez (2016)).[5] Until 1993, the CEVIPOF collected manifestos each election with the government's support. Each departmental administration mailed the manifestos distributed in their district. Unfortunately, this practice finished in 1993, so manifestos between 2002 and 2012 are unavailable.

Platforms for the 1997 elections were digitized from the National Archives. For 2012, the French National Assembly website provides the political platforms of all elected politicians.[6]

For 2017 they come from several sources: during the campaign, the Ministry of Interior shares the manifestos submitted by the candidates on their website[7] and they were web-scraped by a non-profit organization called RegardsCitoyens[8]. In addition, I manually digitized missing manifestos at CEVIPOF (Sciences Po), the National Assembly website also shares the manifestos of all elected MPs[9], some missing manifestos were also found on several local news or candidates websites.

---

[5]They are available at https://archive.org/details/archiveselectoralesducevipof/

[6]They are available at: https://www.assemblee-nationale.fr/dyn/decouvrir-l-assemblee/histoire/barodet2/recueil-des-professions-de-foi-de-la-14eme-legislature.

[7]They are available at: https://programme-candidats.interieur.gouv.fr/

[8]They are available at: https://github.com/regardscitoyens/professions-foi-candidats/tree/master/documents/LG17

[9]They are available at: https://www.assemblee-nationale.fr/dyn/decouvrir-l-assemblee/histoire/barodet2/recueil-des-professions-de-foi-de-la-15eme-legislature

For 2022 and 2024, platforms come from the Ministry of Interior website and several local news and candidates websites.

The dataset comprises 32608 political platforms, 25808 for the first round, and 6800 for the second round. Optical character recognition (Tesseract) transformed these platforms from image to text.

4.3   Legislative work

I web-scraped the speeches from the Assemblée Nationale website covering the 1998-2022 period and transformed them into a novel dataset.[10][11]   I restrict my analysis to elected politicians, excluding presidents and vice-presidents of the Parliament. I eliminate procedural words in parliamentary speech because they appear frequently and their use is unlikely to be informative about group differences (Gentzkow, Shapiro, and Taddy (2019)).[12]   I also remove speeches with less than five words; these speeches tend to be minor reactions to an opponent's speech and are not informative in terms of group or topic. I aggregate speeches so that a document captures all speeches by a given speaker in one agenda of a plenary session; I remove aggregate speeches with less than 15 words; concise speeches are not ideal for detecting group differences or topics. The dataset includes a total of 155,207 documents.

Second, I web-scraped the National Assembly website to obtain the content of the written questions. These questions are directed to a minister to express the MP concern on a specific topic.[13] The dataset includes a total of 590,185 questions.

4.4   Campaign contributions

Data between 1993 and 2017 on campaign expenditures and contributions is from Bekkouche, Cagé, and Dewitte (2022). For 2022, I use data from the National Commission on Campaign Accounts and Political Financing (CNCCFP). In French parliamentary elections, candidates who receive at least 1% of votes in the first round must submit their campaign accounts to the French CNCCFP. This commission was created in 1990, so data before that date is unavailable.

---

[10]They are available at: https://www.assemblee-nationale.fr/dyn/15/comptes-rendus/seance

[11]I do not include the years 2001 and the end of the 11th legislature (2002) because the website during this period is designed differently and is difficult to web-scrape.

[12]I obtain the list of procedural phrases from the following websites: https://www2.assemblee-nationale.fr/decouvrir-l-assemblee/folder/lexique   and   https://www.assemblee-nationale.fr/connaissance/lexique.asp .

[13]They are available at: https://www2.assemblee-nationale.fr/recherche/questions

For each year, electoral district, and candidate, I observe for the whole campaign the total amount spent by the candidate, the total amount of contributions he received, and the amount of each type of contribution: contributions received from the candidate's political party, donations, and personal funds.

5 Methodology

5.1 Topic classification

In this paper, I study gender differences in policies advertised during political campaigns. However, policy topics in the manifestos are not classified as ex-ante, and candidates usually tend to advertise several policies. To overcome this challenge, I rely on topic modeling techniques to retrieve the topics and construct the outcomes of interest, specifically seeded Latent Dirichlet Allocation (seeded LDA).

I pre-processed platforms' content by removing capitalization, punctuation, stop words, and special characters. I tokenize documents at the single-word level and lemmatize each word using Spacy's French model. In addition, I restrict the vocabulary to words used by at least 1% and 50% of the platforms for the whole sample; in the case of the legislative debates and written questions, the minimum number is 0.5% since the number of documents is considerably larger.

LDA (Blei, Ng, and Jordan (2003)) is a generative probabilistic model based on the assumption that each document is a mixture of topics and that latent topics generate the words observed in the document of a corpus. LDA is an unsupervised method, while seeded LDA (Lu, Ott, Cardie, and Tsou (2011), Watanabe and Baturo (2023)) is a semi-supervised machine learning technique. Seeded LDA extracts these topics based on a prior 'seed' of selected words that capture the topic of interest. Watanabe and Baturo (2023) show that this method improves the inconsistency of topics that LDA generally produces.

The central tuning parameter of a LDA model is the number of topics K to be estimated. If K is too small, documents about different topics will be lumped together in the same estimated topic. If K is too large, documents that belong to the same topic are split. I calculate topic coherence in a simple LDA to estimate the number of topics for platforms. The highest value is at 8. I define the following eight topics: economy & employment, environment; health & education; security, justice & foreign policy; local politics, national politics, and the remaining two are other.

For legislative debates, the coherence score advises to use around 12 topics. Note that the number of documents is much higher than in the case of platforms

because the number of observations is much higher. Nonetheless, to ensure comparability with the platforms, I regrouped the topics. At the beginning of each legislative session, the President of the Assembly announces the works of the session; this means that each session tends to be about a specific topic. I classify a document as about a specific topic if the highest value refers to that topic.

I manually reviewed each retained word and assigned the most obvious words to their specific topic. In the Appendix section B, section G and section H I provide further details on the method and the seed words. In addition, to demonstrate the method's validity, I show the top words obtained for each topic in the platforms, legislative debates, and written questions.

5.1.1 Descriptive statistics

Table 1 presents descriptive statistics relative to gender differences in political platforms. I confirm statistically significant differences at 1% in all the characteristics analyzed in both the first and second rounds (except economy in the first round). Regarding gender differences in terms of topics, the most significant difference is on health & education, women give a higher prevalence than males to this topic by 13 percentage points in the first round. In contrast, they give less salience to national politics by 5 percentage points in the first and 13 in the second rounds. Nonetheless, it is premature to refer to gender differences in political campaigns, given that a female presence in a race is endogenous.

Table 1: Differences between female and male political candidates - Summary statistics

	1st Round			2nd Round		
	Female	Male	Difference	Female	Male	Different
% votes	17.62	16.90	0.72***	25.61	27.07	-1.46***
Number words	719.01	660.78	58.23***	658.53	540.22	118.31***
Economy % employment	0.18	0.18	0.00	0.16	0.14	0.02***
Environment	0.04	0.03	0.02***	0.04	0.03	0.02***
Health & education	0.17	0.04	0.13***	0.15	0.03	0.12***
Security & foreign policy	0.10	0.06	0.04***	0.09	0.05	0.05***
Local politics	0.24	0.15	0.09***	0.22	0.13	0.09***
National politics	0.14	0.19	-0.05***	0.23	0.35	-0.13***
Observations	1672	2335	4007	1650	2218	3868

5.2 Empirical strategy

To estimate the causal impact of gender on campaign strategies, I use a regression discontinuity design (RDD) and compare female candidates who narrowly passed to the second round (while a male was not barely eligible in that electoral district) with male candidates who narrowly passed to the second round (while a female was almost eligible). At the cutoff, the female presence is orthogonal to voters' and electoral district characteristics.

To estimate differences between female and male political platforms in the second round, I use a sharp regression discontinuity design and estimate the following equation:

$$Y_i = \beta_0 + \beta_1 T_i + \beta_2 R_i + \beta_3 R_i T_i + \beta_4 X_i + \mu_i \tag{1}$$

The treatment variable T is a dummy equal to one if it is a woman and 0 if it is a male in the second round. The running variable R is the qualifying margin of the candidate in the first round. Remember from Section 3.1, in races where only one or no candidate obtained 12.5% of the votes, only the two most voted candidates are eligible for the second round. In this situation, the running variable must be the difference between the most-voted woman and the second-most-voted man. The unit of observation is the candidate, and there is one observation per electoral district.

X_i is a vector of first-round independent variables that includes controls over the first round's dependent variables. The results do not differ if, in the alternative, the dependent variable is the difference between the outcome in the second and first rounds. I follow Calonico, Cattaneo, and Titiunik (2014) and use a nonparametric approach, fitting a local linear regression on each side of the threshold within an optimal bandwidth selected by the MSERD procedure. I use the mean squared error optimal bandwidths selection procedure proposed by Calonico et al. (2014) in its covariate-adjusted version (Calonico, Cattaneo, Farrell, and Titiunik (2019)). This data-driven procedure implies that bandwidth size varies with the outcome under consideration.[14] In X_i, I include several controls referred to below on the balancing tests. These controls are not necessary for identification but improve efficiency.

5.2.1 Internal validity

The validity of the RDD relies on the assumption that first-round candidates of a particular type (e.g., males) do not systematically sort on the right of the qual-

---

[14]I use the Stata package rdrobust (Calonico, Cattaneo, Farrell, and Titiunik (2017)).

ification threshold. Such manipulation is unlikely since France is a democracy, and international observers qualify the elections as pluralistic, competitive, and respectful of fundamental rights.[15] In addition, manipulation is difficult because it requires predicting the outcome of the first election stage with great accuracy. I test the assumption's validity using the McCrary (2008) test and check if there is a jump in the density of the running variable at the threshold. As Figure C.3 in the Appendix demonstrates, there is no jump at the margin. I also confirm the results with the Cattaneo, Jansson, and Ma (2018) test.

Another implication of the identifying assumption is that districts' characteristics are continuous at the threshold. I run balancing tests for: votes, turnout, number of candidates, number of enrolled voters, victory margin, number of candidates per ideology, and the sum of votes for left and right-wing candidates in the first round. Table C2 in the Appendix shows the sharp RDD estimates of the effect of having a woman in the second round in alternative to a man on first-round electoral district characteristics. Of the 12 regressions, the coefficient on the treatment variable is not statistically significant in any of them. The pre-treatment characteristics of the districts are balanced.

I also test for a jump in individual characteristics. A further concern is the confounding effects due to the ideology of candidates. Male and female candidates may systematically differ in their ideology. For instance, female candidates may be mainly from left-wing parties. Hence, the results obtained from estimating Equation (1) may be due to the ideology rather than the gender of the candidate. However, I do not find significant differences in the ideology of female and male candidates around the cutoff (see Table C3 in the Appendix), except for right-wing ideology at 10 percent. I also test whether female and male candidates differ in other individual characteristics: number of words in their platforms, incumbency status, and a dummy whether they ran in the past. I do not find any significant differences (Appendix Table C4).

The analysis can also be affected by endogenous sample selection. A potential concern is if a political platform is more observed for one gender than the other or when a woman is present (or not) in the second round. Column 1 of Table C4 shows that this is not the case, there is no significant jump in the probability of having a first-round manifesto available at the qualification threshold.

Di Tella et al. (2023) demonstrate empirically that candidates strategically adjust their platform to get closer to their opponent. It could be that the results are confounded because the opponents of female candidates are mostly from a spe-

---

[15]*France. Presidential Election 10 and 24 April 2022. ODIHR Election Assessment Mission Final Report* (2022).

cific ideology. In Table C5, I show the balancing tests regarding the opponents' ideology. None of the ideologies is statistically significant.

I conduct a general balance test by checking whether the candidate's predicted treatment status jumps at the threshold. I first regress actual treatment on the above-mentioned variables to predict treatment status. Figure 1 and Table C6 in the Appendix show that the point estimate is small and insignificant.

Figure 1: General balance test
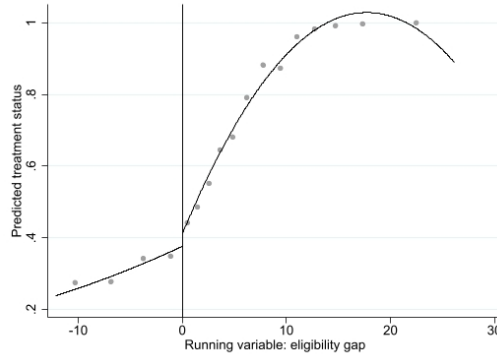
Notes. Dots represent the local averages of the predicted treatment status (vertical axis). Averages are calculated within quantile bins of the running variable (horizontal axis). The outcome is the candidate's predicted treatment status based on observable characteristics listed in the text. The treatment variable is a dummy equal to 1 if a woman qualifies for the second round. The sample is restricted to candidates included in the RDD sample as described in the text. The running variable is the difference between the most voted woman and the second-most-voted man.

Finally, Appendix Table C7 shows no discontinuity in the coverage of topics of the most-voted candidate in the first round. This provides reassuring evidence that the importance of specific topics in some electoral districts does not drive the results.

6 Results: Gender Differences in Political Platforms

In this section, I use the empirical strategy described above to study gender differences in the policy topics covered in the political platforms. Table 2 provides the formal estimates of the effects, while Figure D.4 in the Appendix provides the graphical analysis. I do not find significant differences in the coverage of topics, except for security & foreign policy. I conclude that women write 6.6 percentage

points more about security & foreign policy than males, topics that traditionally are associated with men. These results are not necessarily surprising. Previous experiments have shown that female candidates try to reverse gender stereotypes by portraying themselves as possessing stereotypical masculine traits (Huddy and Terkildsen (1993), Bauer (2017)).[16] These results are robust to other bandwidths and are available in Appendix Tables D8 and D9.

Table 2: Differences between female and male 2nd round political candidates

	(1) Economy & employment	(2) Environment	(3) Health & education	(4) Security & foreign policy	(5) Local politics	(6) National politics
Woman	-0.008	-0.001	-0.025	0.066**	-0.020	0.017
	(0.020)	(0.006)	(0.017)	(0.027)	(0.028)	(0.032)
Observations	836	836	836	836	836	836
Eff. number of obs	322	234	310	214	233	271
Robust p-value	0.828	0.727	0.147	0.024	0.475	0.517
Polyn. order	1	1	1	1	1	1
Bandwidth	4.740	3.202	4.415	2.929	3.151	3.832
Outcome mean	0.187	0.040	0.047	0.093	0.272	0.237

Standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest is a woman running in the second round as an alternative to a man. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure. The dependent variables take values between 0 and 1, and their computation is explained in Section 5.1. The mean gives the average outcome value for male candidates at the threshold.

6.1 Heterogeneity

6.1.1 Gender Differences and Voters' Gender Bias

I next test whether the gender differences in the topics covered in the political platforms are higher in electoral districts where we expect stronger voters' gender bias. To do so, I run a heterogeneity analysis based on the extent of gender discrimination in the electoral district, using two different proxies. First, I exploit the magnitude of the results in electoral districts that never elected a female MP. Second, I follow Le Barbanchon and Sauvagnat (2022), who find that the local gender earnings gap correlates with voters' attitudes toward women. I use the authors'

---

[16]Herrnson et al. (2003) refer several situations in the U.S. where women emphasized issues traditionally thought of as best handled by men to downplay differences perceived by voters between male and female candidates. "In her race for governor of California in 1990, Dianne Feinstein emphasized her support for the death penalty, and as a vice presidential candidate in 1984, Geraldine Ferraro accentuated her tough stance on crime."

gender bias estimation; they computed the residualized local earnings gaps after controlling for age, industry, and occupation.

Table 3 Panel A shows the results for electoral districts that never elected a female MP (at least until that election). In these districts, women talk more about security & foreign policy than males by 8.4 percentage points, statistically significant at a 1 percent level. The higher salience to this topic is compensated by a lower coverage of local politics (6.5 percentage points, statistically significant at 5 percent level) and health and education (3.9 percentage points, statistically significant at 10 percent). Panels B and C present the results separately for districts above and below the mean value of the residualized local earnings gap, respectively. Panel B shows that in districts where the gender wage gap is above average, women give a higher salience to topics of security & foreign policy by 13 percentage points in comparison to males. On the other hand, in districts where the gender wage gap is below average, such a significant difference does not exist. These results further support that female candidates responded to voters' bias rather than being driven by their intrinsic preferences. In districts where the gender wage gap is above average, the higher salience to security & foreign policy is compensated by a lower salience to national politics (by 20.9 percentage points) and health & education (by 3.8 percentage points, statistically significant at 1 percent).

6.2 Gender Differences and Ideology

Studying heterogeneity at the ideology level permits assessing the findings' external validity and whether the effects are specific to certain parties within the French elections. I present results for left and right-wing politicians in Panel A and B, respectively.

Table 4 shows that both left and right-wing women give more salience to security issues & foreign policy than their male colleagues of the same ideology, although the results are considerably stronger for right-wing women. Right-wing women also give less coverage to health & education than right-wing males by 6.5 percentage points and more coverage to economy & employment by 12.7 percentage points.

Table 3: Differences between female and male 2nd round political candidates, by district gender discrimination

	(1) Economy & employment	(2) Environment	(3) Health & education	(4) Security & foreign policy	(5) Local politics	(6) National politics
Panel A: Districts that never elected a woman						
Woman	-0.017	-0.006	-0.039*	0.084***	-0.065**	0.043
	(0.027)	(0.007)	(0.021)	(0.029)	(0.028)	(0.038)
Observations	567	567	567	567	567	567
Eff. number of obs	167	143	205	162	143	172
Robust p-value	0.411	0.313	0.067	0.008	0.022	0.189
Polyn. order	1	1	1	1	1	1
Bandwidth	3.265	2.699	4.012	3.122	2.699	3.385
Outcome mean	0.183	0.041	0.047	0.091	0.289	0.233
Panel B: Gender wage gap above average						
Woman	-0.006	0.009	-0.038***	0.130***	0.020	-0.209***
	(0.037)	(0.007)	(0.011)	(0.045)	(0.032)	(0.057)
Observations	419	419	419	419	419	419
Eff. number of obs	138	118	149	92	128	81
Robust p-value	0.967	0.227	0.019	0.005	0.658	0.000
Polyn. order	1	1	1	1	1	1
Bandwidth	4.211	3.585	4.755	2.750	3.932	2.330
Outcome mean	0.183	0.038	0.047	0.084	0.266	0.219
Panel C: Gender wage gap below average						
Woman	-0.022	-0.017***	-0.023*	0.022	-0.005	0.079**
	(0.020)	(0.006)	(0.013)	(0.028)	(0.032)	(0.033)
Observations	417	417	417	417	417	417
Eff. number of obs	125	121	131	111	140	129
Robust p-value	0.281	0.008	0.151	0.555	0.779	0.030
Polyn. order	1	1	1	1	1	1
Bandwidth	3.182	3.034	3.373	2.792	3.543	3.251
Outcome mean	0.184	0.040	0.047	0.084	0.267	0.233

In Panel A, the sample is restricted to districts that never elected a female MP. In Panel B, the sample is restricted to districts where the gender wage gap is above the mean, while Panel C restricts the sample to districts below the mean. Standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest is a woman running in the second round as an alternative to a man. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure. The dependent variables take values between 0 and 1, and their computation is explained in Section 5.1. The mean gives the average outcome value for male candidates at the threshold.

7 Potential Mechanisms

7.1 Campaign financing

If campaign funds are essential for a politician's success, then large differences in the amount of money that male and female candidates raise might impact how women campaign in their race. A lower amount of funding potentially translates into less funding to hire a team of advisors, implying that this team is smaller

Table 4: Differences between female and male 2nd round political candidates, by ideology

	(1) Economy & employment	(2) Environment	(3) Health & education	(4) Security & foreign policy	(5) Local politics	(6) National politics
	Panel A: Left-wing candidates					
Woman	-0.045	-0.000	0.007	0.014**	-0.001	0.001
	(0.032)	(0.008)	(0.008)	(0.006)	(0.031)	(0.046)
Observations	264	264	264	264	264	264
Eff. number of obs	78	93	97	90	121	97
Robust p-value	0.320	0.796	0.267	0.030	0.948	0.743
Polyn. order	1	1	1	1	1	1
Bandwidth	2.603	3.290	3.478	3.142	4.919	3.498
Outcome mean	0.191	0.040	0.049	0.091	0.260	0.237
	Panel B: Right-wing candidates					
Woman	0.127***	0.008	-0.065**	0.092**	-0.032	-0.063
	(0.035)	(0.008)	(0.032)	(0.044)	(0.039)	(0.042)
Observations	560	560	560	560	560	560
Eff. number of obs	91	108	156	122	109	144
Robust p-value	0.001	0.326	0.043	0.093	0.529	0.328
Polyn. order	1	1	1	1	1	1
Bandwidth	2.150	2.643	3.730	2.978	2.679	3.439
Outcome mean	0.191	0.042	0.048	0.092	0.287	0.239

In Panel A, the sample is restricted to far-left and left-wing candidates. In Panel B, the sample is restricted to liberal, right-wing, and far-right candidates. Standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest is a woman running in the second round as an alternative to a man. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure. The dependent variables take values between 0 and 1, and their computation is explained in Section 5.1. The mean gives the average outcome value for male candidates at the threshold.

and/or of a lower quality. Consequently, women might receive lower-quality advice on the topics of the campaign, what words to choose, and how to personalize their message. Moreover, given that I also have data on funding from the party, it also permits me to understand if there is party bias against women, and part of the results can be explained by less/more support from the party.

I test whether there are gender differences in campaign financing in France that potentially can explain gender differences in political platforms. I estimate these differences by employing the sharp regression discontinuity design described in Section 5.5. As shown in Table 5, I find no significant gender differences in campaign expenditures or contributions.

Table 5: Gender differences in campaign financing

	(1) Total expenditures	(2) Total revenues	(3) Party contribution	(4) Private donations	(5) Personal contribution
Woman	-0.062	-0.047	-0.048	0.041	-0.045
	(0.198)	(0.208)	(0.105)	(0.205)	(0.084)
Observations	1135	1135	1132	1132	1132
Eff. number of obs	436	443	337	356	470
Robust p-value	0.964	0.951	0.500	0.665	0.737
Polyn. order	1	1	1	1	1
Bandwidth	4.212	4.282	3.084	3.314	4.692
Outcome mean	0.849	0.907	0.123	0.288	0.389

Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest is a woman running in the second round instead of a man. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure. Each outcome uses the number of registered voters as the denominator. The mean gives the average outcome value for male candidates at the threshold.

7.2 Are candidates just responding to their opponents?

The results might simply be explained by the fact that women are responding to male proposals, as they are more responsive to their opponents. I test this hypothesis by estimating how similar candidates' manifestos are to the ones of their opponents, using cosine similarity. For any given document vectors x_i and x_j, the cosine similarity is the normalized dot product between the vectors:

$$cosine(x_i, x_j) = \frac{x_i \cdot x_j}{||x_i||||x_j||} \tag{2}$$

Each document contains eight vectors, and they refer to the estimated probability from the LDA of a document talking about each topic.

I build two measures. The first is the similarity between the candidate's manifesto in the second round and her opponent's manifesto in the first round. This measure permits understanding whether candidates in the second round are responding to the policy proposals of their opponent in the first round. The second measure is the similarity between the candidate's manifesto and her opponent in the second round. Di Tella et al. (2023) show that candidates converge to their opponents in ideology and rhetorical complexity; therefore, the results might be an

implication of the median voter theorem.[17]

Table 6 presents the results. In column (1), the dependent variable is a dummy variable, whether the candidate refers to her opponent's name in the platform of the second round or not. I do not find any statistically significant differences between women and men in referring to their opponent's names, which is a first indication if they spend part of their manifesto responding to or criticizing their opponent. In column (2), the dependent variable is the first similarity measure; I do not find that women tend to converge more than males to their opponent's platform in the first round. Column (3) provides a similar conclusion: women do not converge more than males to their opponent's platform in the second round. These results imply that the gender differences found in Section 6 cannot be explained by the fact that women are more responsive or converge more than males to their opponents.

Table 6: Gender differences in convergence to the opponent

	(1)	(2)	(3)
	Opponent name	Similar to opponent 1st round	Similar to opponent 2nd round
Woman	-0.006	-0.069	0.036
	(0.049)	(0.048)	(0.053)
Observations	723	725	723
Eff. number of obs	270	291	261
Robust p-value	0.935	0.174	0.364
Polyn. order	1	1	1
Bandwidth	4.176	4.820	4.040
Outcome mean	0.224	0.508	0.568

Column (1) is a dummy variable equal to one if the candidate cites at least once her opponent name, 0 otherwise. Column (2) is the cosine similarity between the candidate's platform in the second round and the opponent's platform in the first round. Column (3) is the cosine similarity between the candidate's platform in the second round and the opponent's platform in the second round. Standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest is a woman running in the second round as an alternative to a man. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure. The mean gives the average outcome value for male candidates at the threshold.

---

[17]Di Tella et al. (2023) do not find convergence at the level of the topics.

7.3 Electorate Policy Preferences

Another potential explanation for the results is that women give more salience to security & foreign policy issues simply because they are more responsive than males to voters' concerns. Unfortunately, there are no surveys representative at the electoral district level that provide information about voters' preferences. However, security & foreign policy issues were never among the most important issues referred by voters in the French Electoral Studies, who are representative at the national level. In 2007, unemployment and social inequalities were the two most important concerns for voters; 13% of the electorate referred to immigration or delinquency as their most important issue in the election. In 2012, almost 44% of the voters referred to the labor market situation as the most pressing issue, while only 4.7% of voters suggested issues related to security and foreign policy. In 2017, 48.1% of the voters referred unemployment as the most important issue in the country. While in 2022, the most important issues were social security and environmental protection, 18% of the voters referred to security and foreign policy issues.

7.4 Do women strategically adapt to the gender of their opponent?

Women might strategically choose what topics to focus on in their political campaign, depending on the gender of their opponent. As voters have formed beliefs about the ability of candidates to address specific topics depending on their gender, women might strategically change the coverage of certain topics to signal their ability.

I study whether the gender differences found in Table 2 are partly explained by the gender of the candidates' opponent. I start by studying whether women change the coverage of specific topics if they compete against a man instead of a woman. To explore this question causally, I implement a sharp regression discontinuity design. I restrict my sample to races where at least one woman reached the second round, and focus on the most voted woman in an electoral district who passed to the second round. I try to understand how the most voted woman in the first round behaves if she competes against a man that barely passed to the second round instead of competing against a woman that was barely eligible. Therefore, the running variable must be the difference between the most voted man and the second most voted woman; the treatment group is women who compete against a man, and the counterfactual is women who compete against another woman.

The validity of the RDD relies on the key assumption that first-round candi-

dates of a particular type (e.g., female candidates) do not systematically sort on the right of the qualification threshold. I implement the tests proposed by Mc-Crary (2008) and Cattaneo et al. (2018) and verify that there is no discontinuity in the density of the running variable at the threshold (Appendix Figure E.5).

The main implication of the identifying assumption is that electoral districts' characteristics are continuous at the threshold. I run balancing tests for first-round election characteristics in the same spirit of Section 5.5.1. Tables E10 to E13 in the Appendix show the results. Considering 25 balance tests, only three covariates are statistically significant at 10% level. In addition, I test whether women who compete against a man in the second round differ in terms of topics in the first round from women who compete against a woman; only national politics is statistically significant at 10%.

Table 7 shows how women strategically adapt to the gender of their opponent. When women run against men, they give more salience to security & foreign policy by 5.2 percentage points compared to when they compete against women. These results are statistically significant at 10 percent, and with a larger bandwidth, where the number of observations is not small, they are significant at 5 percent. They also provide information about the external validity of the results. In Section 6, I find that women give more coverage to security & foreign policy than males. In the first RDD, women enter the race as *runners-up*, they are competing against a man and were barely eligible for the second round. In this RDD, women are the top runners and are competing against a man who was barely eligible for the second round. In both situations, I find that women give salience to the topics of security & foreign policy.

7.5   Do men strategically adapt to the gender of their opponent?

As women strategically adapt to the gender of their opponent, the same can be expected from men. The gender differences in the topics covered found in Section 6 might be explained not just by the strategic behavior of women but also by men. I study whether men change the coverage of specific topics if they compete against a woman instead of another man. I provide a causal explanation for this answer and implement a sharp regression discontinuity design in the spirit of the previous subsection. I restrict my sample to races where at least one man reached the second round. I analyze whether the most voted man in the first round changes the coverage of the topics in the second round if he competes against a woman who barely reached the second round in contrast to competing with a barely eligible man. Thus, the running variable is the difference between the most-voted woman

Table 7: Impact of a marginal presence of a man on female candidates in the 2nd round

	(1) Economy & employment	(2) Environment	(3) Health & education	(4) Security & foreign policy	(5) Local politics	(6) National politics
Woman competing	-0.023**	-0.005	0.049	0.052*	-0.045	-0.020
against a man	(0.011)	(0.006)	(0.032)	(0.027)	(0.030)	(0.028)
Observations	498	498	498	498	498	498
Eff. number of obs	145	153	180	140	150	174
Robust p-value	0.092	0.347	0.122	0.116	0.145	0.557
Polyn. order	1	1	1	1	1	1
Bandwidth	2.965	3.113	3.692	2.891	3.008	3.597
Outcome mean	0.153	0.043	0.207	0.123	0.243	0.152

Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value and ***, ** and * indicate significance at 1, 5 and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the most voted woman in the first round competes against a man in the second round, 0 if she competes against a woman. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure. The dependent variables take values between 0 and 1, and their computation is explained in Section 5.1. The mean gives the average outcome value for women competing against a woman at the threshold.

and the second most-voted man, implying that a treatment group is a man who competes against a woman, and a counterfactual is a man who competes against another man.

As before, I demonstrate the validity of the RDD. I show in the Appendix that there is no discontinuity in the density of the running variable at the threshold (Figure E.6). In addition, I run again the same balancing tests (Appendix Tables E17-E21). Considering 31 balance tests, I find that 28 are not statistically significant, while the number of far-left candidates, number of left, and a dummy whether the candidate is left-wing are statistically significant at 5, 10, and 1 percent, respectively. Finally, I perform a general balance test that evaluates whether candidates predicted treatment status jumps at the threshold; the point estimate in Table E22 in the Appendix is small and non-significant.

Table 8 provides the results, testing the hypothesis of whether men change the salience of specific topics when they compete against a woman instead of a man. The topics economy & employment, and national politics are statistically significant at 10 percent, while there are no statistically significant differences in other topics. If I exclude as controls the unbalanced covariates, I do not find any statistically significant results (Appendix Table E23). I also test the robustness of the results to a larger bandwidth; Table E24 in the Appendix shows no statistically significant differences between men competing against a woman and men com-

peting against a candidate of the same sex. Therefore, these results demonstrate that male candidates do not adapt to the gender of their opponent. Women give a higher salience to security & foreign policy than males, and this difference can only be explained by the strategic decisions of women.

Table 8: Impact of a marginal presence of a woman on male candidates in the 2nd round

	(1)	(2)	(3)	(4)	(5)	(6)
	Economy & employment	Environment	Health & education	Security & foreign policy	Local politics	National politics
Man competing	0.025*	0.004	0.004	0.013	0.018	-0.054*
against a woman	(0.014)	(0.004)	(0.019)	(0.020)	(0.022)	(0.030)
Observations	802	802	802	802	802	802
Eff. number of obs	241	269	287	317	363	275
Robust p-value	0.049	0.301	0.830	0.607	0.465	0.068
Polyn. order	1	1	1	1	1	1
Bandwidth	2.620	2.964	3.105	3.512	4.138	3.000
Outcome mean	0.154	0.033	0.120	0.066	0.215	0.255

Standard errors are in parenthesis. Statistical significance is computed based on the robust p-value and ***, ** and * indicate significance at 1, 5 and 10, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest is a woman running in the second round in alternative to a man. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.The dependent variables take values between 0 and 1, and their computation is explained in Section 5.1. The mean gives the average value of the outcome for men competing against a man at the threshold.

7.6 Are women behaving strategically during elections or also signaling their type as MPs?

Political campaigns can effectively reduce the problem of asymmetric information, as candidates can send signals to voters, informing them about their types and aiming to adjust voters' beliefs. How reliable is this information? If candidates provide reliable information, political campaigns permit to choose the voters' preferred candidates effectively; in this situation, we should expect women to give more salience to security & foreign policy than males both during political campaigns and after being elected. On the other hand, if candidates do not provide reliable information, adverse selection emerges.

In this section of the paper, I study how reliable the information that candidates provide during elections is. I compare female and male candidates before and after elections in terms of topics; specifically, I study what topics the candidates focus on during legislative debates and in the written questions sent to the ministers

(after being elected) and the topics that these candidates have focused on in their platforms during campaigns. Table 9 provides a summary statistics. Columns (1) and (2) give the mean results for females and males, respectively, and column (3) the t-test of how much they differ. Elected women, on average, give more salience to health & education than males during political campaigns and after being elected; however, although they considerably give more coverage to security & foreign policy during campaigns, this difference disappears in parliamentary work. I also compare elected female and male candidates eligible for the second round with a margin below 5 percentage points; this sample is more similar to the one used in Section 6. These female politicians also gave significantly more coverage to security & foreign policy than males during political campaigns, but this difference disappears once elected. Once elected, women focus more on health & education, both in legislative debates and written questions.

The results above are merely correlations. I also provide causal results on whether elected women and men differ in the policy topics focused on. I adopt a standard sharp regression discontinuity design to estimate the causal impact of gender on participation in legislative work and the topics focused on. Focusing on mixed-gender elections, I use the female margin of victory as a forcing variable in the sharp RD design. I compute the female margin of victory as the difference in the vote share of the female and the male candidates relative to the share of votes obtained by both. This method has been widely used in previous research (e.g., Bhalotra, Clots-Figueras, and Iyer (2018), Casarico et al. (2022), Chauvin and Tricaud (2023)).[18]

Table 10 presents the results. In Panel A, I show the results regarding legislative debates. I find that elected women substantially talk more than males on health & education by 11.6 percentage points, statistically significant at 1 percent. There are no statistically significant differences in the other topics, including in security & foreign policy, or in the number of debates that they participate in. These results might be driven by party influence; since health & education are stereotyped as female topics, the party might force women to participate in debates about these topics, as women are perceived as more legitimate than males to talk about them. I also check gender differences in terms of the written question sent to the members of the Government. Party influence is lower in this case since they tend to refer to

---

[18]I perform the standard internal validity tests. I do not find any evidence of manipulation at the threshold (Figure E.7 in the Appendix). I also conducted 22 balance tests to bring empirical support to the identifying assumption that districts' characteristics are continuous at the threshold (Appendix Tables E25 and E26). Two are statistically significant at 1%: being left and right-wing. Controlling or not controlling for these covariates does not substantially change the results (Appendix Table E28).

Table 9: Differences between elected female and males, during political campaigns and after elections

	All elected politicians			Elected politicians who barely passed to the 2nd round		
	(1)	(2)	(3)	(4)	(5)	(6)
	Female	Male	Difference	Female	Male	Difference
Panel A: Legislative debates						
Number debates	51.65	46.17	5.49*	63.09	54.09	9.01
Economy & employment	0.16	0.17	-0.01	0.17	0.18	-0.01
Environment	0.14	0.14	0.00	0.16	0.17	-0.00
Health & education	0.16	0.10	0.07***	0.18	0.13	0.05**
Security & foreign policy	0.21	0.22	-0.01	0.27	0.24	0.03
Local politics	0.06	0.08	-0.02***	0.06	0.07	-0.00
Panel B: Written questions						
Number questions	73.13	111.94	-8.81***	57.00	68.94	-11.94
Economy & employment	0.27	0.34	-0.07***	0.21	0.27	-0.06**
Environment	0.22	0.20	0.02**	0.24	0.25	-0.01
Health & education	0.29	0.24	0.05***	0.30	0.25	0.05**
Security & foreign policy	0.24	0.24	0.00	0.25	0.24	0.02
Local politics	0.07	0.07	0.00	0.07	0.06	0.01
Panel C: Platforms 2nd round						
Number words	721.72	686.82	34.90***	708.36	741.63	-33.27
Economy & employment	0.14	0.15	-0.01	0.14	0.12	0.02
Environment	0.04	0.03	0.01***	0.04	0.03	0.01
Health & education	0.18	0.10	0.08***	0.14	0.24	-0.10**
Security & foreign policy	0.11	0.07	0.03***	0.14	0.08	0.06**
Local politics	0.25	0.25	0.00	0.29	0.27	0.02
Panel D: Platforms 1st round						
Number words	782.16	772.57	9.58	759.61	826.20	-66.59
Economy & employment	0.16	0.18	-0.02***	0.16	0.13	0.02
Environment	0.04	0.03	0.01***	0.05	0.03	0.02***
Health & education	0.19	0.11	0.08***	0.14	0.25	-0.11**
Security & foreign policy	0.11	0.08	0.03***	0.15	0.09	0.06**
Local politics	0.29	0.29	0.00	0.33	0.29	0.04
Observations	580	1414	1994	54	321	375

Columns (1) and (2) include the mean values for all elected politicians. Columns (4) and (5) include the mean values for all elected politicians who were eligible to the second round with a margin below 5 percentage points. The margin is the difference between the most voted woman and the second-most-voted man, as in Section 5.5.

local issues referred by their constituents. Panel B shows that women work more on health & education (by 6.2 percentage points, statistically significant at 1 percent). In addition, they work less on the economy & employment and write fewer questions. These results are robust to smaller and larger bandwidths (Appendix Tables E29 and E30, respectively).

These results conform with the literature on gender differences in elected politi-

cians (Hessami and da Fonseca (2020), Lippmann (2021)). However, they contrast with the findings about gender differences during political campaigns. The results imply that women's signals during political campaigns are unreliable and do not permit voters to identify their actual type.

Table 10: Differences between female and male MPs during parliamentary work - legislative debates

	(1) Number	(2) Economy & employment	(3) Environment	(4) Health & education	(5) Security & foreign policy	(6) Local politics
Panel A: Legislative debates						
Woman	-6.645	-0.006	-0.019	0.116***	-0.011	-0.019
	(6.562)	(0.029)	(0.026)	(0.030)	(0.030)	(0.021)
Observations	1334	1334	1334	1334	1334	1334
Eff. number of obs	498	722	814	492	778	531
Robust p-value	0.313	0.786	0.604	0.000	0.641	0.413
Polyn. order	1	1	1	1	1	1
Bandwidth	4.400	6.727	7.683	4.322	7.324	4.787
Outcome mean	49.225	0.220	0.159	0.148	0.267	0.092
Panel B: Written questions						
Woman	-74.038**	-0.053***	-0.036	0.062***	0.020	-0.020
	(35.935)	(0.017)	(0.023)	(0.022)	(0.025)	(0.017)
Observations	1451	1451	1451	1451	1451	1451
Eff. number of obs	666	791	646	819	736	680
Robust p-value	0.075	0.006	0.133	0.018	0.475	0.375
Polyn. order	1	1	1	1	1	1
Bandwidth	5.563	6.799	5.340	7.034	6.321	5.648
Outcome mean	133.090	0.329	0.202	0.248	0.240	0.072

Standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest is a woman being elected as an alternative to a man. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure. In columns (2)-(6), the dependent variable takes values between 0 and 1, and the computation is explained in Section 5.1. The mean gives the average outcome value for elected males at the threshold.

8   Conclusion

Leadership has been predominantly a male prerogative in politics. Although women have gained increased visibility as politicians, they remain in minority in leadership positions. To explain this phenomenon, previous research has centered on the idea of a "glass ceiling" - voters discriminate women against men (De Paola et al. (2010), Le Barbanchon and Sauvagnat (2022), Eyméoud and Vertier (2023)). This paper provides new evidence that women and men make different decisions when they campaign and that voters' gender bias permits to explain these differences.

I provide the first causal evidence of gender differences in political campaigns.

I explore the case of French legislative elections for which individual political platforms are available from 1981 to 1997 and 2017-22 for both first and second-round races.

My research, focused on isolating the causal impact of gender on campaigns, highlights the pressing need for policies that account for gender discrimination. By implementing a regression discontinuity design and comparing female candidates who narrowly were eligible for the second round against a male candidate, I causally show that women give the same salience to all topics in comparison to males, except for security and foreign policy. These results, explained by voters' gender bias, are stronger in districts that never elected previously a woman or where the gender wage gap is above average. This underscores the importance of addressing gender discrimination through our educational and institutional policies.

I explore whether women consider the gender of their opponents' when writing their platforms. I compare women who competed against a barely eligible man with women who competed against a woman who barely passed to the second round instead of a man. When women run against a man, they give more coverage to security & foreign policy topics. Women strategically adapt to their opponents and adopt stereotyped male traits to account for voters' bias.

On the contrary, evidence that male politicians adapt to the presence of a woman in the second round is less clear. I conduct a regression discontinuity design and compare races where a woman was barely eligible for the second round with races where she was not present. I do not find robust evidence that men adapt their campaign to the gender of their opponent.

Once elected, women participate in debates as much as men, a task that requires good rhetorical abilities. However, they no longer focus significantly more than males on security and foreign policy but instead on health and education. Again, I prove that focusing on stereotyped male topics during campaigns is a strategic behavior to prevent voter discrimination.

The paper provides several avenues for future research. First, it would be exciting to see how the results extend to other countries, especially with more conservative norms or electoral systems. Future research should also try to understand how women can overcome voter bias. Understanding how voters react to gender in campaign information and how this reflects in votes is crucial to advising future female politicians and helping them be elected.

References

Alesina, A. (1988). Credibility and policy convergence in a two-party system with rational voters. *The American Economic Review*, *78*(4), 796–805.

Ash, E., Krümmel, J., & Slapin, J. B. (2023). Better to be jeered than ignored? gender and reactions during parliamentary debates. *Center for Law & Economics Working Paper Series*, *2*.

Bagues, M., & Campa, P. (2021). Can gender quotas in candidate lists empower women? evidence from a regression discontinuity design. *Journal of Public Economics*, *194*, 104315.

Barber, M., Butler, D. M., Preece, J., et al. (2016). Gender inequalities in campaign finance. *Quarterly Journal of Political Science*, *11*(2), 219–48.

Bauer, N. M. (2017). The effects of counterstereotypic gender strategies on candidate evaluations. *Political Psychology*, *38*(2), 279–295.

Bekkouche, Y., Cagé, J., & Dewitte, E. (2022). The heterogeneous price of a vote: Evidence from multiparty systems, 1993–2017. *Journal of Public Economics*, *206*, 104559.

Besley, T., & Coate, S. (1997). An economic model of representative democracy. *The quarterly journal of economics*, *112*(1), 85–114.

Bhalotra, S., & Clots-Figueras, I. (2014). Health and the political agency of women. *American Economic Journal: Economic Policy*, *6*(2), 164–197.

Bhalotra, S., Clots-Figueras, I., & Iyer, L. (2018). Pathbreakers? women's electoral success and future political participation. *The Economic Journal*, *128*(613), 1844–1878.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *Journal of machine Learning research*, *3*(Jan), 993–1022.

Booth, A., & Nolen, P. (2012). Choosing to compete: How different are girls and boys? *Journal of Economic Behavior & Organization*, *81*(2), 542–555.

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, *17*(2), 372–404.

Calonico, S., Cattaneo, M. D., Farrell, M. H., & Titiunik, R. (2019). Regression discontinuity designs using covariates. *Review of Economics and Statistics*, *101*(3), 442–451.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*(6), 2295–2326.

Casarico, A., Lattanzio, S., & Profeta, P. (2022). Women and local public finance. *European Journal of Political Economy*, *72*, 102096.

Cattaneo, M. D., Jansson, M., & Ma, X. (2018). Manipulation testing based on density discontinuity. *The Stata Journal*, *18*(1), 234–261.

Chattopadhyay, R., & Duflo, E. (2004). Women as policy makers: Evidence from a randomized policy experiment in india. *Econometrica*, *72*(5), 1409–1443.

Chauvin, J. P., & Tricaud, C. (2023). *Gender and electoral incentives: Evidence from crisis response*. Centre for Economic Policy Research.

Croson, R., & Gneezy, U. (2009). Gender differences in preferences. *Journal of Economic literature*, *47*(2), 448–474.

Curini, L., & Vignoli, V. (2021). Committed moderates and uncommitted extremists: Ideological leaning and parties' narratives on military interventions in italy. *Foreign Policy Analysis*, *17*(3), orab016.

Datta Gupta, N., Poulsen, A., & Villeval, M. C. (2013). Gender matching and competitiveness: Experimental evidence. *Economic Inquiry*, *51*(1), 816–835.

De Paola, M., Scoppa, V., & Lombardo, R. (2010). Can gender quotas break down negative stereotypes? evidence from changes in electoral rules. *Journal of Public Economics*, *94*(5-6), 344–353.

Di Tella, R., Kotty, R., Le Pennec, C., & Pons, V. (2023). Keep your enemies closer: Strategic platform adjustments during u.s. and french elections. *NBER Working Paper*(31503).

Djourelova, M. (2023). Persuasion through slanted language: Evidence from the media coverage of immigration. *American Economic Review*, *113*(3), 800–835.

Djourelova, M., Durante, R., & Martin, G. (2021). The impact of online competition on local newspapers: Evidence from the introduction of craigslist.

Downs, A. (1957). An economic theory of political action in a democracy. *Journal of political economy*, *65*(2), 135–150.

Eagly, A. H., & Karau, S. J. (2002). Role congruity theory of prejudice toward female leaders. *Psychological review*, *109*(3), 573.

Eyméoud, J.-B., & Vertier, P. (2023). Gender biases: evidence from a natural experiment in french local elections. *Economic Policy*, *38*(113), 3–56.

Ferreira, F., & Gyourko, J. (2014). Does gender matter for political leadership? the case of us mayors. *Journal of Public Economics*, *112*, 24–39.

Fraccaroli, N., & Pizzigolotto, A. (2021). Credit shocks and populism.

France. presidential election 10 and 24 april 2022. odihr election assessment mission final report. (2022).

Funk, P., & Gathmann, C. (2015). Gender gaps in policy making: Evidence from direct democracy in switzerland. *Economic Policy*, *30*(81), 141–181.

Gallagher, R. J., Reing, K., Kale, D., & Ver Steeg, G. (2017). Anchored correlation explanation: Topic modeling with minimal domain knowledge. *Transactions*

of the Association for Computational Linguistics, *5*, 529–542.

Gaultier-Voituriez, O. (2016). Archelec, les archives électorales de la v e république, du papier au numérique. *Histoire@ Politique*, *3*, 213–220.

Gentzkow, M., Shapiro, J. M., & Taddy, M. (2019). Measuring group differences in high-dimensional choices: method and application to congressional speech. *Econometrica*, *87*(4), 1307–1340.

Gneezy, U., Leonard, K. L., & List, J. A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, *77*(5), 1637–1664.

Granzier, R., Pons, V., & Tricaud, C. (2023). Coordination and bandwagon effects: How past rankings shape the behavior of voters and candidates. *American Economic Journal: Applied Economics*, *15*(4), 177–217.

Hansen, S., McMahon, M., & Prat, A. (2018). Transparency and deliberation within the fomc: A computational linguistics approach. *The Quarterly Journal of Economics*, *133*(2), 801–870.

Herrnson, P. S., Lay, J. C., & Stokes, A. K. (2003). Women running "as women": Candidate gender, campaign issues, and voter-targeting strategies. *The Journal of Politics*, *65*(1), 244–255.

Hessami, Z., & da Fonseca, M. L. (2020). Female political representation and substantive effects on policies: A literature review. *European Journal of Political Economy*, *63*, 101896.

Huddy, L., & Terkildsen, N. (1993). Gender stereotypes and the perception of male and female candidates. *American journal of political science*, 119–147.

Jolly, S., Bakker, R., Hooghe, L., Marks, G., Polk, J., Rovny, J., . . . Vachudova, M. A. (2022). Chapel hill expert survey trend file, 1999–2019. *Electoral studies*, *75*, 102420.

Kanthak, K., & Woon, J. (2015). Women don't run? election aversion and candidate entry. *American journal of political science*, *59*(3), 595–612.

Lawless, J. L. (2004). Women, war, and winning elections: Gender stereotyping in the post-september 11th era. *Political Research Quarterly*, *57*(3), 479–490.

Le Barbanchon, T., & Sauvagnat, J. (2022). Electoral competition, voter bias, and women in politics. *Journal of the European Economic Association*, *20*(1), 352–394.

Le Pennec, C. (2023). Strategic campaign communication: Evidence from 30,000 candidate manifestos. *The Economic Journal*, uead082.

Lippmann, Q. (2021). Are gender quotas on candidates bound to be ineffective? *Journal of Economic Behavior & Organization*, *191*, 661–678.

Lu, B., Ott, M., Cardie, C., & Tsou, B. K. (2011). Multi-aspect sentiment analysis

with topic models. In *2011 ieee 11th international conference on data mining workshops* (pp. 81–88).

Markowsky, E., & Beblo, M. (2022). When do we observe a gender gap in competition entry? a meta-analysis of the experimental literature. *Journal of Economic Behavior & Organization*, *198*, 139–163.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, *142*(2), 698–714.

Niederle, M., & Vesterlund, L. (2007). Do women shy away from competition? do men compete too much? *The quarterly journal of economics*, *122*(3), 1067–1101.

Niederle, M., & Vesterlund, L. (2011). Gender and competition. *Annu. Rev. Econ.*, *3*(1), 601–630.

Osborne, M. J., & Slivinski, A. (1996). A model of political competition with citizen-candidates. *The Quarterly Journal of Economics*, *111*(1), 65–96.

Sutter, M., Bosman, R., Kocher, M. G., & van Winden, F. (2009). Gender pairing and bargaining—beware the same sex! *Experimental Economics*, *12*, 318–331.

Watanabe, K., & Baturo, A. (2023). Seeded sequential lda: A semi-supervised algorithm for topic-specific analysis of sentences. *Social Science Computer Review*, 08944393231178605.

Watanabe, K., Xuan-Hieu, P., & Watanabe, M. K. (2023). Package 'seededlda'.

Weigel, J. L. (2020). The participation dividend of taxation: How citizens in congo engage more with the state when it tries to tax them. *The Quarterly Journal of Economics*, *135*(4), 1849–1903.

A Ideological classification

I allocate candidates into seven political orientations (far-left, left, liberal, right, far-right and other). I use the party classifications from the Chapel Hill Expert Survey (CHES) and when missing from the ParlGov and Granzier, Pons, and Tricaud (2023). I consider the party positioning on a scale between 0 (left) and 1 (right) and family classification.

1981 Parliamentary Elections	
Political label	**Political orientation**
Parti Communiste Français	Far-left
Divers Droite	Right
Divers Gauche	Left
Ecologistes	Left
Extrême Droite	Far-right
Extrême Gauche	Far-left
Indépendants	Other
Non Classés	Other
Rassemblement pour la République	Right
Socialistes	Left
Union pour la Démocratie Française	Right

1988 Parliamentary Elections	
Political label	**Political orientation**
Parti Communiste Français	Far-left
Divers Droite	Right
Ecologistes	Left
Extrême Droite	Far-right
Extrême Gauche	Far-left
Front National	Far-right
Majorité Présidentielle	Left
Radicaux de Gauche	Left
Régionalistes	Other
Rassemblement pour la République	Right
Socialistes	Left
Union pour la Démocratie Française	Right

1993 Parliamentary Elections

Political label	Political orientation
Parti Communiste Français	Far-left
Divers	Other
Divers Droite	Right
Extrême Droite	Far-right
Extrême Gauche	Far-left
Front National	Far-right
Génération Ecologie	Left
Majorité Présidentielle	Left
Radicaux de Gauche	Left
Régionalistes	Other
Rassemblement pour la République	Right
Parti Socialiste	Left
Union pour la Démocratie Française	Right
Les Verts	Left

1997 Parliamentary Elections

Political label	Political orientation
Parti Communiste Français	Far-left
Divers	Other
Divers Droite	Right
Divers Gauche	Left
Ecologistes	Left
Extrême Droite	Far-right
Extrême Gauche	Far-left
Front National	Far-right
Mouvement Pour la France	Far-right
Parti Radical Socialiste	Left
Rassemblement pour la République	Right
Socialistes	Left
Union pour la Démocratie Française	Right

2002 Parliamentary Elections	
Political label	**Political orientation**
Communistes	Far-left
Chasse, Pêche, Nature et Traditions	Right
Divers	Other
Démocratie Libérale	Right
Divers Droite	Right
Divers Gauche	Left
Ecologistes	Left
Extrême Droite	Far-right
Extrême Gauche	Far-left
Front National	Far-right
Ligue Communiste Révolutionnaire	Far-left
Lutte Ouvrière	Far-left
Mouvement des Citoyens	Left
Mouvement National Républicain	Far-right
Mouvement pour la France	Right
Pôle Républicain	Left
Radicaux de Gauche	Left
Régionalistes	Other
Rassemblement pour la France	Right
Socialistes	Left
Union pour la Démocratie Française	Liberal
Union pour un Mouvement Populaire	Right
Les Verts	Left

2007 Parliamentary Elections	
Political label	**Political orientation**
Communistes	Far-left
Chasse, Pêche, Nature et Traditions	Right
Divers	Other
Divers Droite	Right
Divers Gauche	Left
Ecologistes	Left
Extrême Droite	Far-right
Extrême Gauche	Far-left
Front National	Far-right
Majorité Présidentielle	Right
Mouvement pour la France	Right
Radicaux de Gauche	Left
Régionalistes	Other
Rassemblement pour la France	Right
Socialistes	Left
Union pour la Démocratie Française - Mouvement Démocrate	Liberal
Union pour un Mouvement Populaire	Right
Les Verts	Left

2012 Parliamentary Elections	
Political label	**Political orientation**
Alliance Centriste	Liberal
Autres	Other
Centre pour la France	Liberal
Communistes	Far-left
Divers Droite	Right
Divers Gauche	Left
Ecologistes	Left
Extrême Droite	Far-right
Extrême Gauche	Far-left
Front de Gauche	Far-left
Front National	Far-right
Nouveau Centre	Liberal
Parti Radical	Right
Radicaux de Gauche	Left
Régionalistes	Other
Socialistes	Left
Union pour un Mouvement Populaire	Right
Europe Ecologie - Les Verts	Left

2017 Parliamentary Elections

Political label	Political orientation
Parti Communiste Français	Far-left
Debout la France	Far-right
Divers	Other
Divers Droite	Right
Divers Gauche	Left
Europe Écologie les Verts	Left
Extrême Droite	Far-right
Extrême Gauche	Far-left
France Insoumise	Far-left
Front National	Far-right
Les Républicains	Right
Modem	Liberal
Radicaux de Gauche	Left
Régionalistes	Other
République en Marche	Liberal
Parti Socialiste	Left
Union des Démocrats et Indépendants	Liberal

2022 Parliamentary Elections

Political label	Political orientation
Divers	Other
Divers Centre	Other
Divers Droite	Right
Divers Gauche	Left
Divers Extrême-droite	Far-right
Divers Extrême-gauche	Far-left
Droite Souveraniste	Far-right
Écologiste	Left
Ensemble	Liberal
Extrême-droite	Far-right
Les Républicains	Right
Nupes	Left
Radicaux de Gauche	Left
Réconquête	Far-right
Régionalistes	Other
Rassemblement National	Far-right
Union des Démocrats et Indépendants	Liberal

2024 Parliamentary Elections	
Political label	**Political orientation**
Divers	Other
Divers Centre	Other
Divers Droite	Right
Divers Gauche	Left
Divers Extrême-droite	Far-right
Divers Extrême-gauche	Far-left
Droite Souveraniste	Far-right
Écologiste	Left
Ensemble	Liberal
Extrême-droite	Far-right
Hôrizons	Right
Les Républicains	Right
Nupes	Left
Radicaux de Gauche	Left
Réconquête	Far-right
Régionalistes	Other
Rassemblement National	Far-right
Union des Démocrats et Indépendants	Liberal
Union de Gauche	Left
Union Extrême-droite	Far-right

B Data

B.1 Political platforms

I transform the pdf versions of the manifestos into text using optical character recognition: Tesseract.

Table B1: Sampling frame

	First round			Second round		
Year	Races	Total candidates	Platforms collected	Races	Total candidates	Platforms collected
1981	491	2644	2452	333	658	649
1988	577	2820	2585	452	893	830
1993	577	5180	4071	490	977	956
1997	577	6205	2851	553	1170	1068
2017	577	6714	4666	482	964	741
2022	577	4990	3932	454	909	828
2024	577	4009	2749	161	1094	850

Notes: The table indicates the number of races, total candidates, and the manifestos collected for each legislative election included in the data set.

Figure B.1: Florence Blatrix-Contat political platform during 1st round 2017 legislative elections - 1st page

Source: RegardsCitoyens (https://github.com/regardscitoyens)

Translation: Florence BLATRIX-CONTAT Your deputy Michel FONTAINE deputy On the left to make France succeed

Madam, Sir, On May 7, the French people chose a new President of the Republic. The legislative elections of June 11 and 18 will decide the future of France; they will be an opportunity to choose the Republic we want. I am standing for election with the determination to make France a success, to make this five-year term a success. The majority resulting from this vote will have to act for social justice, solidarity, equal opportunities and ecological transition. For this, the left must be strong in the National Assembly. Tomorrow, in the Assembly, I will present a constructive and demanding left. I will ensure the defense of social achievements, respect for social dialogue and will fight for a growth model compatible with the preservation of the planet. With my substitute, Michel FONTAINE, we live, work and invest in this constituency. We will put our experience at your service to develop in this territory: employment, public services, health, solidarity, but also the cultural and sporting activities essential to social ties. As a rural elected official, I know how much public services must be preserved in each municipality. On the ground as in the Assembly, I will put all my energy to accompany and support the projects of our territory and to help those who encounter difficulties. Sunday, June 11, I am counting on your support. You can count on my determination and dedication.

51 years old, married, 3 children I live in my native village in Drom, in the heart of Revermont. My farming grandparents and my parents passed on to me their attachment to these lands; above all, they taught me that school was a means of emancipation and social advancement. After studying accounting, I became a teacher. Associate of economics-management, I teach economics and law. Elected since 1995 in my municipality, then in 2015 in the Auvergne-Rhône-Alpes Region, I am convinced that public action and the will make it possible to to advance. In our communities, I am committed to working with everyone, beyond partisan divisions; it is thanks to the gathering of elected officials that projects can succeed.

Figure B.2: Florence Blatrix-Contat political platform during 1st round 2017 legislative elections - 2nd page

(continuation)

Translation: I will be vigilant: On tax justice: I will oppose the reduction in Wealth Tax and the increase in the general social contribution which will lead to a reduction in pensions for more than 8 million retirees. On social dialogue and the rights of employees: social dialogue must be strengthened and the Labor Code must remain a strong protective base for employees. On public services: "They are the heritage of those who have none"; I will oppose a further drop in the number of civil servants and local authority grants.

Locally, I will defend the projects of our regions: fibre optics, investment in universities, sports and cultural activities, combating medical deserts and maintaining services in rural areas.

Michel FONTAINE First Deputy of the City of Bourg-en-Bresse President of the Agglomeration from 2008 to the end of 2016 and after Vice-President for Economic Development. I have lived in Bourg-en-Bresse for more than 40 years, I shared my professional life between the Carriat high school and my company Fontaine Picard. I have two children and live together. I assumed associative responsibilities before becoming a local elected official. By committing myself alongside Florence Blatrix-Contat, I am choosing success with great loyalty to my convictions.

≪This five-year term will be successful if France does not forget anyone on the way. For that, we need a strong left. I know that Florence and Michel carry these values. I call on you to support them on June 11 and 18.≫ Jean-François DEBAT Candidate of the left, the democrats and the ecologists, I am counting on you from the 1st round.

www.florenceblatrix2017.com - florence.blatrix.contat2017@gmail.com florenceblatrix @Florence-Blatrix Vu le candidat - Agence TOUT&POSSIBLE - Imprimerie du Centre - Bourg-en-Bresse

B.2 Text pre-processing

For all types of text data, I perform the following pre-processing procedures.

I remove a list of words containing party names, party acronyms, parliamentary titles, and terms describing blocs of parties. I also eliminate first and last names.

I pre-process the content of the corpus following standard practices in natural language processing: remove punctuation and numbers, convert all letters to lower-case, lemmatize each word, and restrict the vocabulary to words used by at least 1% and 50% of the documents using spacy French version 3.5.0. In the case of legislative debates, I restrict the vocabulary to 0.5% and 50% of the documents. Last, I convert words from Latin-1 to UTF-8 for three reasons. Given that the OCR sometimes does not detect accents, it permits to reduce error; second, written questions for the ninth legislature are in UTF-8; third, it permits to save memory.

Some politicians opt for using the party platform instead of a personalized platform; this is common in small parties (e.g., Rassemblement National and Green parties) but an infrequent practice among well-established parties (e.g., socialist and republican parties). I do not include manifestos similar to the party platform in the training datasets. Keeping duplicate measures introduces the problem of multicollinearity, and it will make these manifestos count more. I include them in the regression model.

C   Methodology

C.1   Topic classification - Seeded LDA

Latent Dirichlet Allocation (LDA) (Blei et al. (2003)) has been the most widely used topic model (e.g. Hansen, McMahon, and Prat (2018), Weigel (2020), Djourelova (2023)). LDA is an unsupervised method that assumes that documents are composed of words that help determine the topics and maps documents to a list of topics by assigning each word in the document to different topics. The assignment is in terms of conditional probability estimates. Under LDA, a document, d, is generated under the following hierarchical process:

- For each topic k draw a multinomial over words $\phi \sim Dirichlet(\beta)$.

- For each document d:

  - Draw a multinomial over topics $\theta \sim Dirichlet(\alpha)$.

  - For each word w_{N_d}:

* Draw a topic $Z_{N_d} \sim Mult(\theta_D)$,
* Draw a word: $w \sim Mult(\phi_{Z_{d,w}})$.

When the number of documents is not large, the method is not efficient and topics tend to be difficult to interpret. A potential solution is to transform the method into a semi-supervised. Seeded LDA (Lu et al. (2011)) permits to define topics *a priori* through seeded words, before fitting the model.[19] Lu et al. (2011) specify a combined conjugate prior for each seed word, w, in $\phi \sim Dirichlet(\beta + C_w)$, where C_w is a pseudo-count added to the topic to which w belongs. In case there is no prior knowledge for a word w, $C_w = 0$. With a sample obtained via Gibbs sampling, the topic-word distribution phi_k is approximated, for each topic k and the document-topic distribution, θ_d, for each document d.

An alternative to the seeded LDA could be the Correlation Explanation (CorEx) model of Gallagher, Reing, Kale, and Ver Steeg (2017), and implemented by Djourelova, Durante, and Martin (2021). However, this method forces to choose an anchor strength. The anchor strength controls how much weight CorEx puts toward maximizing the mutual information between the seeded words and their respective topics. The authors encourage users to experiment with the anchor strength and determine the values that best suit their needs. Seeded LDA does not need an anchor strength.

Seeded words were gathered from a simple LDA fitted on the same corpus. For the case of manifestos, I start to print the list of selected words in the document. Then, I classify the most obvious words into their respective topics and use them as "seeded words".

Seeded Lda was run in RStudio using the "seededlda" package version 0.9.1 (Watanabe, Xuan-Hieu, and Watanabe (2023)).

C.1.1 List of seeded words per topic - Political platforms

Economy, Employment & Social Security: disability pension (*aah*), money (*argent*), austerity (*austérité*), budget (*budget*), budgetary (*budgétaire*), unemployment (*chômage*), unemployed (*chômeur*), business (*commerce*), trade (*commercer*), competition (*competition*), competitiveness (*competivité*), competition (*concurrence*), general social contribution (*csg*), deficit (*déficit*), tax exemption (*défiscalisation*), reduce taxation (*défiscaliser*), economy (*économie*), economic (*économique*), hire (*employer*), employer (*employeur*), entrepreneur (*entrepreneur*), entrepreneurship (*entrepreneuriat*),

---

[19]Some previous applications: Curini and Vignoli (2021), Fraccaroli and Pizzigolotto (2021) and Ash, Krümmel, and Slapin (2023). For a detailed explanation of the differences between the LDA and the seeded LDA, see Watanabe and Baturo (2023).

entreprise (*firm*), exportation (*exportation*), finance (*finance*), financing (*financement*), fund (*financer*), financial (*financier*), fiscal (*fiscal*), tax (*fiscalité*), gatt (*gatt*), tax (*impôt*), industry (*industrie*), industrial (*industriel*), inflation (*inflation*), wealth tax (*isf*), khomri (*khomri*), monetary (*monetaire*), worker (*ouvrier*), boss (*patron*), employer (*patronal*), bosses (*patronat*), poor (*pauvre*), poverty (*pauvreté*), pension (*pension*), sme (*pme*), small medium industry (*pmi*), precarious (*précaire*), precarity (*précarité*), privatisation (*privatisation*), privatise (*privatiser*), price (*prix*), companies register (*rcs*), recession (*récession*), reform (*réforme*), retirement (*rétraite*), income (*revenu*), income of active solidarity (*rsa*), social security scheme (*rsi*), wage (*salaire*), salary (*salariale*), employee (*salarié*), minimum wage (*smic*), rate (*taux*), tax (*tax*), taxation (*taxation*), tax (*taxer*), worker (*travailleur*), *uberisation*.

Environment: agrarian (*agricole*), farmer (*agriculteur*), agriculture (*agriculture*), agro, agribusiness (*agroalimentaire*), animal (*animal*), bio (*bio*), biodiversity (*biodiversité*), biological (*biologique*), carbon (*carbone*), fuel (*carburant*), carbide (*carbure*), climat (*climate*), climatic (*climatique*), water (*eau*), ecology (*écologie*), ecological (*écologique*), environment (*écologiste*), energetic (*énergétique*), energy (*énergie*), environment (*environnement*), green (*environnemental*), wind (*éolien*), species (*éspece*), forest (*fôret*), nuclear (*nucléaire*), fishing (*pêche*), fisher (*pêcheur*), programme for the endorsement of forest certification (*pefc*), pesticide (*pesticide*), petrol (*petrole*), planet (*planète*), polluting (*polluant*), pollute (*polluer*), pollution (*pollution*), recycling (*recyclage*), recycle (*recycler*), vegetarian (*végétarien*), winegrower (*viticulteur*), viticulture (*viticulture*).

Health & Education: academic (*académique*), class (*classe*), college (*collège*), doctor (*docteur*), school (*école*), educator (*éducateur*), educational (*éducatif*), education (*éducation*), educate (*éduquer*), establishment of accommodation for dependent old persons (*ehpad*), student (*élève*), endocrine (*endocrinien*), childhood (*enfance*), child (*enfant*), confinement (*enfermement*), teacher (*enseignant*), teaching (*enseignement*), teach (*enseigner*), study (*étude*), student (*étudiant*), study (*étudier*), training (*formation*), hospital (*hôpital*), hospitable (*hospitalier*), college (*lycée*), sick (*malade*), disease (*maladie*), maternity (*maternité*), doctor (*médecin*), medecine (*médecine*), medical (*médicale*), medication (*médicament*), patient (*patient*), teacher (*professeur*), blood (*sang*), sanitary (*sanitaire*), health (*santé*), science (*science*), scientific (*scientifique*), academic (*scolaire*), hiv (*sida*), care (*soin*), universitary (*universitaire*), university (*université*).

Security, Justice & Foreign Policy: africa (*afrique*, germany (*allemagne*), american (*americain*), weapon (*arme*), armed (*armée*), weapons (*armement*), asylum (*asile*), brussels (*bruxelle*), clandestin (*clandestine*), crime (*crime*), criminal (*criminel*), cybercrime (*cybercriminalité*), delinquency (*déliquant*), delinquent (*délinquant*), offence (*délit*), drug (*drogue*), foreigner (*étranger*), europe (*europe*), european (*européen*), bor-

der (*frontière*), policeman (*gendarme*), war (*guerre*), immigration (*immigration*), immigrant (*immigré*), insecurity (*insecurité*), maastricht (*maastricht*), magistrat (*magistrate*), world (*monde*), worldwide (*mondial*), nationality (*nationalité*), otan, sentence (*peine*), penal (*pénal*), police (*police*), police (*policier*), prison (*prison*), security (*sécurité*), terrorism (*terrorisme*), terrorist (*terroriste*), treaty (*traité*), court (*tribunal*), ukraine (*ukraine*), victim (*victime*), violence (*violence*).

Local: canton (*canton*), cantonal (*cantonal*), municipal (*communal*), communitarian (*communautaire*), community (*communauté*), municipality (*commune*), decentralisation (*décentralisation*), decentralise (*décentraliser*), departmental (*départementale*), desert (*désert*), desertification (*désertification*), inhabitant (*habitant*), intercommunal (*intercommunal*), municipal (*municipal*), municipality (*municipalité*), region (*région*), regional (*régional*), rural (*rural*), rurality (*ruralité*), land (*terrain*), territory (*territoire*), territorial (*territorial*), city (*ville*), area (*zone*).

Politics: antisocialist (*antisocialiste*), assembly (*assemblée*), campaign (*campagne*), candidature (*candidature*), centrist (*centriste*), coalition (*coalition*), cohabitation (*cohabitation*), constitution (*constitution*), democrat (*démocrate*), democracy (*démocratie*), democratisation (*démocratisation*), sunday (*dimanche*), dissolution (*dissolution*), dissolve (*dissoudre*), right (*droite*), voter (*électeur*), elective (*électif*), electoral (*électoral*), elected (*élu*), inhibit (*empêcher*), left (*gauche*), gaulliste, holland, majority (*majoritaire*), majority (*majority*), presidency (*présidence*), president (*président*), presidential (*présidentiel*), reelection (*reélection*), reelect (*reélire*), republican (*républicain*), republic (*république*), senate (*sénat*), senator (*sénateur*), socialism (*socialisme*).

C.1.2   Top 10 words per topic

Economy, employment & social security: firm (*entreprise*), retirement (*retraite*), economic (*économique*), unemployment (*chômage*), economy (*économie*), tax (*impôt*), wage (*salaire*), reform (*réforme*), worker (*travailleur*), employee (*salarié*).

Environment: environment (*environnement*), energy (*énergie*), ecology (*écologie*), environment (*écologiste*), agriculture (*agriculture*), ecological (*écologique*), farmer (*agriculteur*), agrarian (*agricole*), nuclear (*nucleaire*), water (*eau*).

Health & education: child (*enfant*), health (*santé*), school (*école*), education (*éducation*), training (*formation*), class (*classe*), hospital (*hôpital*), academic (*scolaire*), teaching (*enseignement*), medical (*médicale*).

Security, justice & foreign policy: europe (*europe*), security (*sécurité*), world (*monde*), european (*européen*), immigration (*immigration*), fight against (*lutter contre*), fight (*lutter*), foreigner (*étranger*), police (*police*), insecurity (*insecurité*).

Local: territory (*territoire*), city (*ville*), region (*région*), rural (*rural*), municipal (*municipal*), regional (*régional*), land (*terrain*), inhabitant (*habitant*), defend (*défendre*),

municipality (*commune*).

Politics: majority (*majorité*), left (*gauche*), president (*président*), right (*droite*), assembly (*assemblée*), republic (*république*), elected (*élu*), voter (*électeur*), presidential (*présidentiel*), sunday (*dimanche*).

Other 1: councillor (*counseiller*), general (*général*), general councillor (*counseiller général*, trust (*confiance*), child (*enfant*), future (*avenir*), council (*conseil*), married (*marié*), department (*département*), freedom (*liberté*).

Other 2: citizen (*citoyen*), make (*falloir*), no (*non*), society (*societé*), other (*autre*), man (*homme*), right (*droit*), enter (*entrer*), today (*aujour*), live (*vivre*).

C.2   Empirical strategy

Figure C.3: Manipulation testing: Most voted female - 2nd most voted male
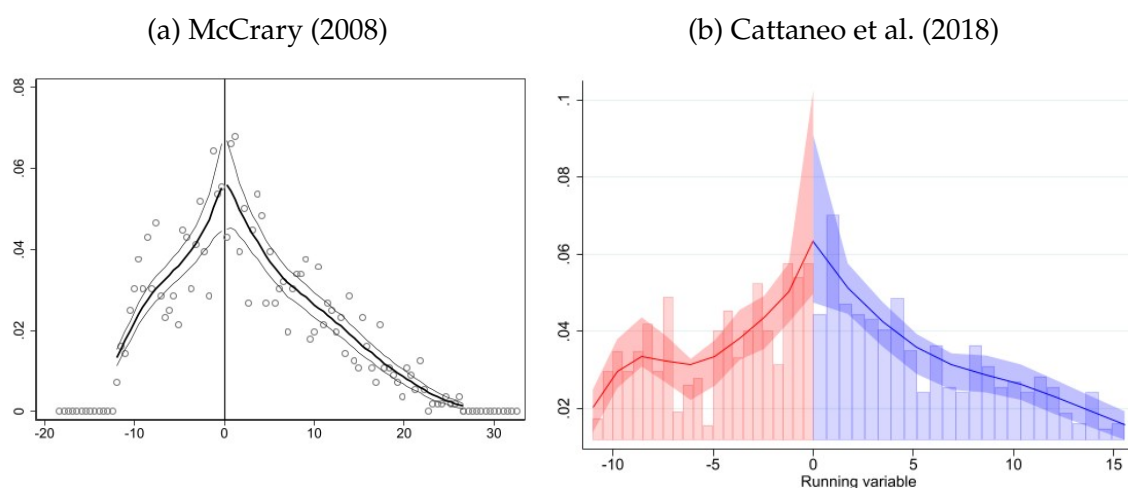
(a) McCrary (2008)   (b) Cattaneo et al. (2018)

Notes. Figures (a) and (b) represent the density test for races where only one candidate (or none) obtained the 12.5%; the margin is the difference between the most-voted woman and the second most-voted man. Figure (a) represents the McCrary density test; discontinuity estimate b: 0.006 (s.e. 0.153). Figure (b) represents the Cattaneo et al. (2018) manipulation test; p-value 0.698 (not reject the null hypothesis of no manipulation).

I conduct placebo tests to examine whether there is discontinuity at the threshold for any of the variables used to predict treatment. I first provide information about the construction of each variable. If the information is missing, it is because the name of the dependent variable is self-explanatory.

Platform available: dummy equal to 1 if the manifesto for the 1st and 2nd round is available, 0 if not.

Votes: number of votes obtained divided by the number of enrolled voters.

Number candidates: number of candidates running in the electoral district.

Number female: number of female candidates running in the electoral district.

Victory margin: margin between the most voted and the second most voted candidate.

Number far-left, left, right, far-right: Number of candidates of the respective ideology.

Sum left/right: sum of the vote share in all left/right candidates.

Number words: total number of words in the manifesto.

Far-left, left, liberal, right, far-right: a dummy equal to 1 if the politician is classified as belonging to that ideology, 0 if not.

Table C2: Balancing tests: 1st round electoral district characteristics

	(1) Votes	(2) Turnout	(3) Number candidates	(4) Number female	(5) Enrolled voters	(6) Victory margin
Woman	0.415	0.309	0.060	-0.029	-5,252	-0.542
	(0.618)	(2.904)	(0.921)	(0.570)	(3,864)	(1.653)
Observations	836	836	836	836	836	836
Eff. number of obs	258	240	186	239	333	259
Robust p-value	0.524	0.916	0.634	0.850	0.150	0.827
Polyn. order	1	1	1	1	1	1
Bandwidth	3.558	3.270	2.446	3.234	4.962	3.575
Outcome mean	9.240	49.424	12.222	4.387	79056	7.416

	(7) Number far-left	(8) Number left	(9) Number right	(10) Number far-right	(11) Sum left	(12) Sum right
Woman	0.740	-0.340	-0.042	-0.097	0.540	-3.022
	(0.332)	(0.364)	(0.287)	(0.199)	(1.765)	(3.572)
Observations	836	836	836	836	836	836
Eff. number of obs	146	284	217	324	260	186
Robust p-value	0.012	0.316	0.863	0.465	0.813	0.257
Polyn. order	1	1	1	1	1	1
Bandwidth	1.885	3.991	2.977	4.806	3.593	2.444
Outcome mean	2.216	3.138	1.849	2.020	16.351	19.272

Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the candidate is a woman. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table C3: Balancing tests - differences in the ideology of female and male candidates

	(1)	(2)	(3)	(4)	(5)
	Far-left	Left	Liberal	Right	Far-right
Woman	0.040	0.118	-0.022	-0.166*	0.064
	(0.071)	(0.108)	(0.102)	(0.085)	(0.100)
Observations	836	836	836	836	836
Eff. number of obs	309	336	318	240	271
Robust p-value	0.690	0.462	0.767	0.187	0.649
Polyn. order	1	1	1	1	1
Bandwidth	4.373	5.013	4.500	3.261	3.833
Outcome mean	0.058	0.259	0.111	0.291	0.230

In columns (1)-(5), the outcome variable is a dummy equal to 1 if the candidate belongs to the referred ideology, 0 otherwise. Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the candidate is a woman. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table C4: Balancing tests - differences in the characteristics of female and male candidates

	(1)	(2)	(3)	(4)
	Available	Number words	Incumbent	Run in the past
Woman	0.008	2.770	-0.009	-0.155
	(0.113)	(64.020)	(0.081)	(0.111)
Observations	1130	836	836	836
Eff. number of obs	351	353	254	345
Robust p-value	0.823	0.975	0.888	0.230
Polyn. order	1	1	1	1
Bandwidth	3.194	5.429	3.489	5.234
Outcome mean	0.659	714.497	0.063	0.365

In column 1, the outcome is a dummy equal to 1 if the manifesto is available in both election rounds. In column 2, the outcome is the number of words of the manifesto in 1st round. In column 3, the dependent variable is a dummy variable, whether the candidate is an incumbent. In column 4, the outcome is a dummy variable equal to 1 if the candidate ran in the past, 0 otherwise. Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the candidate is a woman. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table C5: Balancing tests - differences in the ideology of female and male opponents

	(1) Far-left	(2) Left	(3) Liberal	(4) Right	(5) Far-right
Woman	0.022	-0.102	0.202	0.049	-0.112
	(0.029)	(0.113)	(0.162)	(0.151)	(0.091)
Observations	836	836	836	836	836
Eff. number of obs	282	271	191	189	254
Robust p-value	0.461	0.354	0.133	0.962	0.218
Polyn. order	1	1	1	1	1
Bandwidth	3.955	3.842	2.553	2.534	3.496
Outcome mean	0.014	0.205	0.433	0.159	0.126

In columns (1)-(5), the outcome variable is a dummy equal to 1 if the opponent belongs to the referred ideology, 0 otherwise. Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the candidate is a woman. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table C6: General balance test

	(1) Predicted treatment
Woman	0.057
	(0.052)
Observations	836
Eff. number of obs	344
Robust p-value	0.337
Polyn. order	1
Bandwidth	5.165
Outcome mean	0.393

The outcome is the candidate's predicted treatment status based on observable characteristics listed in the text. The outcome is computed as follows: first, the treatment variable T is regressed on all 25 baseline variables presented in Tables C2-C5, and then the treatment status of each candidate is predicted using the regression coefficients. The sample is restricted to candidates included in the RDD sample as described in the text. The independent variable is an indicator equal to 1 if the candidate is a female. I use a nonparametric estimation procedure and MSERD data-driven bandwidths. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10 percent, respectively. The mean gives the average outcome value for the male candidates at the threshold.

Table C7: Balancing tests - differences in policy topics of the most voted candidate in the 1st round

	(1) Economy & employment	(2) Environment	(3) Health & education	(4) Security & foreign policy	(5) Local politics	(6) National politics
Woman	-0.022	-0.007	0.066	0.026	-0.059	0.033
	(0.052)	(0.030)	(0.044)	(0.047)	(0.060)	(0.050)
Observations	645	645	645	645	645	645
Eff. number of obs	227	145	204	224	168	175
Robust p-value	0.868	0.515	0.114	0.822	0.565	0.445
Polyn. order	1	1	1	1	1	1
Bandwidth	4.178	2.492	3.800	4.137	3.030	3.139
Outcome mean	0.246	0.046	0.071	0.068	0.256	0.130

Column (1)-(5) is the coverage of the referred policy topic by the most voted candidate in an electoral district, the variable is between 0 and 1. Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the candidate is a woman, 0 if it is a man. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure. The dependent variables take values between 0 and 1, and their computation is explained in Section 5.1. The mean gives the average outcome value for male candidates at the threshold.

D Results: Gender Differences in Political Platforms

Figure D.4: Differences between female and male 2nd round political candidates on topics

(a) Economy & employment

(b) Environment
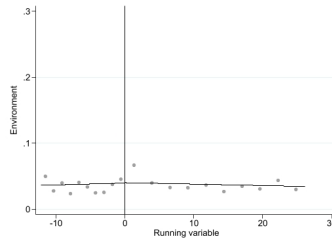
(c) Health & education
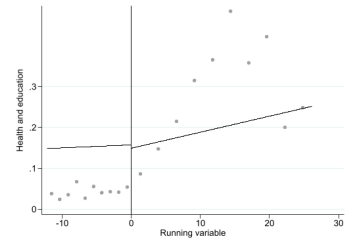
(d) Security & foreign policy
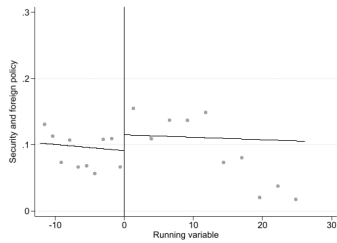
(e) Local politics
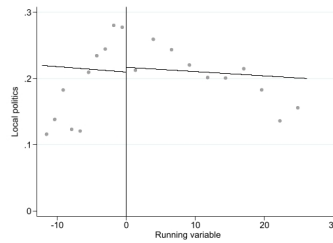
(f) National politics
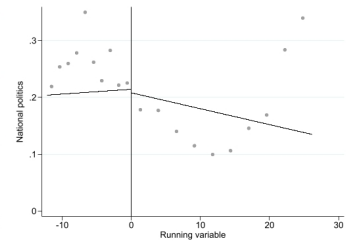
Dots represent the local averages of the topic on a political platform. Averages are calculated within quantile-spaced bins of the running variable. The running variable is the difference between the most-voted woman in a race and the second most-voted man. Positive (negative) values denote that the female (male) candidate passed the second round. All characteristics presented in Tables C1-C5 are included as controls.

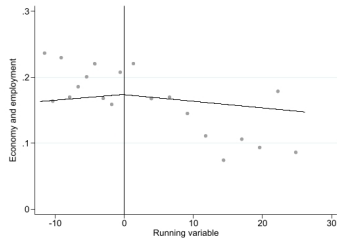Table D8: Differences between female and male 2nd round political candidates (half of MSERD bandwidth)

	(1) Economy & employment	(2) Environment	(3) Health & education	(4) Security & foreign policy	(5) Local politics	(6) National politics
Woman	-0.003	-0.003	-0.034	0.126***	-0.081**	-0.031
	(0.029)	(0.008)	(0.024)	(0.031)	(0.039)	(0.047)
Observations	836	836	836	836	836	836
Eff. number of obs	185	130	170	127	130	148
Robust p-value	0.915	0.660	0.157	0.000	0.041	0.567
Polyn. order	1	1	1	1	1	1
Bandwidth	2.370	1.601	2.207	1.464	1.575	1.916
Outcome mean	0.193	0.045	0.049	0.082	0.277	0.227

Standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest is a woman running in the second round as an alternative to a man. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under half of the MSERD procedure. The dependent variables take values between 0 and 1, and their computation is explained in Section 5.1. The mean gives the average outcome value for male candidates at the threshold.

Table D9: Differences between female and male 2nd round political candidates (double of MSERD bandwidth)

	(1) Economy & employment	(2) Environment	(3) Health & education	(4) Security & foreign policy	(5) Local politics	(6) National politics
Woman	-0.005	0.001	-0.016	0.055***	-0.012	0.004
	(0.015)	(0.005)	(0.012)	(0.019)	(0.021)	(0.024)
Observations	836	836	836	836	836	836
Eff. number of obs	573	406	536	378	399	472
Robust p-value	0.959	0.903	0.133	0.034	0.730	0.911
Polyn. order	1	1	1	1	1	1
Bandwidth	9.480	6.404	8.830	5.858	6.301	7.665
Outcome mean	0.194	0.035	0.048	0.086	0.253	0.251

Standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest is a woman running in the second round as an alternative to a man. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under double of the MSERD procedure. The dependent variables take values between 0 and 1, and their computation is explained in Section 5.1. The mean gives the average outcome value for male candidates at the threshold.

E Potential Mechanisms

E.1 Do Women Strategically Adapt to the Gender of Their Opponent?

Figure E.5: Manipulation testing: Most voted female - 2nd most voted male
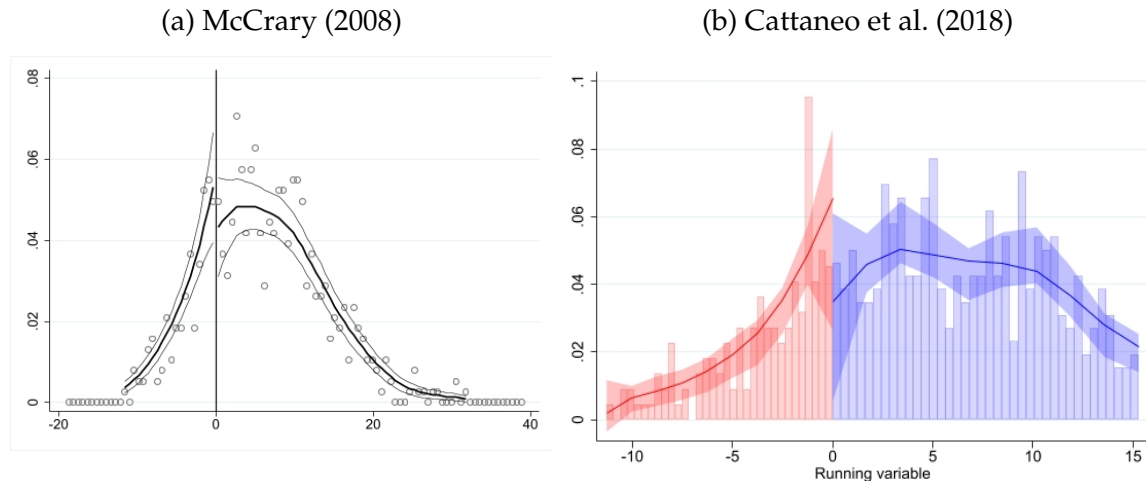
(a) McCrary (2008)   (b) Cattaneo et al. (2018)

Notes. Figures (a) and (b) represent the density test for races where only one candidate (or none) obtained the 12.5%; the margin is the difference between the most-voted man and the second most-voted woman. Figure (a) represents the McCrary density test; discontinuity estimate b: -0.283 (s.e. 0.214). Figure (b) represents the Cattaneo et al. (2018) manipulation test; p-value 0.275 (not reject the null hypothesis of no manipulation).

Table E10: Balancing tests: 1st round electoral district characteristics

	(1) Votes	(2) Turnout	(3) Number candidates	(4) Number female	(5) Enrolled voters	(6) Victory margin
Woman competing against a man	1.845 (1.551)	1.580 (2.151)	-0.330 (1.029)	-0.121 (0.541)	4,398 (5,101)	1.755 (1.862)
Observations	498	498	498	498	498	498
Eff. number of obs	167	153	140	181	189	163
Robust p-value	0.150	0.282	0.767	0.963	0.375	0.258
Polyn. order	1	1	1	1	1	1
Bandwidth	3.381	3.151	2.887	3.699	3.888	3.311
Outcome mean	3.678	48.094	13.591	6.775	82284	7.324

	(7)	(8)	(9)	(10)	(11)	(12)
	Number far-left	Number left	Number right	Number far-right	Sum left	Sum right
Woman competing	0.468	-0.758	0.182	-0.634	-0.902	-0.030
against a man	(0.384)	(0.416)	(0.350)	(0.374)	(2.049)	(2.157)
Observations	498	498	498	498	498	498
Eff. number of obs	165	117	130	111	153	171
Robust p-value	0.230	0.072	0.453	0.072	0.782	0.931
Polyn. order	1	1	1	1	1	1
Bandwidth	3.346	2.440	2.685	2.250	3.139	3.472
Outcome mean	2.278	3.535	2.114	2.240	15.392	17.296

Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the most voted woman in the first round competes against a man in the second round and 0 if she competes against a woman. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table E11: Balancing tests - differences in ideology of women competing against a man vs women competing against a woman

	(1)	(2)	(3)	(4)
	Left	Liberal	Right	Far-right
Woman competing	0.014	0.016	-0.037	0.034
against a man	(0.102)	(0.178)	(0.103)	(0.137)
Observations	498	498	498	498
Eff. number of obs	151	178	150	181
Robust p-value	0.803	0.950	0.880	0.815
Polyn. order	1	1	1	1
Bandwidth	3.061	3.625	3.027	3.704
Outcome mean	0.257	0.067	0.149	0.183

In columns (1)-(5), the outcome variable is a dummy equal to 1 if the candidate belongs to the referred ideology, 0 otherwise. Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the most voted woman in the first round competes against a man in the second round and 0 if she competes against a woman. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table E12: Balancing tests - differences in women competing against a man vs women competing against a woman

	(1) Available	(2) Number words	(3) Incumbent	(4) Run in the past
Woman competing	-0.102	98.560	-0.134	-0.216
against a man	(0.107)	(121.505)	(0.170)	(0.188)
Observations	644	498	498	498
Eff. number of obs	228	164	130	121
Robust p-value	0.404	0.412	0.363	0.221
Polyn. order	1	1	1	1
Bandwidth	4.031	3.331	2.634	2.511
Outcome mean	0.123	705.287	0.019	0.175

Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value and ***, ** and * indicate significance at 1, 5 and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the most voted woman in the first round competes against a man in the second round, 0 if she competes against a woman. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table E13: Balancing tests - differences in the ideology of the opponents of women competing against a man vs women competing against a woman

	(1) Far-left	(2) Left	(3) Liberal	(4) Right	(5) Far-right
Woman competing	-0.079	-0.096	0.044	0.206	-0.072
against a man	(0.083)	(0.157)	(0.142)	(0.144)	(0.111)
Observations	498	498	498	498	498
Eff. number of obs	169.	166	192	174	199
Robust p-value	0.364	0.575	0.659	0.200	0.455
Polyn. order	1	1	1	1	1
Bandwidth	3.432	3.365	3.943	3.603	4.179
Outcome mean	0.143	0.202	0.197	0.263	0.180

In columns (1)-(5), the outcome variable is a dummy equal to 1 if the candidate belongs to the referred ideology, 0 otherwise. Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the most voted woman in the first round competes against a man in the second round and 0 if she competes against a woman. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table E14: Balancing tests - differences in the coverage of topics between women competing against a man and women competing against a woman

	(1) Economy & employment	(2) Environment	(3) Health & education	(4) Security & foreign policy	(5) Local politics	(6) National politics
Woman competing	-0.024	-0.012	0.222	0.000	-0.075	-0.048*
against a man	(0.064)	(0.014)	(0.149)	(0.054)	(0.066)	(0.026)
Observations	498	498	498	498	498	498
Eff. number of obs	159	171	137	220	150	152
Robust p-value	0.664	0.346	0.166	0.880	0.366	0.079
Polyn. order	1	1	1	1	1	1
Bandwidth	3.263	3.480	2.865	4.743	3.016	3.094
Outcome mean	0.234	0.090	0.077	0.075	0.232	0.124

Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the most voted woman in the first round competes against a man in the second round and 0 if she competes against a woman. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table E15: General balance test

	(1)
Woman competing	0.016
against a man	(0.050)
Observations	498
Eff. number of obs	214
Robust p-value	0.824
Polyn. order	1
Bandwidth	4.626
Outcome mean	0.932

The outcome is the candidate's predicted treatment status based on observable characteristics. The outcome is computed as follows: first, regress the treatment variable T on all 30 baseline variables presented in Tables E10-E14 and then predict the treatment status of each candidate using the regression coefficients. The sample is restricted to candidates included in the RDD sample as described in the text. The independent variable is an indicator equal to 1 if the most voted woman in the 1st round competes against a man in the second round and 0 if she competes against a woman. I use a nonparametric estimation procedure and MSERD data-driven bandwidths. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10 percent, respectively. The dependent variables take values between 0 and 1, and their computation is explained in Section 5.1. The mean gives the average outcome value for the women competing against a woman at the threshold.

Table E16: Impact of a marginal presence of a man on female candidates in the 2nd round (double of MSERD bandwidth)

	(1) Economy & employment	(2) Environment	(3) Health & education	(4) Security & foreign policy	(5) Local politics	(6) National politics
Woman competing	-0.016	-0.002	0.009	0.048**	-0.004	-0.022
against a man	(0.012)	(0.005)	(0.027)	(0.020)	(0.026)	(0.022)
Observations	498	498	498	498	498	498
Eff. number of obs	259	266	298	256	259	295
Robust p-value	0.369	0.877	0.393	0.188	0.775	0.345
Polyn. order	1	1	1	1	1	1
Bandwidth	5.930	6.226	7.384	5.782	6.016	7.193
Outcome mean	0.158	0.043	0.193	0.132	0.252	0.145

Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value and ***, ** and * indicate significance at 1, 5 and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the most voted woman in the first round competes against a man in the second round, 0 if she competes against a woman. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under double of the MSERD procedure. The dependent variables take values between 0 and 1, and their computation is explained in Section 5.1. The mean gives the average outcome value for women competing against a woman at the threshold.
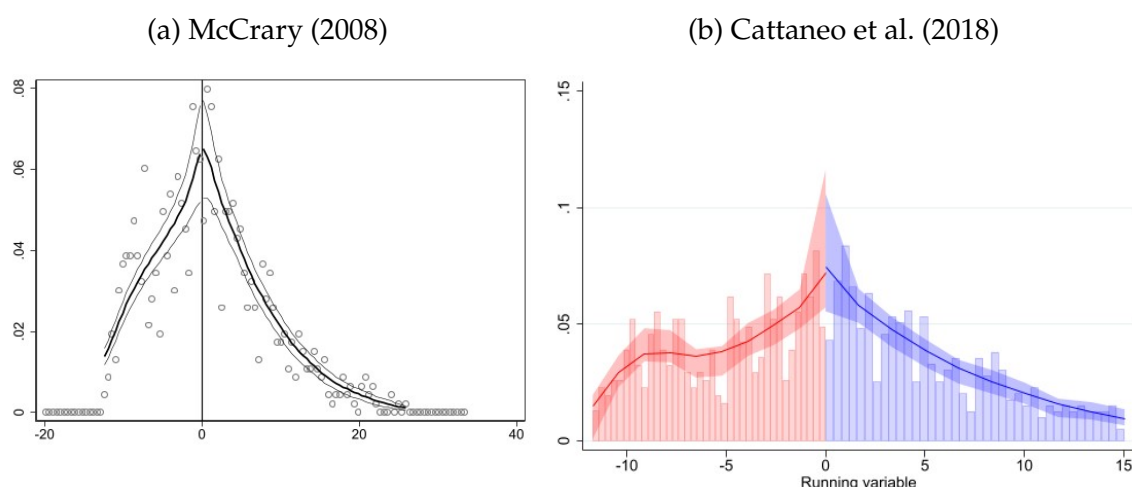
E.2   Do men strategically adapt to the gender of their opponent?

Figure E.6: Manipulation testing: Most voted female - 2nd most voted male

(a) McCrary (2008)   (b) Cattaneo et al. (2018)

Notes. Figures (a) and (b) represent the density test for races where only one candidate (or none) obtained the 12.5%; the margin is the difference between the most-voted woman and the second most-voted man. Figure (a) represents the McCrary density test; discontinuity estimate b: 0.013 (s.e. 0.143). Figure (b) represents the Cattaneo et al. (2018) manipulation test; p-value 0.770 (not reject the null hypothesis of no manipulation).

Table E17: Balancing tests: 1st round electoral district characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
	Votes	Turnout	Number candidates	Number female	Enrolled voters	Victory margin
Man competing	0.075	-0.435	-0.331	0.437	-1,610	-0.146
against a woman	(1.132)	(2.054)	(0.577)	(0.531)	(3,309)	(1.248)
Observations	803	802	803	803	802	803
Eff. number of obs	374	332	408	280	426	392
Robust p-value	0.992	0.803	0.718	0.327	0.599	0.895
Polyn. order	1	1	1	1	1	1
Bandwidth	4.256	3.817	4.867	3.056	5.057	4.482
Outcome mean	4.977	54.361	11.883	4.233	78672	6.742

	(7)	(8)	(9)	(10)	(11)	(12)
	Number	Number	Number	Number	Sum	Sum
	far-left	left	right	far-right	left	right
Man competing	0.666**	-0.636*	-0.091	-0.040	-2.412	-1.540
against a woman	(0.304)	(0.357)	(0.246)	(0.191)	(1.664)	(2.037)
Observations	803	803	803	803	803	803
Eff. number of obs	189	276	293	300	297	361
Robust p-value	0.019	0.093	0.888	0.713	0.168	0.332
Polyn. order	1	1	1	1	1	1
Bandwidth	1.977	3.005	3.170	3.339	3.262	4.104
Outcome mean	2.163	3.230	1.908	1.905	18.540	22.880

Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the most voted man in the first round competes against a woman in the second round, 0 if he competes against a man. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table E18: Balancing tests - differences in the ideology of men competing against a woman vs men competing against a man

	(1)	(2)	(3)	(4)	(5)
	Far-left	Left	Liberal	Right	Far-right
Man competing	0.016	-0.300***	0.199	0.005	0.005
against a woman	(0.051)	(0.107)	(0.128)	(0.099)	(0.082)
Observations	803	803	803	803	803
Eff. number of obs	335	249	255	280	328
Robust p-value	0.637	0.007	0.103	0.923	0.971
Polyn. order	1	1	1	1	1
Bandwidth	3.867	2.676	2.801	3.058	3.741
Outcome mean	0.182	0.261	0.046	0.160	0.169

In columns (1)-(5), the outcome variable is a dummy equal to 1 if the candidate belongs to the referred ideology, 0 otherwise. Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the most voted man in the first round competes against a woman in the second round and 0 if he competes against a man. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table E19: Balancing tests - differences in the characteristics of men competing against a woman vs men competing against a man

	(1) Available	(2) Number words	(3) Incumbent	(4) Run in the past
Man competing	-0.121	48.086	-0.089	0.038
against a woman	(0.088)	(75.088)	(0.120)	(0.114)
Observations	992	803	803	803
Eff. number of obs	339	344	277	330
Robust p-value	0.163	0.478	0.414	0.688
Polyn. order	1	1	1	1
Bandwidth	3.064	3.922	3.029	3.776
Outcome mean	0.122	718.413	0.030	0.220

Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the most voted man in the first round competes against a woman in the second round and 0 if he competes against a man. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table E20: Balancing tests - differences in the ideology of the opponents of women competing against a man vs women competing against a woman

	(1) Far-left	(2) Left	(3) Liberal	(4) Right	(5) Far-right
Man competing	0.048	0.052	0.031	-0.120	0.078
against a woman	(0.062)	(0.124)	(0.099)	(0.086)	(0.094)
Observations	803	803	803	803	803
Eff. number of obs	354	296	360	229	381
Robust p-value	0.544	0.853	0.846	0.374	0.350
Polyn. order	1	1	1	1	1
Bandwidth	4.031	3.228	4.093	2.412	4.363
Outcome mean	0.079	0.245	0.102	0.289	0.217

In columns (1)-(5), the outcome variable is a dummy equal to 1 if the candidate belongs to the referred ideology, 0 otherwise. Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the most voted man in the first round competes against a woman in the second round and 0 if he competes against a man. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table E21: Balancing tests - differences in the coverage of topics between men competing against a woman and men competing against a man

	(1) Economy & employment	(2) Environment	(3) Health & education	(4) Security & foreign policy	(5) Local politics	(6) National politics
Man competing	-0.028	-0.013	0.054	0.050	-0.047	0.011
against a woman	(0.040)	(0.013)	(0.091)	(0.040)	(0.041)	(0.027)
Observations	803	803	803	803	803	803
Eff. number of obs	347	293	257	304	363	322
Robust p-value	0.420	0.340	0.441	0.304	0.213	0.725
Polyn. order	1	1	1	1	1	1
Bandwidth	3.950	3.193	2.843	3.408	4.133	3.646
Outcome mean	0.244	0.082	0.061	0.060	0.194	0.134

Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the candidate. The independent variable is a dummy equal to 1 if the most voted man in the first round competes against a woman in the second round and 0 if he competes against a man. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table E22: General balance test

	(1)
Man competing	0.046
against a woman	(0.049)
Observations	802
Eff. number of obs	315
Robust p-value	0.408
Polyn. order	1
Bandwidth	3.473
Outcome mean	0.835

The outcome is the candidate's predicted treatment status based on observable characteristics. The outcome is computed as follows: first, regress the treatment variable T on all 30 baseline variables presented in Tables E17-E21 and then predict the treatment status of each candidate using the regression coefficients. The sample is restricted to candidates included in the RDD sample as described in the text. The independent variable is an indicator equal to 1 if the most voted man in the 1st round competes against a woman in the second round and 0 if he competes against a man. I use a nonparametric estimation procedure and MSERD data-driven bandwidths. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10 percent, respectively. The mean gives the average outcome value for men competing against men at the threshold.

Table E23: Impact of a marginal presence of a woman on male candidates in the 2nd round (without unbalanced covariates)

	(1)	(2)	(3)	(4)	(5)	(6)
	Economy & employment	Environment	Health & education	Security & foreign policy	Local politics	National politics
Man competing	0.019	0.003	-0.004	0.012	0.014	-0.040
against a woman	(0.015)	(0.004)	(0.017)	(0.017)	(0.022)	(0.030)
Observations	802	802	802	802	802	802
Eff. number of obs	252	293	375	399	375	271
Robust p-value	0.138	0.519	0.946	0.594	0.564	0.187
Polyn. order	1	1	1	1	1	1
Bandwidth	2.700	3.187	4.301	4.683	4.287	2.981
Outcome mean	0.152	0.032	0.116	0.064	0.214	0.255

Standard errors are in parenthesis. Statistical significance is computed based on the robust p-value and ***, ** and * indicate significance at 1, 5 and 10, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest is a woman running in the second round in alternative to a man. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure. The dependent variables take values between 0 and 1, and their computation is explained in Section 5.1. The mean gives the average value of the outcome for men competing against a woman at the threshold.

Table E24: Impact of a marginal presence of a woman on male candidates in the 2nd round (double of MSERD bandwidth)

	(1)	(2)	(3)	(4)	(5)	(6)
	Economy & employment	Environment	Health & education	Security & foreign policy	Local politics	National politics
Man competing	0.009	0.002	-0.009	0.012	0.014	-0.014
against a woman	(0.011)	(0.004)	(0.015)	(0.014)	(0.016)	(0.022)
Observations	802	802	802	802	802	802
Eff. number of obs	432	469	480	523	594	471
Robust p-value	0.085	0.518	0.722	0.711	0.529	0.279
Polyn. order	1	1	1	1	1	1
Bandwidth	5.240	5.927	6.209	7.024	8.275	6.000
Outcome mean	0.146	0.029	0.105	0.065	0.178	0.267

Standard errors are in parenthesis. Statistical significance is computed based on the robust p-value and ***, ** and * indicate significance at 1, 5 and 10, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest is a woman running in the second round in alternative to a man. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under double of the MSERD procedure. The dependent variables take values between 0 and 1, and their computation is explained in Section 5.1. The mean gives the average value of the outcome for men competing against a man at the threshold.
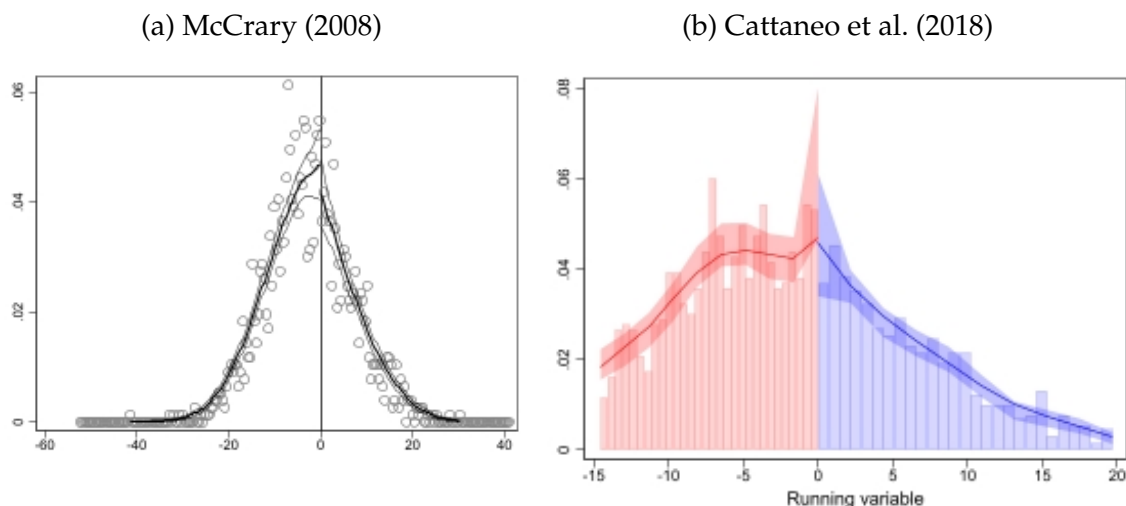
E.3 Are women behaving strategically during elections or also signaling their type as MPs?

Figure E.7: Manipulation testing: Share of votes on woman - Share of votes on man

(a) McCrary (2008)　　　　　　　　(b) Cattaneo et al. (2018)

Notes. This figure tests for a jump in the density of the running variable. The solid line represents the density of the running variable. Thin lines represent the confidence intervals. Figures (a) and (b) represent the density test for mixed-gender races where a woman wins against a man. Figure (a) represents the McCrary density test; discontinuity estimate b: -0.146 (s.e. 0.112). Figure (b) represents the Cattaneo et al. (2018) manipulation test; p-value 0.780 (not reject the null hypothesis of no manipulation).

Table E25: Balancing tests: 1st round electoral district characteristics

	(1) Votes	(2) Turnout	(3) Number candidates	(4) Number female	(5) Enrolled voters	(6) Victory margin
Woman	0.694	2.377	0.179	-0.077	-3,116	-0.084
	(0.812)	(1.446)	(0.494)	(0.298)	(2,376)	(0.477)
Observations	1783	1783	1783	1783	1783	1783
Eff. number of obs	802	757	971	900	864	686
Robust p-value	0.339	0.097	0.803	0.624	0.288	0.885
Polyn. order	1	1	1	1	1	1
Bandwidth	5.739	5.344	7.096	6.580	6.327	4.900
Outcome mean	18.630	53.885	12.087	4.797	78465	7.545

	(7)	(8)	(9)	(10)	(11)	(12)
	Number	Number	Number	Number	Sum	Sum
	far-left	left	right	far-right	left	right
Woman	-0.128	0.145	0.054	0.027	-0.058	0.866
	(0.145)	(0.204)	(0.145)	(0.120)	(1.262)	(1.207)
Observations	1783	1783	1783	1783	1783	1783
Eff. number of obs	921	934	931	875	966	1008
Robust p-value	0.321	0.532	0.714	0.730	0.905	0.435
Polyn. order	1	1	1	1	1	1
Bandwidth	6.718	6.854	6.826	6.404	7.043	7.337
Outcome mean	2.076	3.096	1.928	2.058	18.310	23.165

Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the MP politician. The independent variable is a dummy equal to 1 if it is a woman and 0 if it is a man. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table E26: Balancing tests - differences in the ideology of elected women vs elected men

	(1)	(2)	(3)	(4)	(5)
	Far-left	Left	Liberal	Right	Far-right
Woman	0.008	0.288***	-0.004	-0.248***	-0.048
	(0.021)	(0.069)	(0.056)	(0.036)	(0.036)
Observations	1783	1783	1783	1783	1783
Eff. number of obs	907	827	849	1229	983
Robust p-value	0.758	0.000	0.732	0.000	0.248
Polyn. order	1	1	1	1	1
Bandwidth	6.609	5.918	6.147	9.513	7.164
Outcome mean	0.044	0.224	0.264	0.340	0.084

In columns (1)-(5), the outcome variable is a dummy equal to 1 if the politician belongs to the referred ideology, 0 otherwise. Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the politician. The independent variable is a dummy equal to 1 if it is a woman and 0 if it is a man. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table E27: Balancing tests - differences in the ideology of the opponents of elected women vs elected men

	(1)	(2)	(3)	(4)	(5)
	Far-left	Left	Liberal	Right	Far-right
Woman	0.012	-0.018	-0.094	0.056	-0.006
	(0.012)	(0.059)	(0.060)	(0.062)	(0.044)
Observations	1783	1783	1783	1783	1783
Eff. number of obs	631	938	854	1061	1070
Robust p-value	0.262	0.731	0.080	0.348	0.946
Polyn. order	1	1	1	1	1
Bandwidth	4.382	6.905	6.189	7.795	7.904
Outcome mean	0.039	0.245	0.272	0.299	0.094

In columns (1)-(5), the outcome variable is a dummy equal to 1 if the elected politician is a woman, 0 otherwise. Robust standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. The unit of observation is the politician. Each column reports the results from a separate local polynomial regression. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure.

Table E28: Differences between female and male MPs during parliamentary work (without controlling for unbalanced covariates)

	(1) Debates	(2) Economy & employment	(3) Environment	(4) Health & education	(5) Security & foreign policy	(6) Local politics
Legislative debates						
Woman	-6.505	0.007	-0.026	0.114***	-0.021	-0.021
	(6.622)	(0.028)	(0.026)	(0.030)	(0.031)	(0.021)
Observations	1334	1334	1334	1334	1334	1334
Eff. number of obs	499	746	859	489	751	532
Robust p-value	0.320	0.859	0.410	0.000	0.446	0.377
Polyn. order	1	1	1	1	1	1
Bandwidth	4.425	6.978	8.320	4.289	7.010	4.797
Outcome mean	49.225	0.220	0.159	0.148	0.267	0.092
Written questions						
Woman	-81.309**	-0.050***	-0.037	0.062***	0.018	-0.021
	(35.848)	(0.017)	(0.023)	(0.023)	(0.025)	(0.017)
Observations	1451	1451	1451	1451	1451	1451
Eff. number of obs	677	797	648	807	745	682
Robust p-value	0.049	0.010	0.118	0.017	0.530	0.356
Polyn. order	1	1	1	1	1	1
Bandwidth	5.639	6.864	5.357	6.941	6.406	5.671
Outcome mean	133.090	0.329	0.202	0.248	0.240	0.072

Standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest is a woman being elected as an alternative to a man. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under the MSERD procedure. In columns (2)-(6), the dependent variable takes values between 0 and 1, and the computation is explained in Section 5.1. The mean gives the average outcome value for elected males at the threshold.

Table E29: Differences between female and male MPs during parliamentary work (half of MSERD bandwidth)

	(1) Debates	(2) Economy & employment	(3) Environment	(4) Health & education	(5) Security & foreign policy	(6) Local politics
Legislative debates						
Woman	-5.988	-0.021	-0.030	0.161***	0.004	-0.019
	(7.822)	(0.038)	(0.034)	(0.043)	(0.040)	(0.028)
Observations	1334	1334	1334	1334	1334	1334
Eff. number of obs	271	383	440	268	419	284
Robust p-value	0.445	0.564	0.414	0.000	0.953	0.505
Polyn. order	1	1	1	1	1	1
Bandwidth	2.200	3.364	3.841	2.161	3.662	2.394
Outcome mean	48.861	0.212	0.161	0.133	0.264	0.091
Written questions						
Woman	-61.945	-0.056**	-0.057*	0.072**	0.034	-0.009
	(39.637)	(0.022)	(0.031)	(0.030)	(0.036)	(0.025)
Observations	1451	1451	1451	1451	1451	1451
Eff. number of obs	350	420	338	439	386	356
Robust p-value	0.135	0.014	0.070	0.019	0.364	0.756
Polyn. order	1	1	1	1	1	1
Bandwidth	2.782	3.399	2.670	3.517	3.160	2.824
Outcome mean	113.795	0.315	0.219	0.228	0.260	0.065

Standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest is a woman being elected as an alternative to a man. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under half of the MSERD procedure. In columns (2)-(6), the dependent variable takes values between 0 and 1, and the computation is explained in Section 5.1. The mean gives the average outcome value for elected males at the threshold.

Table E30: Differences between female and male MPs during parliamentary work (double of MSERD bandwidth)

	(1) Debates	(2) Economy & employment	(3) Environment	(4) Health & education	(5) Security & foreign policy	(6) Local politics
Legislative debates						
Woman	-4.647	-0.010	-0.022	0.096***	0.005	-0.015
	(5.350)	(0.021)	(0.021)	(0.022)	(0.023)	(0.015)
Observations	1334	1334	1334	1334	1334	1334
Eff. number of obs	891	1161	1224	884	1192	952
Robust p-value	0.488	0.581	0.729	0.001	0.789	0.512
Polyn. order	1	1	1	1	1	1
Bandwidth	8.800	13.455	15.365 8.643	14.649	9.574	
Outcome mean	52.502	0.224	0.156	0.120	0.266	0.093
Written questions						
Woman	-55.733**	-0.051***	-0.022	0.054***	0.017	-0.025**
	(28.312)	(0.013)	(0.019)	(0.017)	(0.019)	(0.013)
Observations	1451	1451	1451	1451	1451	1451
Eff. number of obs	1146	1266	1121	1273	1219	1158
Robust p-value	0.266	0.009	0.238	0.086	0.447	0.516
Polyn. order	1	1	1	1	1	1
Bandwidth	11.127	13.597	10.680	14.068	12.641	11.295
Outcome mean	126.395	0.321	0.221	0.237	0.243	0.070

Standard errors are in parenthesis. Statistical significance is computed based on the robust p-value, and ***, **, and * indicate significance at 1, 5, and 10, respectively. Each column reports the results from a separate local polynomial regression. The variable of interest is a woman being elected as an alternative to a man. Separate polynomials are fitted on each side of the threshold. The polynomial order is 1, and the optimal bandwidths are derived under double of the MSERD procedure. In columns (2)-(6), the dependent variable takes values between 0 and 1, and the computation is explained in Section 5.1. The mean gives the average outcome value for elected males at the threshold.