CESifo AREA CONFERENCES 2020

Public Economics

Online, 8–9 December 2020

Gender Gaps in Political Careers: Evidence from Competitive Elections

Davide Cipullo

CESifo

Gender Gaps in Political Careers: Evidence from Competitive Elections*

Davide Cipullo†

November 2, 2020

Click here for latest version

Job market Paper

Abstract

This paper investigates the impact of voter support on the representation of women in the political profession. The empirical analysis exploits two-stage elections in the United States and Italy to hold the selection of candidates constant. In two-stage elections, candidates are admitted to the second round of voting based on the outcome of the first round. I find that among candidates who marginally qualify for the final round, women are 20 percent less likely than men to be elected to the US House of Representatives and 40 percent less likely to be elected mayor in Italian municipalities. Using a difference-in-discontinuities design, I then show that the gender gap in the probability of being elected has long-lasting effects on career trajectories. Women are substantially less likely than men to win future elections and to climb the political hierarchy. My findings suggest that one of the reasons that few women reach the top in politics is that female candidates face hurdles at the beginning of their careers.

JEL Classification: C24, D02, D72, J16
Keywords: Gender gaps, Self-selection, Political careers, Sticky floor, Voting

*I would like to thank Eva Mörk and Luca Repetto for extensive feedback and Michael Blanga-Gubbay, Alessandra Bonfiglioli, Pamela Campa, Enrico Cantoni, Felipe Carozzi, Oscar Erixson, Jon H. Fiva, Olle Folke, Peter Fredriksson, Georg Graetz, Christoph Hedtrich, Matthew O. Jackson, Ina Jäkel, Sebastian Jävervall, Mounir Karadja, Arizo Karimi, Daniel Knutsson, Barton E. Lee, Nicola Maaser, Elena Mattana, Torben Mideksa, Mariona Mas Montserrat, Daniele Nosenzo, Yoko Okuyama, Vincent Pons, Erik Prawitz, André Reslow, Alex Solis, Alexandra Spitz-Oener, Anna Thoresson, Björn Tyrefors, Ekaterina Zhuravskaya, an anonymous discussant at the ZEW Public Finance Conference 2020, and the participants in the seminars held at the 35^{th} Annual Congress of the European Economic Association (EEA), the 76^{th} annual Congress of the International Institute of Public Finance (IIPF), the 32^{nd} Annual Conference of the Italian Society of Public Economics (SIEP), Harvard University, Aarhus University, the Research Institute for Industrial Economics in Stockholm (IFN), Instituto Tecnológico Autónomo de México, Uppsala Centre for Fiscal Studies, Uppsala University, and the Stockholm-Uppsala Political Economy Meeting for their dialogues and comments. All errors are mine. A research grant from the Jan Wallander och Tom Hedelius Foundation (P2017-0185:1) is gratefully acknowledged.

†Department of Economics, Uppsala University, and Uppsala Centre for Fiscal Studies. Address: Box 513, SE-75120, Uppsala, Sweden. E-mail: davide.cipullo@nek.uu.se.

1 Introduction

Despite substantial progress in female labor force participation and earnings over the last decades, women remain underrepresented in top positions. For instance, even if women account for as much as 57 percent of all college graduates in the US (OECD, 2020), in 2017, only one-third of lawyers (Azmat and Ferrer, 2017) and one-fourth of senior managers (Thortnton, 2017) were women.[1] Lack of female representation in top jobs is particularly pervasive in the political arena, as less than eight percent of countries have a female head of government (Bell, 2020). Additionally, women are underrepresented in parliaments and are less likely than men to run for political office (Fox and Lawless, 2004).

Identifying the determinants of gender gaps in political representation presents several empirical challenges. In particular, individual returns from running in an election – which depend on voters' support – interact with the selection of candidates. Indeed, the probability of being elected depends on the interaction of decisions made by candidates, parties, and voters: voters choose among candidates who have decided to run and who have received the support of their party. Failing to empirically separate returns and selection may lead to inaccurate conclusions about gender gaps. For instance, suppose women are less likely to participate in politics because they expect parties or voters not to reward them fairly. In such a scenario, only the best women ultimately end up on a ballot, which one may erroneously interpret as evidence of a gender gap in the propensity to participate in politics. Additionally, the positively selected women may end up performing as well as or even better than their male opponents in an election, leading one to mistakenly conclude that there is no gender gap in voters' support.[2]

This paper exploits two-stage elections to estimate the impact of voters on the representation of women in the political profession, holding (self- or party-appointed) selection of candidates constant. I provide evidence based on two different settings: partisan primaries for the election of members of the US House of Representatives and runoff elections of mayors in Italy. The research design exploits the fact that in both systems, politicians who barely make it to the final round and those who just miss the qualification threshold are arguably comparable in observed and unobserved characteristics. Combining evidence from two settings strengthens the external validity of the empirical results. More specifically, my empirical analysis allows me to assess whether gender differences in voters' support are present across heterogeneous contexts in terms of institutions and voter characteristics.

I find that qualifying to the final round increases the chances of winning an election

[1]Moreover, the gender wage gap is larger at the top of the wage distribution (e.g., Albrecht et al., 2003), a phenomenon referred to as the "glass ceiling".

[2]Hsieh et al. (2019) make an analogous argument in describing the underrepresentation of women and individuals from minority groups among lawyers and doctors.

substantially more for men than for women. Female candidates who barely win the primary are 7 percentage points less likely to be elected to the House than their male counterparts. Similarly, female candidates who barely qualify for the second round of a runoff election in Italy are 9 percentage points less likely to become mayor than males. In relative terms, my findings imply that women have a 20 percent lower likelihood than men of becoming a member of the House of Representatives and a 40 percent lower probability of winning a mayoral election in Italy. To ensure that my results capture voter behavior, and not men and women competing in different settings, I show that barely qualified male and female candidates run in comparable elections. More precisely, the predicted probability of winning the seat based on first-round determinants and outcomes is equal for female and male candidates who barely qualify.

In addition to showing that male and female candidates who run in comparable elections differ in the support that they receive from voters (and the channels through which such divergence may occur), this paper investigates whether the gender gap in the probability of being elected has long-lasting effects on career trajectories. I implement a cross-sectional difference-in-discontinuities design (Eggers et al., 2018; Grembi et al., 2016). More specifically, I compare gender differences in the future careers of candidates who marginally qualify for the second round with the corresponding differences among those who just miss the threshold to qualify.

The gender gap in the probability of winning the election has strong and persistent effects on career trajectories. Winning the primary increases both the probability of receiving a nomination and being elected representative in future elections by 4-6 percentage points less for female than for male politicians. Gender gaps in career returns are strong and persistent, as female candidates are penalized at least until the third downstream election. In Italian municipalities, qualification for the second round increases the probability of winning the election in the next term by 6 percentage points less for women than for men. Moreover, the returns from qualifying for the final round in terms of the probability of promotion to higher levels of the political hierarchy are 5 percentage points lower for women. In relative terms, future returns from electoral competition are 30-50 percent higher for men than for women in both contexts. The results for the US House of Representatives also indicate that gender differences in voter support discourage women from participating in subsequent elections: while winning the primary increases the probability of competing in the next election for men and women equally, it gives rise to a gender gap in the decision to run starting from the second downstream election.

Qualification for the final round may have a more positive impact on the future careers of males either because men are more likely to be elected – and thus to exploit future incumbency advantages – or because women also have a lower probability of being reelected, conditional on winning the seat once. In line with the *sticky floor* hypothesis (e.g., Bjerk, 2008), according to which individuals from underrepresented groups are less likely to climb the career ladder to the top because they face early challenges, I find that a marginal victory in an election increases the probability of future success and promotion equally for men and women.

The existing literature proposes a range of potential explanations for gender gaps. Potential determinants include discrimination (e.g., Becker, 1957; Charles et al., 2018; Goldin and Rouse, 2000); differences in productivity, potentially due to parenthood (e.g., Azmat and Ferrer, 2017; Kleven et al., 2019); sorting across firms (e.g., Barth et al., 2017; Card et al., 2015); and differences in attitudes towards competition (e.g., Bertrand, 2011) and negotiation (e.g., Card et al., 2015).[3] Taking advantage of the two empirical settings, I rule out the possibility that my results are driven by differences in campaigning and negotiating behavior between male and female candidates. On the one hand, I find no gender differences in the amount of campaign contributions raised and invested by US congressional candidates before and after the primary. On the other hand, I document that male and female candidates receive the endorsement of equally many parties in between the two rounds of runoff elections in Italy.

My findings appear consistent with models of statistical discrimination (e.g., Arrow et al., 1973; Phelps, 1972) against female candidates among voters. I exploit a number of dimensions of heterogeneity with respect to candidates' and districts' characteristics to correlate estimated gaps in voters' and candidates' backgrounds. The results show that gender differences are driven by candidates without previous experience and by districts where no women previously served in the same position and where women are relatively less integrated in the labor market. In contrast, male and female candidates perform equally well if they previously competed for the same elected office in the past and in places where a woman previously held the same office. Moreover, I find that voters are more likely to support female candidates in districts where the economy is growing less than in neighboring districts (when, arguably, voters may seek a change) than in good times (when voters confirm the current political leadership).

This paper makes several contributions to the literature. First and foremost, by proposing a novel empirical strategy to separate gaps in returns from gaps in selection, I contribute to the literature on gender gaps in the economy. Second, I contribute to the literature analyzing the causes and consequences of gender differences in political success. Potential determinants include a lack of political ambition (e.g., Fox and Lawless, 2004, 2014), differences in opportunity costs (e.g., Folke and Rickne, 2020), incumbency advantages penalizing newcomers (e.g., Shair-Rosenfield, 2012), and voter or party bias (e.g., Gonzalez-Eiras and Sanz, 2018; Karpowitz et al., 2017; Lawless and Pearson, 2008).[4]

[3]See Gneezy et al. (2003) and Buser et al. (2014) for experimental evidence of gender differences in attitudes towards bargaining. See also Biasi and Sarsons (2020) for evidence of gender differences in the propensity to negotiate after the introduction of individual wage bargaining for teachers in Wisconsin.

[4]Casas-Arce and Saiz (2015) exploit the introduction of gender quotas in Spain to document that before the introduction of quotas, parties were not nominating enough women as candidates to maximize

I show that early gaps in the probability of winning an election lower the probability of climbing the career ladder in the future even after taking into account gender differences in ambition, party support, and campaigning effort. This evidence suggests that few women make it to the top in politics because they lack voter support in early rather than in later steps of their career.

My findings are also relevant to policy outcomes since the identity of political leaders is predicted to matter for economic outcomes (e.g., Besley and Coate, 1997; Osborne and Slivinski, 1996). Moreover, women and men have been shown to have different political preferences over redistribution (e.g., Alesina and Giuliano, 2011; Miller, 2008), size of the government (Lott and Kenny, 1999) and health, environmental, and defense policy (Funk and Gathmann, 2014).[5] Understanding how voting decisions depend on the gender of competing candidates is valuable both for the design of policies that aim to reduce gender gaps and for the interpretation of how elections shape the economy.

The paper proceeds as follows. In Section 2, I show under which assumptions two-stage elections allow us to separate voter support for given candidates (i.e., the returns from running) from selection into an electoral competition. Section 3 endows the reader with relevant information on women's participation in politics and voting systems in the two countries included in the analysis. In Section 4, I outline the empirical strategy, while in Section 5, I present the empirical results and conduct a battery of robustness checks. Finally, Section 6 discusses the results in light of the existing literature and of differences between the two empirical settings, and Section 7 concludes.

2 Exploiting two-stage elections to identify the returns from electoral competition

In this section, I first describe why a comparison of the probability of winning an election of male and female candidates who self-selected into an electoral competition does not separate the propensity to participate in politics and returns from participation in an election. Second, I show how two-stage elections can be used to identify gender differences in returns from the electoral competition.

the electoral results. On the other hand, Le Barbanchon and Sauvagnat (2018), exploiting the introduction of gender quotas in France, find that parties are more likely to nominate male candidates as a response to voter bias against women.

[5] It remains an open question whether the election of a female or a male politician leads to alternative policy outcomes. On the one hand, Ferreira and Gyourko (2014) and Carozzi and Gago (2017) do not find any differences across municipalities in the US and Spain, respectively. On the other hand, Chattopadhyay and Duflo (2004) in India, Svaleryd (2009) in Sweden, and Brollo and Troiano (2016) in Brazil find that male and female officials implement different policies. Lastly, Casarico et al. (2020) document that Italian municipalities under the runoff system spend and raise more when a woman is elected mayor. In addition, male and female legislators prioritize different policies, as estimated in the contexts of close elections and of the introduction of gender quotas (e.g. Bagues and Campa, 2020; Lippmann, 2019).

Suppose that the probability of winning an election depends on the following process:

$$Y_i = \eta + \gamma F_i + u_i, \tag{1}$$

where Y_i is an indicator equal to 1 if candidate i is elected and F_i is an indicator equal to 1 if the individual is a woman. The error component u_i represents all determinants of electoral success that do not depend on gender. Thus, if voters are asked to cast their vote based on a randomly determined pool of candidates, the gender difference in the probability of winning the election would identify γ. The coefficient γ measures the presence of gender differences in returns from participating in an electoral competition and indicates whether male or female candidates are more likely to receive support from voters. γ may reflect the effect of personal traits correlated with gender, productivity, discrimination, or a mix of the above on the probability of winning the election. For instance, a negative γ would imply that voters are less likely to vote in favor of a woman than a man or similarly that returns from participating in an electoral competition are, on average, lower for women than for men.

However, the decision to participate in an electoral competition is not random. It is a consequence of an unobserved selection process that may depend on personal traits, party support, and expectations of voter support.[6] To simplify, suppose that the participation decision depends on

$$z_i = \mathbf{1}\{\alpha F_i + v_i > 0\}, \tag{2}$$

where z_i is an indicator equal to 1 if the individual competes in the election. Consistent with the interpretation of γ in equation (1), α summarizes gender differences in the propensity to compete, while the error term v_i captures any other determinant of the decision to compete.

If the error terms u_i and v_i are correlated, which is the case if the participation decision is a function of the expected probability of winning the election, the average gender gap among candidates (i.e., individuals with $z_i = 1$) fails to identify the differences in returns γ. Formally,

$$\mathbb{E}(Y_i|F_i = 1, z_i = 1) - \mathbb{E}(Y_i|F_i = 0, z_i = 1) = \gamma + \underbrace{\mathbb{E}(u_i|F_i = 1, \alpha + v_i > 0)}_{\text{Selection for female}} - \underbrace{\mathbb{E}(u_i|F_i = 0, v_i > 0)}_{\text{Selection for male}}. \tag{3}$$

As long as the selection stage in equation (2) is different across genders (i.e., $\alpha \neq 0$),

[6] See, e.g., Osborne and Slivinski (1996) and Besley and Coate (1997) for a theoretical model of electoral competition among an endogenously determined set of candidates. Citizen candidates pay a cost attached with the decision to run based on the expected probability of victory, given the distribution of policy preferences, and are systematically different from citizens who do not compete for office.

the gender gap in equation (3) consists of two components: first, the gap in returns from participation γ and, second, the direct effect of the selection process of male and female candidates on the probability of being elected. More specifically, a simple comparison of mean election rates overestimates γ when $\alpha < 0$ and underestimates γ if $\alpha > 0$. The case is analogous to that of labor markets, where returns are measured by wages. A simple comparison of average wages of men and women who work in the same profession fails to identify gender gaps in returns if selection into the profession is unobserved and differs for men and women.

Heckman (1979) proposes a correction for sample selection using a two-step procedure. First, the probability of selection is predicted based on observable characteristics. Second, among individuals with $z_i = 1$, equation (1) is estimated, augmented with a term that controls for the correlation between selection and the outcome. The Heckman (1979) method rests on strong assumptions, which include i) the necessity of observing all nonrandom determinants of selection for individuals who are not in the profession and ii) imposition of a specific structure on the joint distribution of u_i and v_i.[7] This paper proposes a different approach that rests on milder assumptions than those required for the Heckman (1979) correction. In two-stage elections, where multiple candidates compete in a first round and only the top vote recipients among them qualify for the final stage, selection into the final round is observed.[8] Moreover, in a narrow window around the threshold, nonparametric approximations can proxy for any functional relationship between selection and an outcome in the selected subsample.

Formally, in the presence of two rounds of elections, equation (2) can be expressed as $z_i = \mathbf{1}\{h_i > 0\}$, where h_i is the *qualification margin*, defined as the vote share distance between candidate i and the closest candidate who fails to qualify (if candidate i qualifies for the final round) or the vote share distance between the closest candidate who qualifies for the final round and candidate i (if candidate i does not qualify). If the qualification margin absorbs the unknown relationship between selection and outcome among candidates who barely qualify for the final round (i.e., with $h_i \to 0^+$), so that

$$\lim_{h_i \to 0^+} \mathbb{E}(u_i | F_i = 1, z_i = 1) = \lim_{h_i \to 0^+} \mathbb{E}(u_i | F_i = 0, z_i = 1), \qquad \text{(Assumption I)}$$

[7] More specifically, the Heckman (1979) selection method estimates the probability of selection using the probit model and augments the outcome regression in the selected sample with a control for the inverse Mills ratio $\lambda_i = \frac{\phi(Z_i)}{1-\Phi(Z_i)}$, where $Z_i = -\frac{\mathbf{X_i}\beta}{\text{var}(v_i)^{\frac{1}{2}}}$ if $z_i = \mathbf{X_i}\beta + v_i$.

[8] A second round restricted to two candidates is the norm in runoff elections. France is an exception to this norm. In France, all candidates receiving at least 12.5 percent of the votes qualify to the final round and can decide whether to run or to endorse another candidate. See Pons and Tricaud (2019) for details.

then, equation (3) can be expressed as

$$\lim_{h_i \to 0^+} \big[\mathbb{E}(Y_i|F_i = 1, z_i = 1) - \mathbb{E}(Y_i|F_i = 0, z_i = 1)\big] = \gamma. \tag{4}$$

Assumption I requires that men and women who marginally qualify for the final round enter that round with equal prior probability of winning the election.[9] The validity of this assumption can be assessed by predicting the probability of winning an election based on first-round determinants and showing that around the threshold, this probability is uncorrelated with the gender of the candidate.

A comparison of marginally qualified candidates identifies the gender gap in returns from participating in an electoral competition and allows me to locally measure whether voters are more likely to support a man than a woman among candidates facing comparable competitions.[10] However, note that the presence of a gender gap in the probability of winning an election does not necessarily imply that voters are biased against women. Gender differences in the probability of being elected may also imply that men are more attractive to voters than women because of other characteristics correlated with gender.

In Appendix C, I show, using numerical simulations, that in male-dominated professions, where unobserved determinants of selection are highly correlated with gender, relatively small gender gaps in outcomes may hide substantial gender gaps in returns.

2.1 Future political career

While electoral success depends on selection into an electoral competition and on the returns from that electoral competition, future career success also depends on selection into the next electoral contest, which may vary by gender as well.[11]

To see this formally, suppose that the probability of being elected in the future can be described by the following process:

$$Y_i^c = \phi^c F_i + z_i\big(\eta^c + \gamma^c F_i\big) + u_i^c, \tag{5}$$

where Y_i^c is an indicator equal to 1 if the individual is elected in the future. ϕ^c measures gender differences in the decision to participate in the future, while $z_i = \mathbf{1}\{\alpha F_i + v_i > 0\}$ reflects unobserved selection into the current election as in equation (2). Lastly, γ^c measures gender differences in career returns from participation in an election. The

[9]Notice that $h_i = \alpha F_i + v_i$ is correlated with the gender of the candidate and controls for the differential selection process of males and females into the first round as well as any differences in performance during the first round.

[10]Kline and Walters (2019) show that the Heckman (1979) method identifies a local average treatment effect for individuals close to the *latent* participation threshold. Similarly, the research design of this paper identifies the returns from participation in the election around the explicit qualification threshold.

[11]See, e.g., Baskaran and Hessami (2020), who establish the existence of a gender recontest gap, and Wasserman (2018), who finds that women are more likely to quit a competition than men.

difference in the probability of being elected in the future between men and women who currently compete for office (i.e., with $z_i = 1$) is equal to

$$\mathbb{E}(Y_i|F_i = 1, z_i = 1) - \mathbb{E}(Y_i|F_i = 0, z_i = 1) = \\ \gamma^c + \underbrace{\phi^c}_{\text{Future selection}} + \underbrace{\mathbb{E}(u_i^c|F_i = 1, \alpha + v_i > 0)}_{\text{Current selection for female}} - \underbrace{\mathbb{E}(u_i^c|F_i = 0, v_i > 0)}_{\text{Current selection for male}}. \tag{6}$$

The approach introduced above would identify the sum $\gamma^c + \phi^c$ in this case but would not separate the two. Intuitively, the solution to separate γ^c from ϕ^c is to find a group of male and female individuals who are not exposed to returns from participating in an electoral competition, i.e., with the same ϕ^c as the individuals participating but with $\gamma^c = 0$.

In the case of two-stage elections, candidates who just miss the qualification threshold are the ideal control group for candidates who marginally qualify for the final round and are exposed to returns from it. More specifically, candidates of the same gender on opposite sides of the *qualification threshold* are arguably equal in all characteristics, including ϕ^c, except the exposure to an additional round of voter choice.

Formally, if the error components in equation (5) in a narrow window around $h_i = 0$ are such that

$$\lim_{h_i \to 0^+} \mathbb{E}(u_i^c|F_i, z_i = 1) = \lim_{h_i \to 0^-} \mathbb{E}(u_i^c|F_i, z_i = 0), \tag{Assumption II}$$

then

$$\lim_{h_i \to 0^+} \left[\mathbb{E}(Y_i^c|F_i = 1, z_i = 1) - \mathbb{E}(Y_i^c|F_i = 0, z_i = 1)\right] - \\ \lim_{h_i \to 0^-} \left[\mathbb{E}(Y_i^c|F_i = 1, z_i = 0) - \mathbb{E}(Y_i^c|F_i = 0, z_i = 0)\right] = \gamma^c. \tag{7}$$

Equation (7) takes the form of a cross-sectional difference-in-discontinuities design (Eggers et al., 2018; Grembi et al., 2016) and identifies gender differences in career returns from qualifying to the final round of voting. Equation (7) estimates the differential impact of voter support on the career outcomes of male and female candidates, holding constant all characteristics correlated with the gender of the candidate that do not interact with voter decisions. Consistent with equation (4), a gender difference in future career returns may imply that voters are biased against women or that different characteristics of male and female candidates that voters deem valuable also have a long-run impact. Assumption II is less restrictive than Assumption I because it does not require that male and female candidates at the threshold have the same expected returns from participating in the election.

In Section C in the Appendix, I compare the gap in future returns estimated using

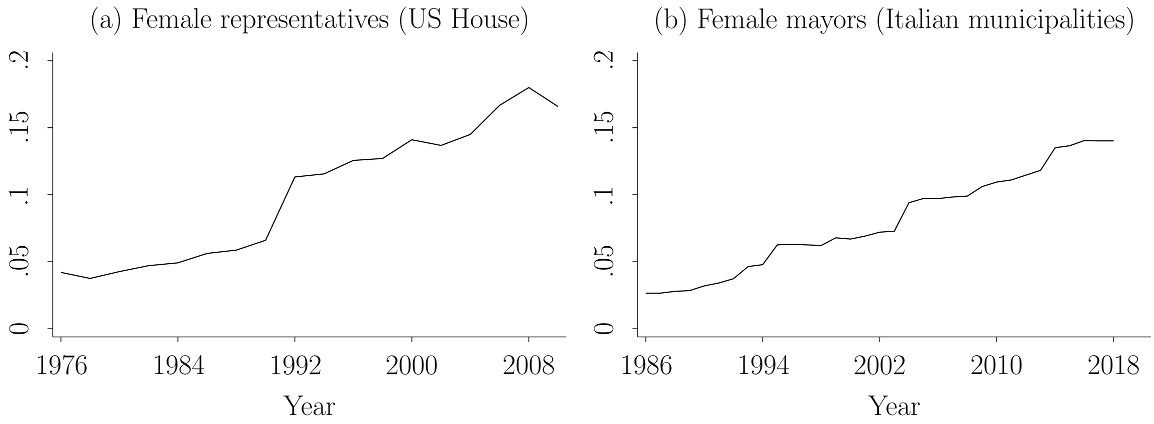

Figure 1: Share of female office holders, US House of representatives and Italian municipalities

Notes: Share of women in the US House of Representatives between 1976 and 2010 (Panel (a)) and serving as mayors in Italian municipalities between 1986 and 2019 (Panel (b)).

difference-in-discontinuities with promotion gaps estimated not taking into account the two selection stages as in equation (6) or not taking into account gender differences in ϕ^c. The results of the simulation show that not accounting for selection when gender affects both current and future decisions to participate in an electoral competition leads to substantial overestimation of differences in returns.

3 Background and data

The empirical analysis builds on the features of the two most widely implemented types of two-stage electoral institutions: primary and runoff elections. First, I exploit the election of members of the US House of Representatives, for which each party nominates a candidate to the general election based on the outcome of a party primary. Second, I exploit the election of mayors in large Italian municipalities, where mayors are elected using a runoff voting system. The two settings differ in several institutional features, allowing me to explore a wide range of mechanisms and to strengthen the external validity of the estimates. Each setting allows me to apply the intuition presented in Section 2 and represents a case of democracy in which the representation of women in politics is steadily increasing but still low (see Figure 1).

3.1 Party primaries for the US House of Representatives

The US House of Representatives is the lower branch of the US Congress. A total of 435 voting members are elected every other year using the plurality system in single-member districts of approximately equal population size. Parties select their candidates for the general election by organizing primaries.[12] Primary elections are held on a day determined by the state and take place several months before the general election, which takes place on the first Tuesday of November. Candidates who lose their party primary are not forced to support the winner, but they do so in the vast majority of cases. Members of the House do not face term limits.

Although the first woman was elected to the US House in 1916, women are still underrepresented in US politics. Between 1976 and 2010, women accounted for only approximately 10 percent of candidates in US House primaries and of elected members of the House. Figure A1 shows that female representatives are mainly elected in states in the northeast and west, although women are substantially underrepresented throughout the country.

The analysis in this paper builds on individual-level primary election returns from Pettigrew et al. (2014), publicly available on Harvard University's *Dataverse*, and general election returns from the Massachusetts Institute of Technology *Election Data + Science Lab*. I observe information on the congressional district, gender, experience in elected office, incumbency status, and party affiliation as well as the total number of valid votes and votes received by each candidate.[13] I follow the career path of each candidate over time in terms of the decision to compete in future primaries, the likelihood of winning future primaries, and the likelihood of future election to the House. In the final dataset, I observe 24,080 candidates, among which 2,727 women competed in 13,774 primaries (equivalent to 7,478 seats in the US House of Representatives). I complement these data with information on campaign activity and financing from the Federal Election Commission for a subsample of candidates competing between 1980 and 2010 (see Appendix D for details). Table B1 reports the descriptive statistics of all relevant variables. Table B1 shows that 28 percent of candidates are incumbents and that the number of valid votes increases on average from 35,000 to 184,000 between the primary and the general election. Thus, the outcome of the general election mostly depends on voters who did not participate in the selection of party nominees.

[12]Primary rules are decided at the state level, with nonnegligible variation, even if voters are usually unable to vote in the primary election of more than one party. The main sources of variation are the electoral formula, with the single-ballot plurality rule being the most used system; the composition of the population eligible to vote, which can either be equal to the universe of registered voters in the district or constrained to the voters affiliated with the organizing party; and the holding of either a primary election or multiple caucuses.

[13]The dataset also contains information about the characteristics of each primary (whether it was an open or a closed primary and whether the outcome was determined in a regular primary or by caucuses).

3.2 Runoff municipal elections in Italy

Italy is divided into three subnational administrative levels (regions, provinces, and municipalities), of which municipalities represent the lowest level in the hierarchy. As of March 2020, Italy has 7,904 municipalities.

The head of the executive power at the municipal level is the mayor, who is directly elected under majority rules. Mayors are elected using either the single-ballot plurality rule or the dual-ballot runoff system based on the official population computed in the last census. The runoff voting system is in place in municipalities with more than 15,000 residents; these municipalities are approximately 10 percent of the total. The candidate with the largest vote share is elected mayor only if she receives more than 50 percent of the valid votes in the first round. Otherwise, the two candidates with the most votes compete in a second round two weeks later. On the same election day, voters elect the members of the municipal council with a proportional rule. Each candidate is aligned with a coalition of parties running for seats in the municipal council. The parties of the elected mayor are assigned a majority premium that guarantees them a stable majority in the assembly.[14]

Between the first and the second round, parties supporting excluded candidates have the possibility of endorsing one of the two remaining candidates, conditional on her acceptance. As shown in Figures A2 and A3, the voting ballot signals endorsements to voters, as the endorsing parties appear on the ballot beside the candidate.[15] In the case of victory, endorsing parties receive a share of the majority premium seats. Terms last five years, and mayors are limited to serving two consecutive terms. If the mayor resigns, early elections are held for both the mayor and the members of the municipal council.[16]

The representation of women in politics in Italy has traditionally been meager at all levels, especially in leadership positions. Since the transition to democracy, all of the presidents, prime ministers (*Presidente del Consiglio*), and finance ministers of the Republic have been men. Despite a steady increase since the eighties, women serve

[14]The coalition of the elected mayor receives 60 percent of the seats. Voters are allowed to cast a *disjoint* vote (i.e., to vote for a party in the coalition of another mayoral candidate), and the allocation of seats among the parties in the coalition of the elected mayor takes place on a proportional basis. The mayor is not awarded the majority premium in two limit cases: first, when the absolute majority of votes in the election for the council go to a coalition of parties not supporting the elected mayor and, second, if the mayor is elected during the first round and her coalition receives less than 40 percent of the votes. Between 1993 and 1999 [Law 120/1999], the latter vote threshold was fixed to 50 percent. In both cases, all seats of the municipality council are assigned proportionally to lists that received more than 3 percent of the votes.

[15]Figures A2 and A3 show an example of ballots used in the first and second rounds of municipal elections. Michelangelo Betti appeared as the candidate for a coalition of six parties in the first round. Before the second round, he received the endorsement of the parties in the coalition of the excluded candidates, Cristiano Masi and Fabio Poli. In turn, Michelangelo Betti was the candidate for a coalition of ten parties in the second round.

[16]Gagliarducci and Paserman (2011) find that female mayors are more likely than their male colleagues to be unseated before the natural end of the term. Example of cases leading to early polls include votes of no confidence and the resignation of the majority of the members of the municipal council.

as mayors in only approximately 14 percent of municipalities. Although the average proportion of female mayors is rather low, it hides substantial geographical variation, as presented in Figure A4. Figure A4 shows that the share of women serving as mayors is higher in Northern and Central Italy than in the South, where the percentage of female mayors has always been below 5 percent.

The main data source covers election results of all municipal elections held between 1993 and 2019 in the fifteen *ordinary regions*.[17] The data contain information on the identity of each candidate, her vote share and party affiliation, the number of parties in her coalition, the allocation of seats in the municipal council, and voter turnout. For municipalities in which a second round took place, I observe all variables for both rounds of the electoral competition. The final dataset covers 19,037 candidates, of which 2,245 are women, competing in 3,572 elections. I complement these data with information from the *Anagrafe degli amministratori locali* collection, which includes education, occupation, age, and other individual characteristics of politicians elected to any office since 1986 (see Appendix D for details). Since mayoral candidates who lose an election are automatically elected to the municipal council if they receive at least 3 percent of the valid votes, I also observe this information for the large majority of losing candidates.

Table B2 reports the descriptive statistics of the variables used in the empirical analysis. Only 1.3 percent of voters cast a *disjoint* vote between the candidate for the position of mayor and the lists of candidates for the municipal council. Endorsements are relatively common, with coalitions during the second round formed by 0.3 more parties than during the first round. However, turnout falls from 75 percent of the eligible population in the first round to 61 percent in the second round.

4 Empirical strategy

The two-stage nature of partisan primaries and runoff elections allows me to account for individual and party selection into an electoral competition because participation in the final round depends on the first round, which is observable. If male and female candidates who closely qualify for the final round enter it with the same prior probability of being elected, then it is possible to estimate whether voters are more likely to support men than women by comparing the returns from participating in the electoral competition. Moreover, two-stage elections allow me to separate voter support from individual and party participation decisions in future electoral competitions, using marginally excluded

[17] Municipalities in the fifteen *ordinary regions* have responsibilities including urban planning, waste disposal, environmental policy, road maintenance, and welfare and finance their expenditures through transfers from the central and regional governments as well as by levying a real estate tax and labor income tax. The five regions whose special autonomy is guaranteed by the national constitution (Friuli–Venezia Giulia, Sardinia, Sicily, Trentino–Alto Adige, and Valle d'Aosta) are excluded because their municipalities are subject to different voting systems.

candidates as a control group. Specifically, I estimate the following model:

$$Y_{i,t} = \beta_0 + \beta_1 F_i \times Qualified_{i,t} + \beta_2 Qualified_{i,t} + \beta_3 F_i + f(F_i; h_{i,t}) + \delta_t + \varepsilon_{i,t}, \quad (8)$$

where the running variable $f(F_i; h_{i,t})$ is a flexible function of the qualification margin $h_{i,t}$ and allowed to vary across genders and on either side of the qualification threshold and δ_t is a set of election-year dummies. The coefficient of interest is β_1, which measures the differential effect of qualification for female candidates in comparison with the effect for their male colleagues. For instance, a negative β_1 would imply – holding constant the selection of candidates and gender differences in the propensity to remain in the political profession in the future – that qualifying for the final round improves the career outcomes of male candidates more than those of female candidates.

The qualification margin measures the vote share distance between each candidate and the candidate closest to the qualification threshold but on the opposite side of it. Consistent with Pons and Tricaud (2019), I define the qualification margin $h_{i,t}$ as follows:

$$h_{i,t} = \begin{cases} VoteShare_{i,t} - VoteShare_{A,t} < 0 & \text{if } VoteShare_{i,t} < VoteShare_{A,t} \\ VoteShare_{i,t} - VoteShare_{B,t} > 0 & \text{if } VoteShare_{i,t} > VoteShare_{B,t} \end{cases}, \quad (9)$$

where $VoteShare_{A,t}$ and $VoteShare_{B,t}$ are the share of valid votes received by the closest candidate above and below the threshold, respectively. In the case of party primaries, where only the top candidate qualifies for the final round, $h_{i,t}$ captures the distance between candidate i and the runner-up if candidate i won the primary and the distance between candidate i and the top candidate if candidate i did not win the primary. In runoff elections, $h_{i,t}$ is the vote share distance between candidate i and the runner-up if candidate i did not qualify for the second round and the difference between the vote share of candidate i and the vote share of the candidate with the third-largest share of valid votes if candidate i qualified for the second round. In Appendix E, I extensively discuss the definition of $h_{i,t}$ and propose alternative specifications, which yield the same empirical results.[18]

As dependent variables, I use either an indicator equal to 1 if the candidate wins the seat or indicators of electoral success later in the candidate's career, depending on the analysis conducted. In the case of the US House of Representatives, where incumbent representatives are not subject to term limits, I focus on the decision to run for election, winning a primary, and being elected to the House for up to a third downstream election. For Italian municipalities, where mayors are limited to serving two consecutive terms, I focus on the next election and on an indicator equal to 1 if the candidate is elected to

[18] See Fiva et al. (2018) for an extended discussion of the definition of running variables for close-election regression-discontinuity designs when more than two candidates compete for a seat.

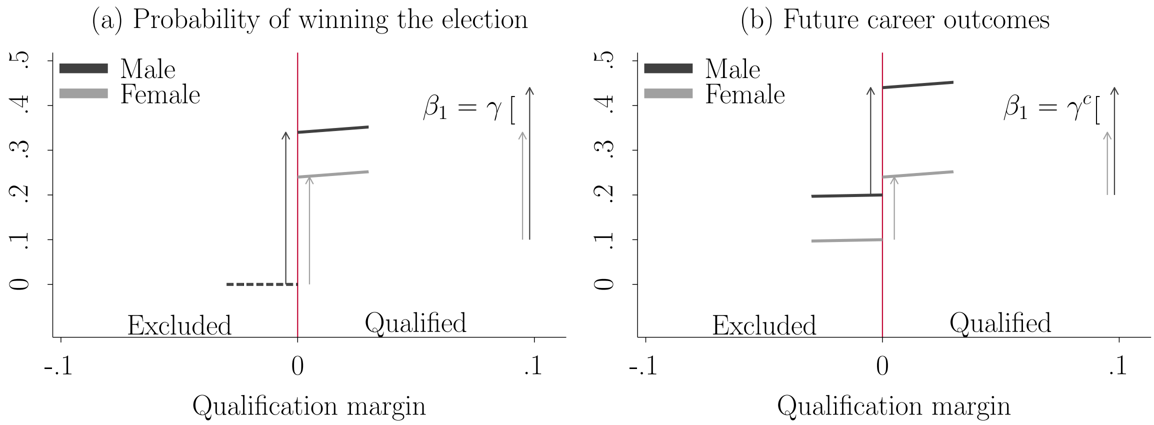

Figure 2: Illustration of the empirical strategy

Notes: This figure illustrates the intuition behind the empirical strategy used to estimate gender differences in returns from qualification. The horizontal axis represents the qualification margin $h_{i,t}$. In Panel (a), the dependent variable is an indicator equal to 1 if the candidate is elected in the final round. In Panel (b), the dependent variable is an indicator measuring the future career trajectory of the candidate.

office at the province level in the future.

For all outcomes, I graphically present the estimation results, controlling nonparametrically for the qualification margin. In regression tables, I fit local-linear regressions within the Calonico et al. (2014) optimal bandwidth as argued by Gelman and Imbens (2019).[19] Inference is based on standard errors robust to clustering at the congressional district-by-party level and at the municipal level.

Figure 2 summarizes the intuition behind the empirical strategy. Candidates who do not qualify for the final round are on the left of the qualification threshold, while candidates who do qualify for the final round are on the right. The light and dark gray arrows display the returns from qualifying for the final round for female and male candidates, respectively. In Panel (a), where the dependent variable is an indicator equal to 1 if the candidate is elected (and thus is always 0 for candidates excluded from the final round), gender differences among candidates just above the threshold estimate the average difference in returns (γ in equation (4)). In this case, $\widehat{\beta_1}$, the OLS estimator of β_1 in equation (8), is analogous to a difference estimator among barely qualified candidates since the coefficient β_3 is by construction restricted to zero. In Panel (b), where the dependent variable also varies between 0 and 1 among candidates excluded from the final

[19]For the sake of comparability and validation, I do not estimate different optimal bandwidths for male and female candidates. Therefore, in the regression tables, I require $f(F_i; h_{i,t})$ to be equal for candidates of both genders who lie on the same side of the qualification threshold. Specifications in which $f(F_i; h_{i,t})$ is allowed to vary across genders show that the interaction coefficients $F_i \times h_{i,t}$ and $F_i \times Qualified_{i,t} \times h_{i,t}$ are nonsignificant and have a t-statistic below one and a lower Akaike information criterion score than specifications in which $f(F_i; h_{i,t})$ does not vary across genders. Moreover, $F_i \times h_{i,t}$ and $F_i \times Qualified_{i,t} \times h_{i,t}$ are excluded in a post-Lasso regression. The results are available upon request.

round, controlling for differences in career outcomes among losers allows me to separate the effect of voter support from the propensity to compete in the future. In this case, β_1 (equivalent to γ^c in equation (7)) is estimated using a difference-in-discontinuities design, as summarized by the difference between the two arrows.

4.1 Validation of the identifying assumptions

As discussed in Section 2, the empirical strategy relies on different identifying assumptions depending on the outcomes of interest. When the dependent variable is the indicator equal to 1 if the candidate wins the election, the gender gap in returns is locally identified among candidates around the qualification threshold if Assumption I holds. Assumption I requires that male and female candidates who only just qualify for the final round approach the competition with equal probability of winning. When the dependent variable is the indicator equal to 1 if the candidate is elected or promoted in the future, differences in returns are identified under the less restrictive Assumption II.

Assumption II requires that i) absent qualification for the second round, the career paths of the marginal qualifying candidates and marginal losers of a given gender would have evolved equally and that ii) absent the second round, the career paths of politicians of the two genders would have followed local parallel trends with respect to the running variable (Grembi et al., 2016). Intuitively, Assumption II requires that candidates of the same gender on opposite sides of the threshold are equal in all characteristics and that the difference in outcomes among marginal losers reflects differences among barely qualified candidates had the final round not taken place. I present evidence in support of the continuity of potential outcomes in Figure A5 and Tables B3 and B4. In Figure A5, I document the results of formal McCrary (2008) tests, which do not show any evidence of manipulation of the qualification margins around the thresholds. In Tables B3 and B4, I show that candidate and election characteristics are balanced across the qualification threshold.

The validity of Assumption I, which is the most restrictive assumption, can be empirically assessed by showing that the predicted probability of being elected (based on a rich set of determinants and outcomes of the first round of the electoral competition) does not jump at the threshold. This approach is similar to that in Kirkeboen et al. (2016). More specifically, I estimate the following equation:

$$Y_{i,t} = \eta_e + \delta_t + \mathbf{X_{i,t}}\beta + \epsilon_{i,t}, \tag{10}$$

where η_e represents district-by-party and municipality fixed effects, δ_t is a set of election-year dummies, and $\mathbf{X_{i,t}}$ is a vector of controls for all first-round results at the individual candidate and election level. Table B5 and Table B6 show the results from estimating equation (10) among candidates for the US House of Representatives and for the position

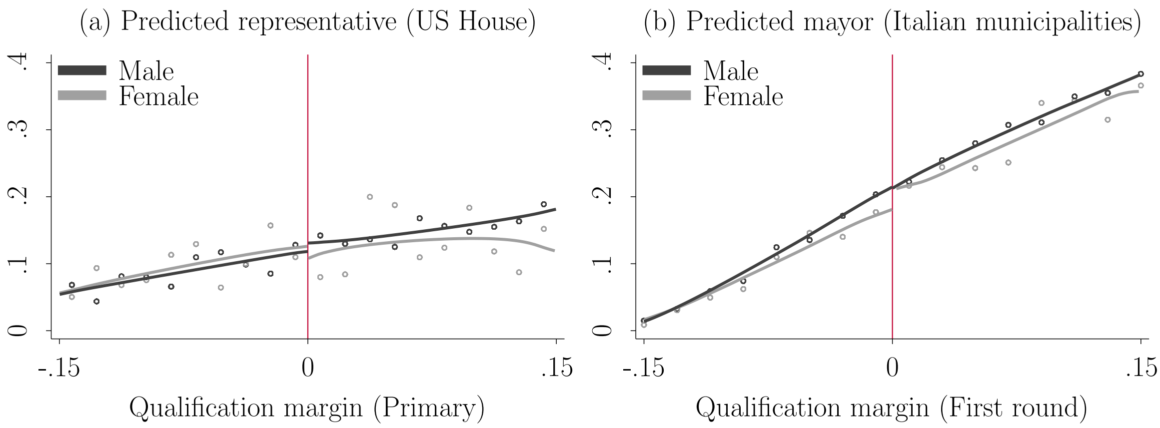

Figure 3: Gender of the candidate and predicted probability of winning the election

Notes: Panel (a): The dependent variable is the predicted probability of winning the election to the position of member of the US House of Representatives, estimated using a linear probability model as in equation (10). Panel (b): The dependent variable is the predicted probability of winning the election to the position of mayor, estimated using a linear probability model as in equation (10). In both panels, the solid lines represent non-parametric smoothers of the qualification margin, separately estimated on either side of the qualification threshold and for male and female candidates. Light gray lines and markers refer to female candidates, while dark gray lines and markers refer to male candidates. Markers represent, for each gender, sample averages within bins of the running variable $h_{i,t}$ equal to 0.015.

of mayor in Italian municipalities under a runoff. For each of the two settings, institutional details and data availability constrain the set of variables included in the regression. Nevertheless, in both tables, the included covariates account for 50–70 percent of the variation in the probability of winning the election.

The results are presented in Figure 3 and Tables B7 and B8.[20] In both settings, the predicted probability of winning the election given the first-round returns is uncorrelated with the qualification threshold and with the gender of the candidate. This evidence confirms that Assumption I is likely satisfied in both contexts and that elections in which male and female candidates compete in the second round are virtually identical. Therefore, any estimated gender gaps in the probability of winning an election among candidates that only just make it into the final round identify differences in voter support during the final round. Figure 3 and Tables B7 and B8 also provide reassurance on the validity of Assumption II because candidates that are only just excluded from the final round would have entered it with the same likelihood of success as candidates of the same gender who managed to qualify.

Additionally, Figures A8 and A9 in the Appendix show the gender balancing of first-round determinants and returns in the full sample of candidates and among candidates who were marginally admitted to the final round. The figures show that male and female

[20]See also Figures A6 and A7 for analogous results obtained when equation (10) is estimated using a probit or logit model.

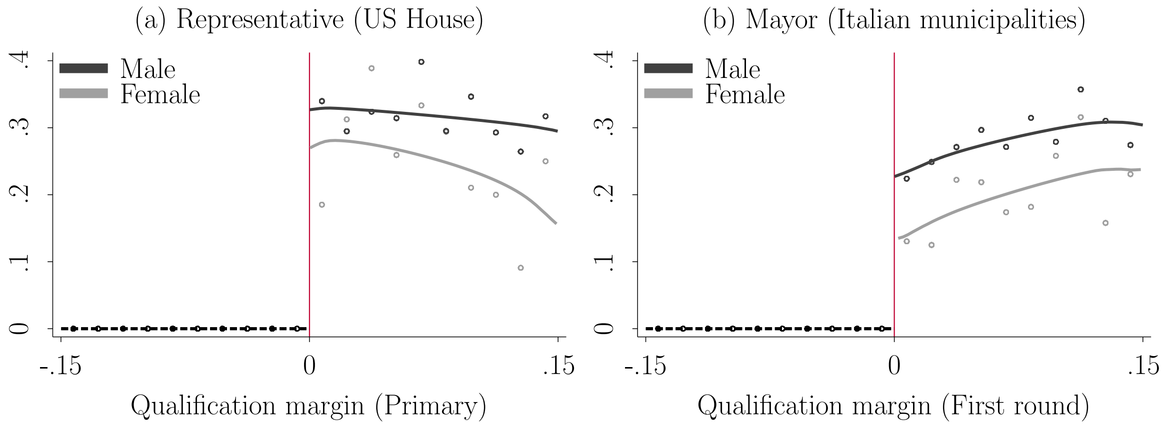

Figure 4: Qualification to final round and probability of winning the election, by gender

Notes: Panel (a): The dependent variable is an indicator equal to 1 if the candidate is elected as member of the US House of Representatives during the general election. Panel (b): The dependent variable is a dummy equal to 1 if the candidate is elected as mayor during the second round. In both panels, the solid lines represent non-parametric smoothers of the qualification margin, separately estimated on either side of the qualification threshold and for male and female candidates. Light gray lines and markers refer to female candidates, while dark gray lines and markers refer to male candidates. Markers represent, for each gender, sample averages within bins of the running variable $h_{i,t}$ equal to 0.015.

candidates marginally admitted to the final round perform equally well during the first round and are exposed to the same competitive environments. All estimated coefficients are small and not significantly different from zero. Notice that the same conclusion would not hold in the full sample of candidates admitted to the final round, where, on average, male and female candidates face different competitive environments. In Section 5.4, I present a battery of robustness checks, which further confirm the fulfillment of Assumption I.

5 Results

5.1 Gender gaps in the probability of winning an election

Figure 4 shows the estimated gender gap in the probability of being elected to the US House of Representatives and to the position of mayor in Italian municipalities. The results show that voters are significantly less likely to support women than men among candidates who are quasirandomly admitted to the competition and have the same prior probability of winning the election.

Compared to male candidates, who are elected as representatives to the US House in 33 percent of the elections and are elected as mayors in Italy in 23 percent of the elections, females who marginally qualify for the second round are 7–9 percentage points less likely to win. In relative terms, female runners-up are 40 percent less likely to win the second round than male runners-up in the context of Italian municipal elections, while female primary winners are 20 percent less likely than male primary winners to be elected to the US House of Representatives.[21]

Table 1 shows that the gender difference in the probability of being elected cannot be explained by observable characteristics at the individual and election level. More specifically, in Column (2), I augment the specification in equation (8) with observable individual characteristics that include, for the US House of Representatives, party affiliation, an indicator equal to 1 if the candidate is the incumbent representative, and the number of previous terms in which the candidate competed for office or received the party nomination for the general election. For Italian municipalities, I control for age, education, occupation, an indicator equal to 1 if the candidate is the incumbent mayor, the number of previous terms in which the candidate competed for office, and party ideology indicators.

In Column (3), I control for determinants of electoral success in the primary election and the first round. In Column (4), I simultaneously add personal and election characteristics at the candidate level to the specification. Finally, in Column (5), I control for congressional district fixed effects in Panel (a). In Panel (b), I replace election-year effects with a set of election fixed effects, thus holding constant all the observable and unobservable characteristics of each race and each municipality. The results presented in Table 1 show that the estimated coefficients are not sensitive to the introduction of any of these sets of additional covariates, providing reassurance that Assumption I is likely satisfied in both contexts.[22]

The results presented in Figure 4 and Table 1 indicate that voters are less likely to vote for women who performed as well as their male counterparts during the first round in comparable competitive environments. These results rely on a variation that holds constant the individual decision to run for office and the party selection of the leader. Gender differences in the probability of winning an election are not explained by observable characteristics, including previous political experience and incumbency status.

[21]The baseline probability of winning the election for marginally qualified male candidates is well below 50 percent in the context of the US House of Representatives because incumbent representatives – who have a high probability of being reelected – are unlikely to win their primary election by a narrow margin. The low baseline probability of winning the election for marginally qualified male candidates in the context of runoff elections in Italian municipalities comes as a direct consequence of the fact that in the second round, the marginal candidate faces a competitor who received more votes in the first round.

[22]Note that the specification in Column (5) of Panel (b) exploits only within-election variation, hence identifying the coefficients β_1 and β_2 based on 281 elections in which both the runner-up and the top recipient are within the Calonico et al. (2014) optimal bandwidth.

Table 1: Qualification to final round and probability of winning the election, by gender

(a) US House of Representatives

	(1)	(2)	(3)	(4)	(5)
	Representative	Representative	Representative	Representative	Representative
Female × Qualified	-0.071*	-0.060*	-0.079**	-0.069**	-0.065*
	(0.036)	(0.035)	(0.036)	(0.034)	(0.037)
Qualified	0.347***	0.342***	0.350***	0.346***	0.347***
	(0.026)	(0.026)	(0.026)	(0.026)	(0.028)
Observations	3,403	3,403	3,403	3,403	3,403
R^2	0.211	0.256	0.224	0.274	0.360
Individual Char.		✓		✓	✓
Primary Returns			✓	✓	✓
District Fixed Effects					✓
Election Year Effects	✓	✓	✓	✓	✓
Mean Dep. Var.	0.130	0.130	0.130	0.130	0.130
Bandwidth	0.138	0.138	0.138	0.138	0.138

(b) Italian municipalities

	(1)	(2)	(3)	(4)	(5)
	Mayor	Mayor	Mayor	Mayor	Mayor
Female × Qualified	-0.097***	-0.094***	-0.090***	-0.089***	-0.107**
	(0.032)	(0.032)	(0.031)	(0.031)	(0.054)
Qualified	0.232***	0.231***	0.231***	0.230***	0.245***
	(0.021)	(0.021)	(0.020)	(0.020)	(0.027)
Observations	3,582	3,582	3,582	3,582	3,582
R^2	0.200	0.209	0.275	0.284	0.507
Individual Char.		✓		✓	✓
First round Returns			✓	✓	✓
Election Fixed Effects					✓
Election Year Effects	✓	✓	✓	✓	
Mean Dep. Var.	0.129	0.129	0.129	0.129	0.129
Bandwidth	0.115	0.115	0.115	0.115	0.115

Notes: Panel (a): The dependent variable is a dummy equal to 1 if the candidate is elected as member of the US House of Representatives during the general election. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Column (2) includes controls for *Incumbent, Candidacy counter, Nomination counter,* and *Democratic party.* Column (3) includes controls for *Vote share (Primary).* Column (4) includes the sets of controls in (2) and (3). Column (5) includes the set of controls in (4) and congressional district fixed effects. Panel (b): The dependent variable is a dummy equal to 1 if the candidate is elected as mayor during the second round. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable. Column (2) includes controls for *Age, College, High-skill job, Incumbent, Candidacy counter, Left-wing, Right-wing*, and *Independent.* Column (3) includes controls for *Vote share (First round), Vote share coalition (First round),* and *No. parties in coalition (First round).* Columns (4) and (5) jointly include the sets of controls in (2) and (3). Columns (1)–(5) include election year dummies, while Column (6) includes individual election fixed effects. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the congressional district-by-party level (Panel (a)) and municipality level (Panel (b)) are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

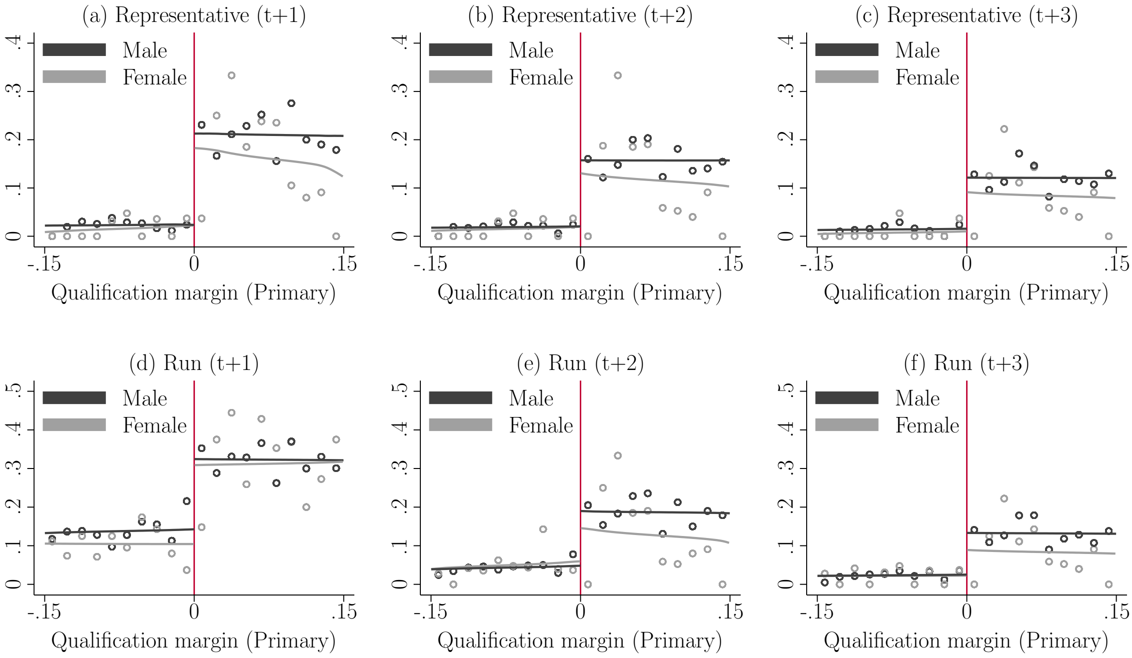

Figure 5: US Primary victory and future career outcomes, by gender

Notes: In Panel (a), the dependent variable is an indicator equal to 1 if the candidate is elected to the position of member of the US House of Representatives during the next election, held after two years. In Panels (b) and (c), the dependent variables are indicators equal to 1 if the candidate is elected as a member of the US House of Representatives during the elections held after four and six years, respectively. In Panels (d), (e), and (f), the dependent variable is an indicator equal to 1 if the candidate runs in the primary election for the position of member of the US House of Representatives during the elections held after two, four, and six years, respectively. In all panels, solid lines represent non-parametric smoothers of the qualification margin, separately estimated on either side of the qualification threshold and for male and female candidates. Purple lines and markers refer to female candidates, while green lines and markers refer to male candidates. Markers represent sample averages within bins of the running variable $h_{i,t}$ equal to 0.015.

In Section 5.3, I exploit data on campaign expenditures and revenues for the US House of Representatives and data on endorsements in runoff elections to explore whether the estimated gender gaps depend on differences in campaigning and negotiation behavior. To conclude, I investigate the possibility that voters are biased against women by exploiting heterogeneities in individual and election characteristics. In what follows, I show how the lower probability of winning an election impacts the whole career trajectories of female candidates, who are less likely than male candidates to be elected or promoted to higher levels of the hierarchy in the future.

5.2 Future career outcomes

Figure 5, Table B9 and Table B10 show that the returns from winning a primary election for the US House of Representatives are lower for female candidates than for male candidates for up to the third downstream election. Qualifying for the general election

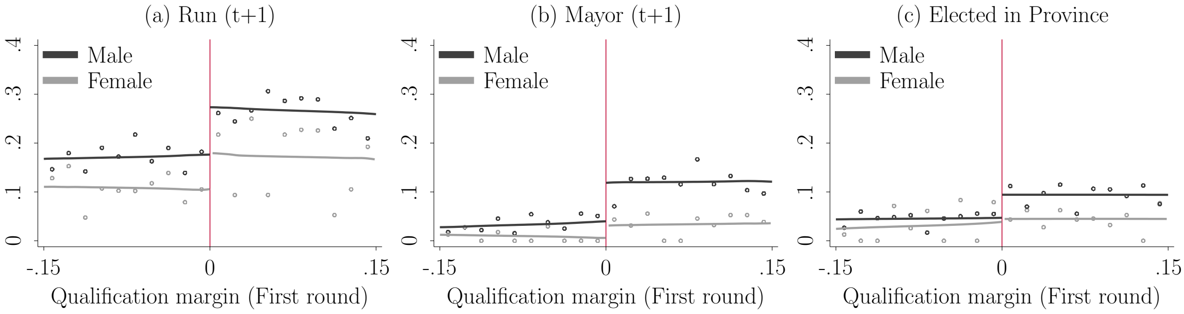

Figure 6: Qualification to second round and future career outcomes, by gender (Italy)

Notes: In Panel (a), the dependent variable is a dummy equal to 1 if the candidate runs for office in the occasion of elections held in term $t+1$. In Panel (b), the dependent variable is a dummy equal to 1 if the candidate is elected as mayor during elections held in term $t+1$. In Panel (c), the dependent variable is a dummy equal to 1 if the candidate is in the future elected as a member of the Provincial Council or to the position of President of the Province. In all panels, solid lines represent non-parametric smoothers of the qualification margin, separately estimated on either side of the qualification threshold and for male and female candidates. Light gray lines and markers refer to female candidates, while dark gray lines and markers refer to male candidates. Markers represent, for each gender, sample averages within bins of the running variable $h_{i,t}$ equal to 0.015.

increases the probability of becoming a member of the US House of Representatives in the election held two years later by 20 percentage points if the candidate is a male. The change in probability of winning the election is four percentage points lower if the candidate is a woman, although this gender difference is not statistically significant. The gender difference in returns is stable at four percentage points and statistically significant at the 10 percent level in subsequent terms.

Figure 5 and Tables B9 and B10 also show that the effect of qualifying for the general election on the probability of running for office in the next term is equal for male and female candidates, while fewer women decide to continue competing and receive their party nomination in subsequent terms. Taken together, these results suggest that the lower probability of winning the election for female candidates has a dual impact on future careers. In the short run, women are less likely to be elected even if the qualification does not increase the probability of running more for men. In subsequent terms, many women decide not to run, plausibly because they anticipate that they will face challenging competition.

I find similar and even stronger results in the context of runoff elections in Italy. Figure 6 and Table B11 show that during the next election, male runners-up are 10 percentage points more likely to compete for office and 6 percentage points more likely to be elected mayor than marginal losers. Male runners-up are also 3 percentage points more likely than marginal losers to be promoted to higher levels of the administrative hierarchy than male losers. In contrast, female candidates who marginally rank as the runner-up are more likely than female losers to compete for office in the future. However, they are not more likely than female losers to be elected during the next election or to be promoted.

***Sticky floor* or *leaky pipeline*?**

Future career trajectories could be worse for female candidates either because women are less likely to be elected in the first place or because female candidates who are elected have a lower incumbency advantage than men. These two alternatives reflect two common hypotheses to explain why women are less likely than men to reach the top in prestigious positions. On the one hand, the *sticky floor* hypothesis (e.g., Bjerk, 2008) posits that individuals belonging to an underrepresented group face early hurdles that affect the whole career trajectories. On the other hand, the *leaky pipeline* hypothesis (e.g., Thomsen and King, 2020) purports that individuals from an underrepresented group face more challenges than other individuals at every stage of their career trajectories. To distinguish between the two channels, I investigate whether there is evidence of gender differences in incumbency advantages by exploiting close elections in Italian municipalities with less than 15,000 inhabitants, subject to a single-ballot plurality system. The empirical strategy is akin to that of Wasserman (2018), who investigates whether an election defeat impacts the future career of male and female politicians differently in the context of local elections in California. The intuition behind this strategy is analogous to that of the empirical strategy presented in Section 4 of this paper. In this context, however, treated and control candidates are selected based on a comparison between the top vote recipient, who is elected mayor, and the runner-up.

Formally, I estimate a difference-in-discontinuities model of the form

$$Y_{i,t} = \gamma_0 + \gamma_1 F_i \times Mayor_{i,t} + \gamma_2 Mayor_{i,t} + \gamma_3 F_i + f(F_i; WinningMargin_{i,t}) + \lambda_t + u_{i,t}, \tag{11}$$

where $Mayor_{i,t}$ is an indicator equal to 1 if the candidate is marginally elected to the position of mayor (i.e., she has a positive winning margin) and $f(F_i; WinningMargin_{i,t})$ is the running variable, such that $WinningMargin_{i,t}$ is the vote share distance between the candidate and the minimal vote share required to win the election (e.g., Lee, 2008). Analogously to in Figure 6 and Table B11, the dependent variables are the probability of running for office during the next election, the probability of being elected to the position of mayor during the next election and the probability of being promoted to the position of member of the provincial council or president of the province in the future.[23]

[23] Consistent with the specification in equation (8), I estimate equation (11) nonparametrically in the figures, use local-linear regressions within the Calonico et al. (2014) optimal bandwidth in the tables, and draw inferential conclusions based on standard errors robust to clustering at the municipal level. See Figure F1, Table F1 and Figure F2 in the Appendix for the McCrary (2008) test for these running variables, as well as for balancing checks and tests of the sensitivity of the estimated coefficients to the bandwidth selector. See Table F2 in the Appendix for the results obtained when I use alternative definitions of the winning margin as defined in Appendix E.

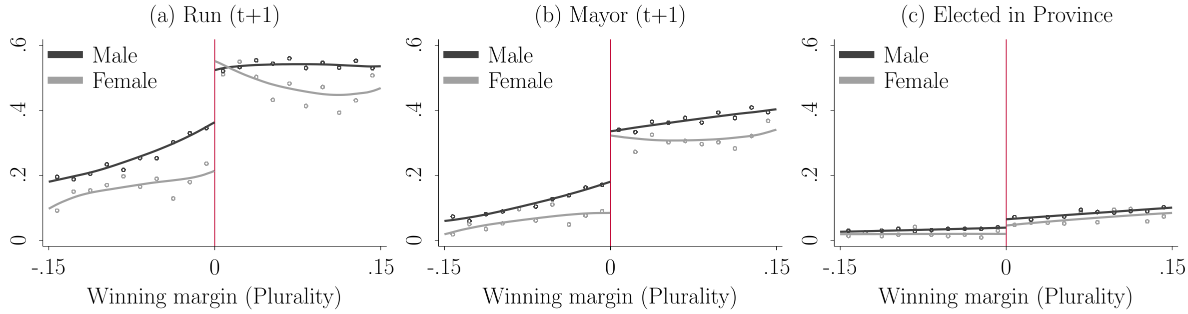

Figure 7: Electoral success and future career outcomes, by gender

Notes: All mayoral candidates in the municipalities belonging to the ordinary regions running for election between 1993 and 2019 in municipalities having less than 15,000 residents. Data are at the term level. In Panel (a), the dependent variable is a dummy equal to 1 if the candidate runs for office in the occasion of elections held in term $t+1$. In Panel (b), the dependent variable is a dummy equal to 1 if the candidate is elected as mayor during elections held in term $t+1$. In Panel (c), the dependent variable is a dummy equal to 1 if the candidate is in the future elected as a member of the Provincial Council or to the position of President of the Province. In all panels, solid lines represent non-parametric smoothers of the winning margin, separately estimated on either side of the victory threshold and for male and female candidates. Light gray lines and markers refer to female candidates, while dark gray lines and markers refer to male candidates. Markers represent, for each gender, sample averages within bins of the running variable $WinningMargin_{i,t}$ equal to 0.015.

A negative γ_1 would imply that incumbency advantages are lower for women than for men. Therefore, it would suggest that the results presented in Figure 6 and B11 are a combination of gender differences in the probability of becoming the new incumbent and gender differences in incumbency advantages. A nonnegative γ_1 would instead indicate that gender differences in future returns from participating in an election depend crucially on the challenges that women face in being elected in the first place, as predicted by the *sticky floor* hypothesis.

I present the results of this analysis in Figure 7 and in Table B12. First, consistent with the established evidence of incumbency advantages (e.g., Lee, 2008), marginal winners are more likely than marginal runners-up to run for office in the future as well as to be elected or promoted to higher levels of the administrative hierarchy. Marginally elected mayors are 16 percentage points more likely to be elected in the next election than marginal losers. Incumbency advantages in Italy are weaker than in other countries because of the presence of term limits, which prevent incumbent mayors from running for office after they have served two consecutive terms. Second, winning the election increases the probability of running for office more for women than it does for men.[24] Female marginal

[24] Brown et al. (2019) and Wasserman (2018) implement difference-in-discontinuities strategies to estimate gender differences in the returns to electoral success or defeat and find contrasting results. Brown et al. (2019) find that being elected as a member of a state legislature increases the probability of running for a seat in the US House more for men than for women. In contrast, Wasserman (2018) finds that losing an election decreases the probability of running for the same position in the future more for women than for men. My results are consistent with those of Wasserman (2018).

losers are less likely than male marginal losers to compete in the future, while male and female mayors are equally likely to compete in the next election. Third, election to the position of mayor increases the probability of being elected after the next election and of being promoted to the province level equally for marginally elected candidates of both genders. Women are less likely to be elected and promoted than men, but being elected does not differentially affect male and female candidates.

These results indicate that women who reach elected office and become the new incumbent have career developments comparable to those of male mayors. Therefore, the long-term returns from qualifying to the final round of a two-stage election are lower for women than for men arguably because female candidates do not receive support from voters in the final round. Consistent with the *sticky floor* hypothesis, women are therefore less likely than men to become incumbents early on in their careers, and in turn, this affects the chances of success in subsequent elections.[25]

5.3 Why are women less likely to win an election?

There are many potential explanations for why women are less successful in elections. First, male and female candidates may differ in their attitudes towards competition and negotiation (e.g., Bertrand, 2011). Gender differences in attitudes towards competition and negotiation may result in difficulties in raising money to conduct a campaign or in gathering sufficiently broad support from parties. Second, some voters may be biased against female candidates (e.g., Gonzalez-Eiras and Sanz, 2018; Le Barbanchon and Sauvagnat, 2018).

Other explanations for gender gaps established in the literature do not appear to be a key driver of my findings. Since candidates are quasirandomly assigned to participate in the election and compete in comparable competitions, sorting into the profession cannot explain my results. Moreover, gender differences in childrearing responsibilities (e.g., Azmat and Ferrer, 2017; Kleven et al., 2019) are unlikely to explain my results, at least in the context of municipal elections in Italy, where I observe the age of each candidate. First, politicians are relatively old (with a median age of 49 in my sample). Second, the youngest individuals in the sample do not drive the estimated gender gap (see Figure A11).

Voter turnout and local mass media

In Table 2, I present additional evidence in line with both the possibility that voters discriminate against female candidates and the possibility that gender gaps are driven by

[25]These findings are consistent with those of Hall (2015) and Fowler and Hall (2017), who document in the contexts of the US House, US Senate and US state legislatures that parties that nominate an extremist candidate in a close election suffer a backlash that prevents them from winning the seat for at least a decade after treatment because of incumbency advantages.

Table 2: Impact of qualification on voter turnout and newspaper reports, by gender

	(1)	(2)	(3)	(4)
	Turnout	Male turnout	Female turnout	Log(Newspaper articles)
		Δ (Second round - First round)		
Female × Qualified	-0.010**	-0.019**	0.002	-0.256*
	(0.005)	(0.008)	(0.008)	(0.132)
Qualified	-0.145***	-0.148***	-0.140***	
	(0.004)	(0.006)	(0.005)	
Observations	2,905	2,973	3,316	508
R^2	0.817	0.646	0.645	0.555
Election Year Effects	✓	✓	✓	✓
Mean Dep. Var.	-0.0652	-0.0614	-0.0595	2.581
Bandwidth	0.131	0.152	0.171	0.139

Notes: In Column (1), the dependent variable is the change in voter turnout between the first and the second round of a runoff election; in Column (2), the dependent variable is the change in voter turnout among male voters between the first and the second round of a runoff election; in Column (3), the dependent variable is the change in voter turnout among female voters between the first and the second round of a runoff election; in Column (4), the dependent variable is the natural logarithm of the difference between the number of articles on local newspaper that mention the candidate before the second round and the number of articles that mention the candidate before the first round. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

differences in campaigning and negotiation behavior. Table 2 shows how voter turnout and local media coverage of candidates varies with the gender of the marginally qualified candidates in runoff elections in Italian municipalities.[26] In Columns (1)-(3), the dependent variable is the difference between voter turnout in the first round and voter turnout in the second round, computed for the whole population of voters and separately for men and women.

Voter turnout is approximately 14 percentage points lower in the final round than in the first round if the marginally admitted runner-up is a man.[27] When, instead, the marginal candidate is a woman, turnout drops by one additional percentage point. The results presented in Columns (2) and (3) show that male voters drive the observed variation in turnout. Male participation in the final round drops by 17 percentage points – instead of 15 percentage points – when a woman competes. However, the participation of female voters in the final round does not vary with the gender of the marginal candidate. The estimates indicate that voters are less inclined to go to the polling station if a female

[26] Media coverage data at the individual candidate level are available only from 2007. See Appendix D for additional details.

[27] This result, consistent with that of Bordignon et al. (2016), represents a specificity of the Italian context. Fiva and Smith (2017) shows that turnout in Norway increased from 49 percent in the first round to 60 percent in the second round between 1909 and 1918.

candidate competes in the election. Reduction in voter turnout could be consistent with both gender differences in campaigning strength and with voter discrimination against female candidates. First, if female candidates do not campaign as fiercely as men, fewer voters are likely to vote. Second, if some voters are biased against women, those who are willing to support neither the female candidate nor the opposing party may decide not to vote.

Column (4) of Table 2 investigates whether the number of newspaper articles reporting on a candidate varies as a function of the gender of the candidate. The estimated coefficient for $Female \times Qualified$ is rather imprecise because these data are available only from 2007. Nevertheless, I find that at the threshold, female candidates are mentioned in local newspapers 25 percent less often than male candidates. On the one hand, voter bias against women impacts media outlets if demand for news on male candidates dominates demand for news on female candidates. In turn, market incentives could induce mass media to strategically report more news about male candidates than female candidates.[28] On the other hand, differences between the number of newspaper articles reporting about male and female candidates may reflect gender differences in campaigning behavior.

In what follows, I take advantage of rich data from each of the two settings to carefully separate these two potential mechanisms. First, to address whether male and female candidates differ in their campaigning and negotiation behavior, I exploit information on the funding that candidates invest and raise in advance of primary and general elections for the US House of Representatives. Second, I employ information on the number of endorsements that candidates receive between the first and second rounds of elections in Italian municipalities to investigate whether male candidates conduct more efficient negotiations. Lastly, I exploit several dimensions of heterogeneity with respect to candidate and district characteristics to detect whether the results are consistent with the conjecture that voters are biased against women.

5.3.1 Gender differences in campaigning and negotiation

Gender differences in campaigning and negotiation behavior may prevent women from being successful in the final round of two-stage elections. For instance, Bagues et al. (2020) find that parties with a plurality of seats in Spanish municipal councils are less likely to form a majority coalition when the leader is a woman. Biasi and Sarsons (2020) estimate that the introduction of individual bargaining over wages among school teachers widened gender wage differences.[29]

Each of the two contexts on which the empirical analysis builds offers unique variation

[28] Gentzkow and Shapiro (2006) and Gentzkow et al. (2018) show that market incentives induce media to adapt their reporting if individuals prefer to read news that confirms their priors.

[29] Gender differences in psychological attitudes towards bargaining and competition have been estimated in a number of lab experiments. See the handbook chapter Bertrand (2011) for a literature review.

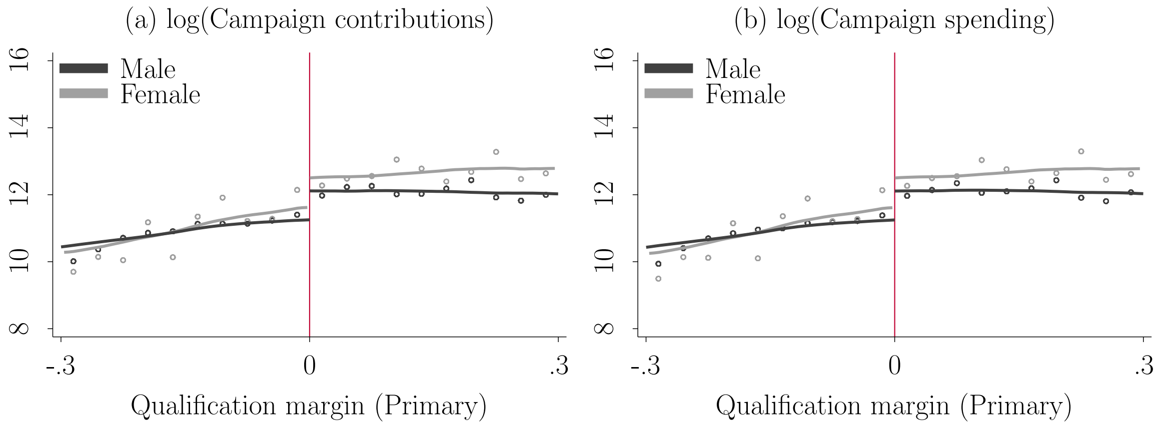

Figure 8: US Primary victory and campaign financing, by gender

Notes: In Panel (a), the dependent variable is the natural logarithm of the campaign contributions received by the candidate; in Panel (b), the dependent variable is the natural logarithm of the amount of resources invested by the candidate to finance her campaign. In both panels, the solid lines represent non-parametric smoothers of the qualification margin, separately estimated on either side of the qualification threshold and for male and female candidates. Light gray lines and markers refer to female candidates, while dark gray lines and markers refer to male candidates. Markers represent, for each gender, sample averages within bins of the running variable $h_{i,t}$ equal to 0.03.

to explore this channel thoroughly. More specifically, I observe the total amount and distribution of financial resources that each candidate raises and spends during his or her electoral campaign for the US House of Representatives both before the party primary and in between the two election stages. Second, I observe the number of endorsements received after the first round and the vote share of each party endorsing a candidate admitted to the second round of a runoff election in Italy. Suppose gender differences in attitudes towards negotiation or campaigning are a driver of the result. Then, we should expect female candidates to be endorsed by fewer parties than male competitors and to raise or invest a lower amount of funding in their campaign.

Male and female candidates raise and invest equal amounts of money

Starting with the US House of Representatives, Figure 8 and Table B13 offer no evidence of gender differences in the total amount of financial resources that a candidate raises and spends either before the primary election or between the two stages. The precision of the estimates is rather low because this information is available only for one-third of the candidates in the sample. Nevertheless, the point estimates suggest that gender differences in campaigning activity are not a key driver of the results. Figure A12 and Table B13 also show that the composition of donors does not vary significantly across genders. If anything, female candidates invest a lower amount of their own resources than male candidates and compensate by borrowing additional resources and by receiving

Table 3: Impact of qualification on coalition size and endorsements, by gender

	(1)	(2)	(3)	(4)
	No. parties in coalition	Vote share coalition	No. endorsements	Vote Share of endorser parties
Female × Qualified	0.091	0.013	-0.084	0.001
	(0.234)	(0.012)	(0.103)	(0.007)
Qualified	0.539***	0.100***	0.614***	0.097***
	(0.136)	(0.007)	(0.088)	(0.006)
Observations	2,188	2,106	1,682	2,396
R^2	0.312	0.541	0.127	0.381
Election Year Effects	✓	✓	✓	✓
Mean Dep. Var.	2.896	0.242	0.209	0.0404
Bandwidth	0.0941	0.0910	0.0735	0.105

Notes: In Column (1), the dependent variable is the total number of parties in support of the candidate during the second round; in Column (2), the dependent variable is the total vote share for the election of the municipality council of all parties who support the candidate during the second round; in Column (3), the dependent variable is the number of endorsements received by each candidate between the first and the second round; in Column (4), the dependent variable is the vote share received by endorsing parties in the election of the municipality council. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

slightly more contributions from the political party and individual donors.

Male and female candidates are equally likely to be endorsed by other parties

Turning next to Italian municipalities, Table 3 shows that male and female mayoral candidates receive the same number of endorsements before the second round of the runoff competition and, more generally, are supported by a coalition formed by the same number of parties. Parties endorsing male and female candidates are equal in strength, as measured by the share of total votes received in the election for the municipal council. In Columns (1) and (3) of Table 3, the dependent variables are the number of parties in the coalition supporting the candidate during the second round and the number of parties endorsing the candidate between the two rounds, respectively. The results show that qualifying for the final round leads to an increase in coalition size equal to 0.5–0.6 parties. However, the coefficients for the gender difference are not significantly different from zero and are small in magnitude in both columns. In Columns (2) and (4), the dependent variables are the sum of the vote shares received by all parties supporting the candidate in the second round and the vote share received by parties who endorse the candidate between the two rounds. I estimate a positive and nonsignificant coefficient for $Female \times Qualified$, excluding the possibility that parties endorsing female candidates are less representative than parties supporting male candidates.

Table 4: Heterogeneity by district characteristics (Italian municipalities)

	Female mayor in the past		Local GDP growth		Difference in LFP	
	Yes	No	Above median	Below median	Above median	Below median
	(1)	(2)	(3)	(4)	(5)	(6)
	Mayor	Mayor	Mayor	Mayor	Mayor	Mayor
Female × Qualified	0.162	-0.141***	-0.155***	-0.011	-0.175***	-0.051
	(0.100)	(0.033)	(0.050)	(0.069)	(0.057)	(0.041)
Qualified	0.142**	0.242***	0.261***	0.200***	0.247***	0.226***
	(0.067)	(0.023)	(0.038)	(0.040)	(0.032)	(0.028)
Observations	430	3,130	1,021	1,036	1,372	2,188
R^2	0.346	0.219	0.241	0.212	0.250	0.215
Election Year Effects	✓	✓	✓	✓	✓	✓
Province Fixed Effects	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.144	0.126	0.127	0.112	0.141	0.119
Bandwidth	0.115	0.115	0.115	0.115	0.115	0.115

Notes: The dependent variable is a dummy equal to 1 if the candidate is elected as mayor during the second round. In Column (1), the sample is restricted to elections held in municipalities in which at least one woman had been previously elected to the position of mayor, while in Column (2), the sample is restricted to elections held in municipalities in which no women had previously served as the mayor. In Column (3), the sample is restricted to municipalities in which GDP growth during the election year is above the median GDP growth of the province, while in Column (4), the sample is restricted to municipalities in which the GDP growth is below the median GDP growth of the province. In Column (5), the sample is restricted to municipalities in which the difference between the labor force participation of men and women is above the median difference of the province, while in Column (6), the sample is restricted to municipalities in which the difference between the labor force participation of men and women is below the median difference of the province. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable, election year dummies and province fixed effects. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

5.3.2 Voter bias against women

Voter bias may result in a lower probability of female candidates being elected both in the case of taste-based discrimination (Becker, 1957) and in the case of statistical discrimination (e.g., Arrow et al., 1973; Phelps, 1972). In the absence of a sharp test for discrimination, which would require analyzing the gender of the candidate independently of any other characteristics, I estimate the main results in subsamples of localities and candidates differing in a number of predetermined characteristics. The aim is, on the one hand, to detect patterns consistent with the conjecture that voters discriminate against female candidates. On the other hand, I aim to provide suggestive evidence on whether the bias depends on statistical or taste-based discrimination.

Heterogeneity by district characteristics

In Table 4, I show that gender differences in the probability of winning a seat are heterogeneous in terms of a number of characteristics of the local electoral district.[30] In Columns (1) and (2) of Table 4, I compare the results in Italian municipalities where at least one woman previously served as a mayor and municipalities in which only men were elected mayor in past elections, consistent with Gonzalez-Eiras and Sanz (2018).[31] In Columns (3) and (4), I compare municipalities in which GDP growth in the election year is above the median GDP growth of the province to which the municipality belongs and municipalities in which GDP growth is below its provincial median.[32] Finally, in Columns (5) and (6), I compare municipalities in which the difference between male and female labor force participation is above or below the provincial median.[33,34] If voters are biased against women, we should expect stronger gaps in places where voters are more likely to discriminate (i.e., places in which voters have not supported female candidates in the past or places in which women face hurdles in the local labor market). Analogously, we should observe a reduced gap if voters seek a change in leadership because of poor economic performance.[35]

The results in Table 4 are strongly consistent with the conjecture that female candidates are less likely to be elected because some voters discriminate against women. First, the gap in the probability of being elected is significantly larger in municipalities where no women previously held the same political office. There are at least two potential explanations for these findings. On the one hand, exposure to a female office holder may increase voter support for women (e.g., Bhavnani, 2009; Shair-Rosenfield, 2012). On the

[30] I limit this analysis to runoff elections in Italian municipalities for data availability reasons. Congressional districts for the US House of Representatives do not correspond one to one with counties at which data on labor market participation are available. Moreover, redistricting and the unavailability of electoral returns before 1976 do not allow me to assess whether other women represented the same district in the past. In Italy, instead, I rely on a stable number of administrative units for which labor market data are available at the municipal level and that have been observed ever since the introduction of direct election of mayors in 1993.

[31] Bhavnani (2009) and Shair-Rosenfield (2012) find that observable characteristics of elected politicians generate spillover effects. Localities with a woman in power are more likely to vote for another woman in the future.

[32] GDP data at the municipality level are released yearly from the Italian Ministry for the economy and finance by aggregating individual tax returns. Data are available from year 2000.

[33] Data on labor force participation at the municipality level are from the 1991, 2001, and 2011 population censuses.

[34] In all specifications, I augment equation (8) with province dummies to limit the comparison to municipalities belonging to the same province. This strategy, unlike a simple comparison of municipalities across the country, allows me to draw conclusions not driven by heterogeneities between regions in Northern and Southern Italy. Figure A13 in the Appendix reports the gender gap based on estimating equation (8) separately for each region. We can see a systematically larger gender gap in regions in the South, apart from Liguria, than in the North. The smallest gap, often indistinguishable from zero, is estimated for the regions of Central Italy, where the left parties are particularly strong (e.g., Fontana et al., 2017).

[35] See, e.g., Markus (1988) and Bagues and Esteve-Volart (2016) for empirical evidence on the association between local economic conditions and support for incumbents seeking reelection.

other hand, the results may reflect that voters in different localities have different gender preferences. Second, women are more likely to be elected mayor in municipalities experiencing relatively low GDP growth than in municipalities in which GDP is growing at a higher rate. These results indicate that voters are more willing to support women – who are arguably outsiders – in bad times, while voters are likely to confirm male-dominated leadership in good times. Third, stronger gender differences are found in municipalities where women are less integrated into the local labor market, as measured by gender differences in labor force participation.

Overall, the results in Table 4 show that women and men are equally likely to win elections only under specific circumstances that make voters more inclined to support a candidate from the underrepresented gender. In particular, a comparison of the estimates in Columns (3) and (4) indicates statistical discrimination as a key channel. Seemingly, the gender preferences of voters are attenuated when voters receive signals on the poor productivity of the current political elite. In Table B14, I propose an additional piece of evidence in support of this conjecture. More specifically, Table B14 shows that in elections for both members of the US House of Representatives and for mayors in Italian municipalities, the results are stronger in the earlier years of the sample. This evidence suggests that exposure to increasingly more women in the political arena may reduce voter bias against female candidates.

Gender gaps are driven by inexperienced candidates

In Table 5, I compare gender difference in the probability of winning an election among candidates who never competed for the same position in the past and the estimates for the subsample of candidates who previously ran for office. The results show that female candidates have a lower probability of winning an election than their male colleagues only the first time that they compete for office. Among first-time runners, women are 9 percentage points less likely than men to be elected to the US House of Representatives and more than 10 percentage points less likely to be elected mayor in Italian municipalities. On the other hand, among candidates with past political experience at the same level, the estimated coefficients for $Female \times Qualified$ are not significantly different from 0, and the point estimates are negligible.

These results should be interpreted with caution since they rely on samples selected based on an outcome. However, the coefficients further confirm that gender gaps in the probability of winning an election estimated in this paper are consistent with models of statistical discrimination. Statistical discrimination predicts that voters use gender as a signal for future productivity or a policy platform when information about an individual candidate is limited. When, instead, candidates run for a second time, voters have had the opportunity to observe more precise indicators of productivity or policy preferences during a term in office or leading the opposition. Therefore, the gender of the candidate

Table 5: Heterogeneity by previous experience of the candidate

	(a) US House of Representatives First experience		(b) Italian municipalities First experience	
	Yes	No	Yes	No
	(1)	(2)	(1)	(2)
	Representative	Representative	Mayor	Mayor
Female × Qualified	-0.091**	-0.014	-0.103***	-0.001
	(0.040)	(0.076)	(0.032)	(0.106)
Qualified	0.342***	0.354***	0.235***	0.227***
	(0.029)	(0.053)	(0.023)	(0.048)
Observations	2,525	879	2,850	710
R^2	0.210	0.228	0.205	0.235
Election Year Effects	✓	✓	✓	✓
Mean Dep. Var.	0.122	0.150	0.125	0.141
Bandwidth	0.138	0.138	0.115	0.115

Notes: Panel (a): The dependent variable is a dummy equal to 1 if the candidate is elected as member of Congress in the general election. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Panel (b): The dependent variable is a dummy equal to 1 if the candidate is elected as mayor during the second round. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. In both panels, in Column (1), the sample is restricted to candidates who are running for office for the first time while in Column (2), the sample is restricted to candidates who had previously run for the same office. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the congressional district-by-party level (Panel (a)) and municipality level (Panel (b)) are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

no longer contributes to the formation of voter beliefs.

In Table B15, I corroborate these results by showing that gender differences in the probability of winning the election are i) marginally stronger among Republican candidates than among Democratic candidates and ii) stronger among independent candidates than among candidates with a party affiliation. These results suggest that sharing a party affiliation with a female candidate induces more voters to support her. In contrast, voters take gender into account more when evaluating independent candidates.[36]

[36]Figures A14 and A15 document evidence of heterogeneity across the educational level and the previous occupation of political candidates. More specifically, I find that women are less likely to be elected mayor than men among both candidates having less than a high school education and college degree holders. In contrast, I do not find any evidence of gender differences in the probability of being elected among candidates holding a high school degree. Moreover, gender differences in the probability of winning an election are larger among lawyers, entrepreneurs, school teachers and university professors – which are common professions of successful career politicians – than among candidates with different backgrounds, such as clerks and medical doctors. These results further document that gender gaps are heterogeneous with respect to a number of individual characteristics. However, the gaps are present among a relatively large pool of candidates.

5.4 Robustness checks

As already indicated, the McCrary (2008) test supports that candidates are not able to manipulate the running variable around the qualification threshold (see Figure A5), and several predetermined covariates are balanced on the two sides of the threshold (see Tables B3 and B4). I perform a battery of additional robustness checks to validate the main results and to ensure that crossing the qualification threshold affects the career of politicians only through the mechanism of having qualified for the final round.

First, in Figures A16–A19, I replicate the main results using several bandwidths to fit local-linear regressions, and in Tables B16 and B17, I replicate the main results using quadratic and cubic controls of $f(F_i; h_{i,t})$. The estimated coefficients are very stable across bandwidths and polynomial fits, and the coefficients for $F_i \times Qualified_{i,t}$ are statistically significant even for relatively large bandwidths.[37]

Second, Tables B20 and B21 replicate the main results in the subsample of elections in which at least one male and one female candidate compete. The results presented in Tables B20 and B21 are in line with the findings obtained by exploiting all elections and are stronger for the US House of Representatives. The results ensure that gender differences in the probability of being elected measure voter support for given candidates in the final round and not unobserved determinants of selection. Moreover, in Table B22, I replicate the results in Table 1 by replacing the $F_i \times Qualified_{i,t}$ dummy with an indicator equal to 1 if the closest candidate to the threshold who does not qualify for the final round is a woman. The gender of excluded candidates should not impact the probability of winning an election for candidates who qualify if men and women compete in comparable elections. Indeed, the results presented in Table B22 confirm that the gender of excluded candidates does not have any impact on the probability of being elected.

Third, to further ensure that Assumption I is satisfied, I report the results from estimating an alternative regression-discontinuity specification. Specifically, I restrict the focus to elections in which only one woman runs and compare elections in which the female candidate qualifies by a narrow margin and elections in which the woman is excluded from the final round on a close margin. I estimate

$$\mathbf{1}(Elected = MargQualified)_{e,t} = \kappa \mathbf{1}(Female = MargQualified)_{e,t} + f(h^F_{e,t}) + \iota_t + \varepsilon_{e,t}, \quad (12)$$

where $\mathbf{1}(Female = MargQualified)_{e,t}$ is an indicator equal to 1 if the female candidate qualifies for the second round and ι_t is a set of election-year fixed effects. The dependent

[37] Tables B18 and B19 in the Appendix show that the results do not change when the qualification margin is defined according to alternative approaches, when the sample is restricted to include only the runner-up and the third largest vote recipients, or when each election has been reweighted by the inverse of the number of candidates.

variable $\mathbf{1}(Elected = MargQualified)_{e,t}$ is an indicator equal to 1 if the marginally qualified candidate is elected to the office. $f(h^F_{e,t})$ is the running variable, defined as the vote share distance between the female candidate and the male candidate closest to the qualification threshold on the opposite side of it.[38]

The results of this check are reported in Table B23. The precision is substantially reduced as the number of observations drops by 90 percent. However, the estimated coefficients for κ are in line with those in Table 1 and statistically significant at the 10 percent level in the richer specifications.[39]

Lastly, Anagol and Fujiwara (2016) find for Brazil, India, and Canada that candidates who marginally rank as runners-up under the plurality system have better future career development than candidates who rank by a narrow margin as the third-largest vote recipients. Similarly, Folke et al. (2016) find that when voters are allowed to rank candidates belonging to the same party, candidates who rank marginally better than their colleagues have better career development. Neither the Anagol and Fujiwara (2016) runner-up effect nor the Folke et al. (2016) primary effect effect depend on differences in the probability of winning the election, which is zero on either side of the threshold separating the two candidates in the former case and one for both candidates in the latter case.

To ensure that gender differences in the runner-up effect or the primary effect do not threaten the validity of my estimates, I estimate equation (8) in cases in which a final round of voting does not take place. In my sample, I observe Italian municipalities subject to the plurality system and municipalities under the runoff system where the top candidate received more than 50 percent of the votes in the first round. In both cases, the runner-up and the third-largest vote recipient have a probability of being elected equal to zero, and voters do not impact the career path of candidates on either side of the threshold. The results of this check are presented in Table B24. Table B24 shows that the career path of candidates does not vary by gender around the placebo thresholds. Interestingly, the Anagol and Fujiwara (2016) runner-up effect does not seem to be present in Italy. Candidates who marginally rank as the runner-up under the plurality system are 2 percentage points more likely to compete for the position of mayor in the future, but they are not more likely to be either elected or promoted to the province administration

[38]This approach has the advantage of holding constant all characteristics that vary at the election and district level but comes with limitations that make it not the first-best option for this empirical analysis. First, it does not allow me to rely on individual-level outcomes since the identifying variation comes at a more aggregate level. Second, it cannot be used to study long-run outcomes, which also depend on gender differences in future selection, as introduced in equation (6). Third, it reduces the number of observations by approximately 90 percent from the number used in the main analysis.

[39]See Figure F3 in the Appendix for a graphical representation of the relationship. See Figure F4 in the Appendix for the McCrary (2008) test showing no evidence of manipulation of this running variable as well as Tables F3 and F4 and Figure F5 for evidence on the balancing of municipal-level characteristics and for the sensitivity of the estimates to the chosen bandwidth, respectively.

in the future.[40]

6 Discussion

The empirical results presented in Section 5 provide new evidence on why so few women reach top positions in the political profession. To start, female candidates have a significantly lower probability of winning an election than their male colleagues at the beginning of their careers. As a result, women are less likely to become the new incumbent and to climb the political hierarchy in the future. Failing to enter the profession the first time a candidate attempts to do so has a direct impact on future success and promotions.

The empirical evidence documented in this paper also suggests an additional channel making female politicians more likely than men to leave the political profession during their career. Research has thus far focused mainly on supply-side determinants such as gender differences in political ambition (e.g., Fox and Lawless, 2004), opportunity costs (e.g., Folke and Rickne, 2020), and within-family allocation of childrearing responsibility (e.g., Kleven et al., 2019). My results suggest that women are more likely to desist from running for office when they recognize that voters are less likely to support them than male competitors. My results on the US House of Representatives are informative along this line. Women are less likely to be elected than men, but they are not less likely to compete in the following election. However, as the gender difference in the probability of winning the election propagates to the following term, fewer women decide to run in subsequent elections.

My use of two empirical settings allows me to strengthen the external validity of the estimates given the widely differing institutions and to explore a broad range of mechanisms. First and foremost, candidates who only just manage to qualify for the second round of a runoff election have a systematically lower probability of being elected than their opponent who received a larger share of total votes during the first round. In the case of party primaries for the US House of Representatives, instead, the marginal winner of a party primary is not necessarily weaker than her opponent who won the other party's primary.

Second, individual characteristics arguably impact elections for members of the US House of Representatives to a lesser extent than mayoral elections in Italy. In the former, the winner becomes a member of a legislative assembly at the national level, and voters might trade off their preferences over individual candidates and parties. In contrast, national considerations are likely to be less important in local elections for mayor. Third, the group of voters participating in the first round differs substantially between the

[40] See Figure F6, Table F5, and Figure F7 in the Appendix for the McCrary (2008) test for these running variables as well as for balancing checks and for tests of the sensitivity of the estimated coefficients to the bandwidth selector.

two settings. Voters participating in party primaries are a small and selected group of individuals with firm policy or ideological preferences in support of the party, while the majority of voters only participate in the general election. In runoff elections in Italy, voter turnout is higher in the first round than in the second since a subset of voters who supported excluded candidates usually do not participate in the final round.

Fourth, the campaign ahead of the final round differs in length between the two systems. While general elections for the US House take place several months after the primary, the final round in runoff elections takes place only two weeks after the first round. Fifth, losing candidates are free to endorse any of the two remaining challengers in Italian municipal elections. Conversely, excluded candidates are expected to support their party nominees in the case of party primaries. Lastly, incumbent members of the US House of Representatives are always allowed to compete for an additional term, while mayors in Italian municipalities are limited to two consecutive terms in office.

Gender gaps in the probability of winning an election and in future career returns from participating in an electoral competition are found in both countries and settings considered in the analysis. However, the magnitude of the divergence differs substantially. In relative terms, women have a 20 percent lower probability than men of winning election to the US House of Representatives, while female candidates are 40 percent less likely to be elected mayor in Italian municipalities. A cross-country comparison of the estimates cannot conclude that voters in the United States are more biased against women than voters in Italy since the two empirical settings differ along several dimensions.

7 Concluding remarks

Several prestigious professions are male-dominated environments. Women are less likely than men to enter those professions and to be promoted to the top positions. The key to understanding what prevents women from being successful in such environments is to identify gender gaps in returns, taking into account that selection into such professions varies by gender. In this paper, I exploit electoral institutions featuring two-stage elections, where candidates are deterministically admitted to the final round based on first-round returns. This institutional rule allows me to compare the probability of electoral success and future career outcomes of candidates around the qualification threshold, who are neither self-selected nor appointed by the party, and to exploit marginally excluded candidates as an ideal control group.

I find that female primary winners are 7 percentage points less likely than male primary winners to be elected to the US House of Representatives. Analogously, female runners-up in Italian municipalities are 9 percentage points less likely than male runners-up to be elected to the position of mayor in the second round. The baseline probabilities that a male candidate at the threshold is elected are 33 and 23 percent in the US and

Italian settings, respectively, indicating that female candidates are heavily penalized by voters, as their chances of election are 20–40 percent lower than those of their male counterparts. The gender gap in the probability of being elected affects the future career trajectories of female candidates. Accounting for the career trajectories of men and women who marginally failed to qualify for the final round, I find that female candidates are less likely to be elected and promoted to higher levels of the administrative hierarchy in the future.

Gender differences in negotiation or campaigning cannot explain my findings. Male and female candidates attract the same number of endorsements between the two rounds and invest the same amount of funding in their campaign. Additionally, I exclude the possibility that future career outcomes are lower for women because of gender differences in incumbency advantages. Using data from Italy, statistical discrimination appears to be the key mechanism behind the empirical results since the subpopulation of inexperienced candidates drives the gap. Moreover, female candidates are penalized more in areas of the country in which women are not integrated into the labor market and politics.

My results contribute to the vast literature estimating gender gaps in the economy and its main drivers. Not accounting carefully for sample selection bias may lead to misleading conclusions, as no gaps in outcomes might still hide gaps in selection or returns if these sources of bias are not taken into account. In particular, my results advise reconsidering earlier evidence showing that men and women who compete for a seat in the US House have a comparable probability of being elected (see, e.g., Thomsen and King, 2020, for a literature review). Additionally, my results indicate that the worse career outcomes of female politicians than of male politicians are significantly affected by the challenges faced by women early on in their political careers.

References

Albrecht, J., Björklund, A., and Vroman, S. (2003). Is There a Glass Ceiling in Sweden? *Journal of Labor Economics*, 21(1):145–177.

Alesina, A. and Giuliano, P. (2011). Preferences for Redistribution. In *Handbook of Social Economics*, volume 1, pages 93–131. Elsevier.

Anagol, S. and Fujiwara, T. (2016). The Runner-up Effect. *Journal of Political Economy*, 124(4):927–991.

Arrow, K. et al. (1973). The Theory of Discrimination. *Discrimination in Labor Markets*, 3(10):3–33.

Azmat, G. and Ferrer, R. (2017). Gender Gaps in Performance: Evidence from Young Lawyers. *Journal of Political Economy*, 125(5):1306–1355.

Bagues, M. and Campa, P. (2020). Can Gender Quotas in Candidate Lists Empower Women? Evidence from a Regression Discontinuity Design. Forthcoming at *Journal of Public Economics*.

Bagues, M., Campa, P., and Palguta, J. (2020). Gender and Political Coalitions. Working Paper.

Bagues, M. and Esteve-Volart, B. (2016). Politicians' Luck of the Draw: Evidence from the Spanish Christmas Lottery. *Journal of Political Economy*, 124(5):1269–1294.

Barth, E., Kerr, S. P., and Olivetti, C. (2017). The Dynamics of Gender Earnings Differentials: Evidence from Establishment Data. NBER Working Paper No. 23381.

Baskaran, T. and Hessami, Z. (2020). A Leaky Pipeline: Recontest Rates and the Underrepresentation of Women in Politics. Working Paper.

Becker, G. S. (1957). *The Economics of Discrimination*. University of Chicago press.

Bell, C. (2020). The Rulers, Elections, and Irregular Governance Dataset (REIGN). https://oefresearch.org.

Bertrand, M. (2011). New Perspectives on Gender. In *Handbook of Labor Economics*, volume 4, pages 1543–1590. Elsevier.

Besley, T. and Coate, S. (1997). An Economic Model of Representative Democracy. *The Quarterly Journal of Economics*, 112(1):85–114.

Bhavnani, R. R. (2009). Do Electoral Quotas Work After They are Withdrawn? Evidence from a Natural Experiment in India. *American Political Science Review*, 103(1):23–35.

Biasi, B. and Sarsons, H. (2020). Flexible Wages, Bargaining, and the Gender Gap. Working Paper.

Bjerk, D. (2008). Glass Ceilings or Sticky Floors? Statistical Discrimination in a Dynamic Model of Hiring and Promotion. *The Economic Journal*, 118(530):961–982.

Bordignon, M., Nannicini, T., and Tabellini, G. (2016). Moderating Political Extremism: Single Round versus Runoff Elections under Plurality Rule. *American Economic Review*, 106(8):2349–2370.

Brollo, F. and Troiano, U. (2016). What Happens when a Woman Wins an Election? Evidence from Close Races in Brazil. *Journal of Development Economics*, 122:28–45.

Brown, R., Mansour, H., O'Connell, S. D., and Reeves, J. (2019). Gender Differences in Political Career Progression: Evidence from US Elections. CESifo Working Paper No. 7821.

Buser, T., Niederle, M., and Oosterbeek, H. (2014). Gender, Competitiveness, and Career Choices. *The Quarterly Journal of Economics*, 129(3):1409–1447.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust Nonparametric Confidence Intervas For Regression-Discontinuity Designs. *Econometrica*, 82(6):2295–2326.

Card, D., Cardoso, A. R., and Kline, P. (2015). Bargaining, Sorting, and the Gender Wage Gap: Quantifying the Impact of Firms on the Relative Pay of Women. *The Quarterly Journal of Economics*, 131(2):633–686.

Carozzi, F., Cipullo, D., and Repetto, L. (2020). Political Fragmentation and Government Stability. Evidence from Local Governments in Spain. Working Paper.

Carozzi, F. and Gago, A. (2017). Do Female Leaders Promote Gender-Sensitive Policies? Working Paper.

Casarico, A., Lattanzio, S., and Profeta, P. (2020). Women and Local Public Finance. Working Paper.

Casas-Arce, P. and Saiz, A. (2015). Women and Power: Unpopular, Unwilling, or Held Back? *Journal of Political Economy*, 123(3):641–669.

Charles, K. K., Guryan, J., and Pan, J. (2018). The Effects of Sexism on American Women: The Role of Norms vs. Discrimination. NBER Working Paper No. 24904.

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

Eggers, A. C., Freier, R., Grembi, V., and Nannicini, T. (2018). Regression Discontinuity Designs Based on Population Thresholds: Pitfalls and Solutions. *American Journal of Political Science*, 62(1):210–229.

Ferreira, F. and Gyourko, J. (2014). Does Gender Matter for Political Leadership? The Case of US Mayors. *Journal of Public Economics*, 112:24–39.

Fiva, J. H., Folke, O., and Sørensen, R. J. (2018). The Power of Parties: Evidence from Close Municipal Elections in Norway. *The Scandinavian Journal of Economics*, 120(1):3–30.

Fiva, J. H. and Smith, D. M. (2017). Local Candidates and Voter Mobilization: Evidence from Historical Two-Round Elections in Norway. *Electoral Studies*, 45:130–140.

Folke, O. (2014). Shades of Brown and Green: Party Effects in Proportional Election Systems. *Journal of the European Economic Association*, 12(5):1361–1395.

Folke, O., Persson, T., and Rickne, J. (2016). The Primary Effect: Preference Votes and Political Promotions. *American Political Science Review*, 110(3):559–578.

Folke, O. and Rickne, J. (2020). All the Single Ladies: Job Promotions and the Durability of Marriage. *American Economic Journal: Applied Economics*, 12(1):260–287.

Fontana, N., Nannicini, T., and Tabellini, G. (2017). Historical Roots of Political Extremism: The Effects of Nazi Occupation of Italy. IGIER Working Paper No. 618.

Fowler, A. and Hall, A. B. (2017). Long-Term Consequences of Election Results. *British Journal of Political Science*, 47(2):351–372.

Fox, R. L. and Lawless, J. L. (2004). Entering the Arena? Gender and the Decision to Run for Office. *American Journal of Political Science*, 48(2):264–280.

Fox, R. L. and Lawless, J. L. (2014). Uncovering the Origins of the Gender Gap in Political Ambition. *American Political Science Review*, 108(3):499–519.

Funk, P. and Gathmann, C. (2014). Gender Gaps in Policy Making: Evidence from Direct Democracy in Switzerland. *Economic Policy*, 30(81):141–181.

Gagliarducci, S. and Paserman, M. D. (2011). Gender Interactions within Hierarchies: Evidence from the Political Arena. *The Review of Economic Studies*, 79(3):1021–1052.

Gelman, A. and Imbens, G. (2019). Why High-order Polynomials Should not be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, 37:447–456.

Gentzkow, M. and Shapiro, J. M. (2006). Media Bias and Reputation. *Journal of Political Economy*, 114(2):280–316.

Gentzkow, M., Wong, M. B., and Zhang, A. T. (2018). Ideological Bias and Trust in Information Sources. Working paper.

Gneezy, U., Niederle, M., and Rustichini, A. (2003). Performance in Competitive Environments: Gender Differences. *The Quarterly Journal of Economics*, 118(3):1049–1074.

Goldin, C. and Rouse, C. (2000). Orchestrating Impartiality: The Impact of "Blind" Auditions on Female Musicians. *American Economic Review*, 90(4):715–741.

Gonzalez-Eiras, M. and Sanz, C. (2018). Women's Representation in Politics: Voter Bias, Party Bias, and Electoral Systems. Banco de Espana Working Paper No. 1834.

Grembi, V., Nannicini, T., and Troiano, U. (2016). Do Fiscal Rules Matter? *American Economic Journal: Applied Economics*, 8(3):1–30.

Hall, A. B. (2015). What Happens When Extremists Win Primaries? *American Political Science Review*, 109(1):18–42.

Heckman, J. J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47(1):153–161.

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The Allocation of Talent and US Economic Growth. *Econometrica*, 87(5):1439–1474.

Karpowitz, C. F., Monson, J. Q., and Preece, J. R. (2017). How to Elect More Women: Gender and Candidate Success in a Field Experiment. *American Journal of Political Science*, 61(4):927–943.

Kirkeboen, L. J., Leuven, E., and Mogstad, M. (2016). Field of Study, Earnings, and Self-selection. *The Quarterly Journal of Economics*, 131(3):1057–1111.

Kleven, H., Landais, C., and Søgaard, J. E. (2019). Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*, 11(4):181–209.

Kline, P. and Walters, C. R. (2019). On Heckits, LATE, and Numerical Equivalence. *Econometrica*, 87(2):677–696.

Lawless, J. L. and Pearson, K. (2008). The Primary Reason for Women's Underrepresentation? Reevaluating the Conventional Wisdom. *The Journal of Politics*, 70(1):67–82.

Le Barbanchon, T. and Sauvagnat, J. (2018). Voter Bias and Women in Politics. CEPR Discussion Paper 13238.

Lee, D. S. (2008). Randomized Experiments from Non-random selection in US House Elections. *Journal of Econometrics*, 142(2):675–697.

Lippmann, Q. (2019). Gender and Lawmaking in Times of Quotas. Working Paper.

Lott, Jr, J. R. and Kenny, L. W. (1999). Did Women's Suffrage Change the Size and Scope of Government? *Journal of Political Economy*, 107(6):1163–1198.

Markus, G. B. (1988). The Impact of Personal and National Economic Conditions on the Presidential Vote: A Pooled Cross-Sectional Analysis. *American Journal of Political Science*, pages 137–154.

McCrary, J. (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics*, 142(2):698–714.

Miller, G. (2008). Women's Suffrage, Political Responsiveness, and Child Survival in American History. *The Quarterly Journal of Economics*, 123(3):1287–1327.

OECD (2020). Distribution of Graduates and Entrants by Field.

Osborne, M. J. and Slivinski, A. (1996). A Model of Political Competition with Citizen-Candidates. *The Quarterly Journal of Economics*, 111(1):65–96.

Pettigrew, S., Owen, K., and Wanless, E. (2014). U.S. House Primary Election Results (1956-2010). https://doi.org/10.7910/DVN/26448.

Phelps, E. S. (1972). The Statistical Theory of Racism and Sexism. *American Economic Review*, 62(4):659–661.

Pons, V. and Tricaud, C. (2019). Coordination and Bandwagon Effects of Candidate Rankings: Evidence from Runoff Elections. NBER Working Paper No. 26599.

Shair-Rosenfield, S. (2012). The Alternative Incumbency Effect: Electing Women Legislators in Indonesia. *Electoral Studies*, 31(3):576–587.

Svaleryd, H. (2009). Women's Representation and Public Spending. *European Journal of Political Economy*, 25(2):186–198.

Thomsen, D. M. and King, A. S. (2020). Women's Representation and the Gendered Pipeline to Power. *American Political Science Review*, pages 1–12.

Thortnton, G. (2017). Women in Business 2017: New Perspectives on Risk and Reward. https://grantthornton.global.

Wasserman, M. (2018). Gender Differences in Politician Persistence. Working Paper.

Appendix

A Figures

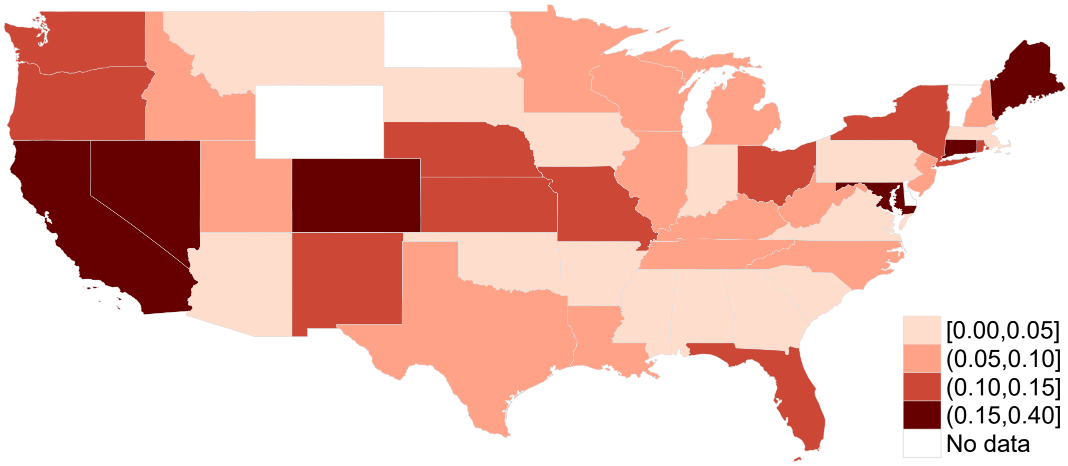

Figure A1: Share of female members of the US House of Representatives (1976–2010), by state

Notes: All members of the US House of representatives elected between 1976 and 2010. The map reports the share of female representatives in each state. Source: author's elaboration on the US Census Bureau geodata with reference to the 116^{th} Congress.

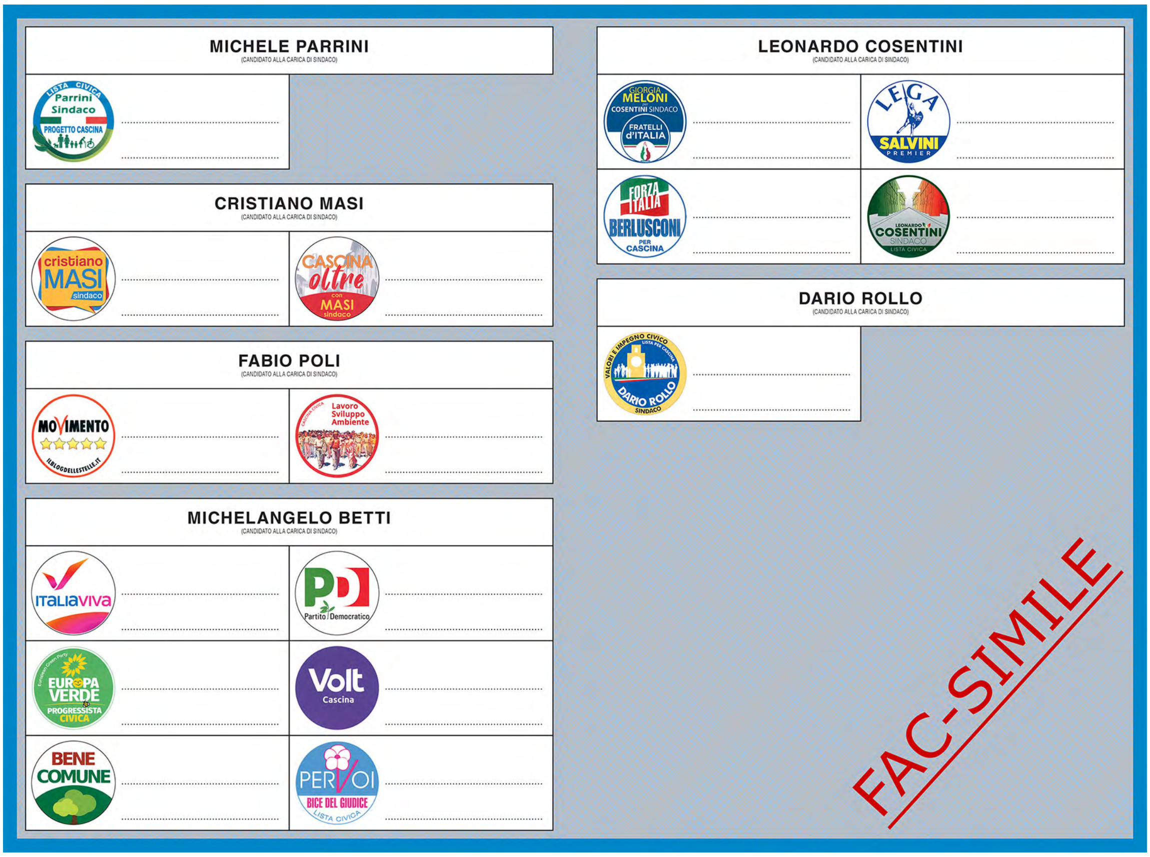

Figure A2: Example of voting ballot (First round)

Notes: Example of voting ballot used in the first round of a runoff election in Italian municipalities having more than 15,000 residents. Source: Italian Ministry of Internal Affairs.

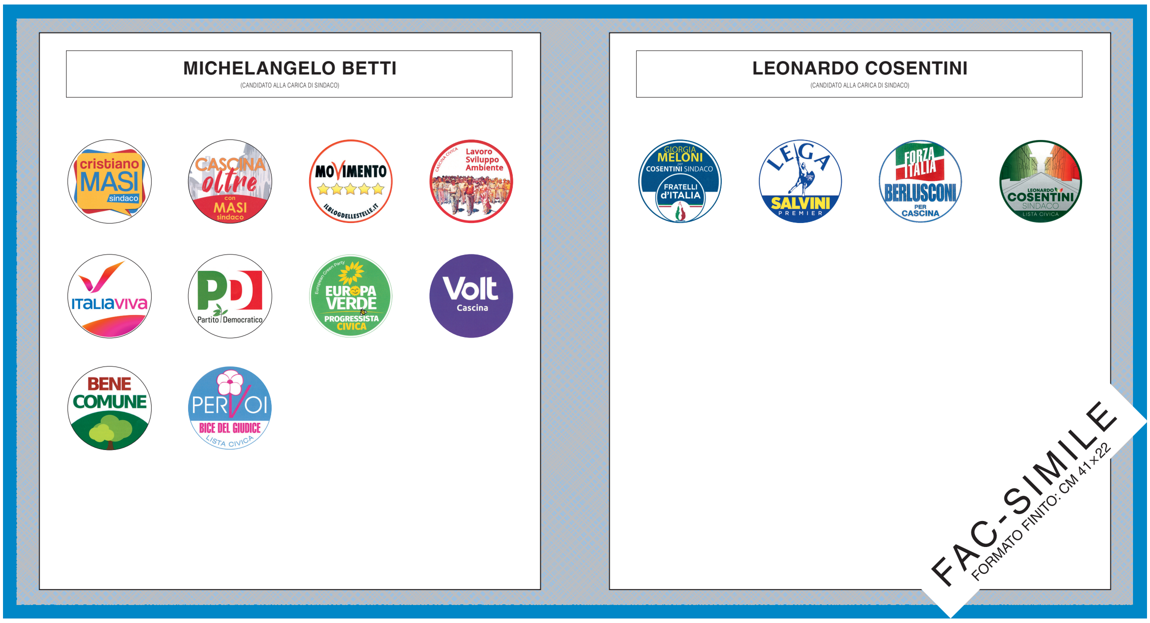

Figure A3: Example of voting ballot (Second round)

Notes: Example of voting ballot used in the second round of a runoff election in Italian municipalities having more than 15,000 residents. Source: Italian Ministry of Internal Affairs.

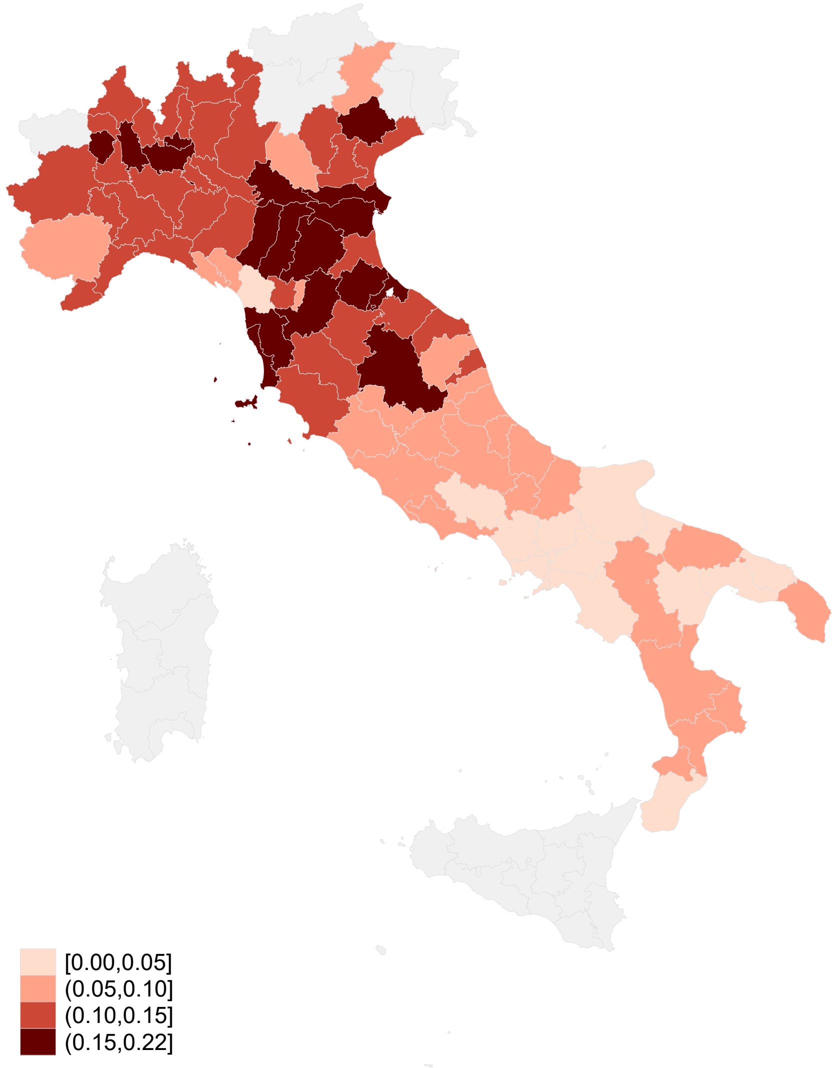

Figure A4: Share of female mayors (1993–2019), by province

Notes: All mayors elected in Italian municipalities elected between 1993 and 2019. The map reports the share of female mayors by each province in the sample. Special regions: *Friuli-Venezia Giulia, Sardegna, Sicilia, Trentino-Alto Adige* and *Valle d'Aosta* are excluded. Source: author's elaboration on ISTAT geodata.

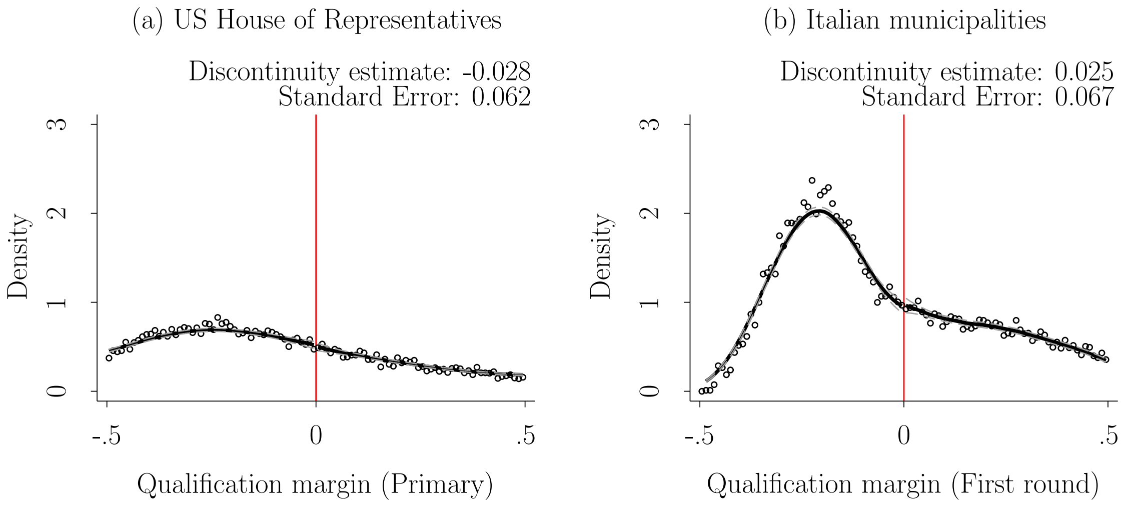

Figure A5: McCrary (2008) test

Notes: Notes: Log-density discontinuity and standard errors are computed performing a formal McCrary (2008) with optimal bandwidth. Markers represent sample averages within bins of the running variable $h_{i,t}$ equal to 0.01.

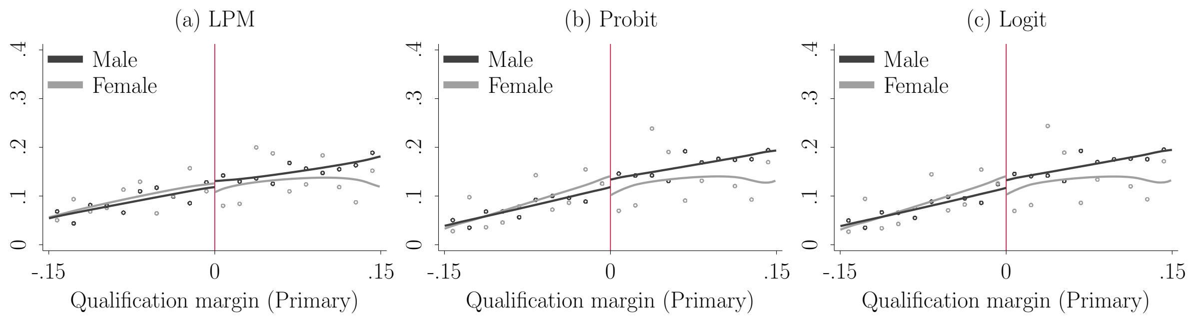

Figure A6: Gender of the candidate and predicted probability of winning the election (US House of Representatives)

Notes: The dependent variable is the predicted probability of winning the election to the position of member of the US House of Representatives, estimated using a linear probability model (Panel (a)), a Probit model (Panel (b)) and a Logit model (Panel (c)) as in equation (10). In all panels, solid lines represent non-parametric smoothers of the qualification margin, separately estimated on either side of the qualification threshold and for male and female candidates. Light gray lines and markers refer to female candidates, while dark gray lines and markers refer to male candidates. Markers represent, for each gender, sample averages within bins of the running variable $h_{i,t}$ equal to 0.015.

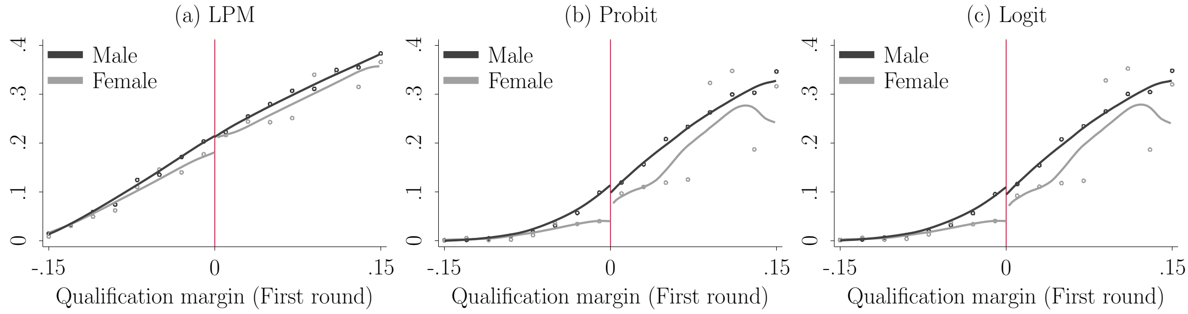

Figure A7: Gender of the candidate and predicted probability of winning the election (Italian municipalities)

Notes: The dependent variable is the predicted probability of winning the election to the position of mayor, estimated using a linear probability model (Panel (a)), a Probit model (Panel (b)) and a Logit model (Panel (c)) as in equation (10). In all panels, solid lines represent non-parametric smoothers of the qualification margin, separately estimated on either side of the qualification threshold and for male and female candidates. Light gray lines and markers refer to female candidates, while dark gray lines and markers refer to male candidates. Markers represent, for each gender, sample averages within bins of the running variable $h_{i,t}$ equal to 0.015.

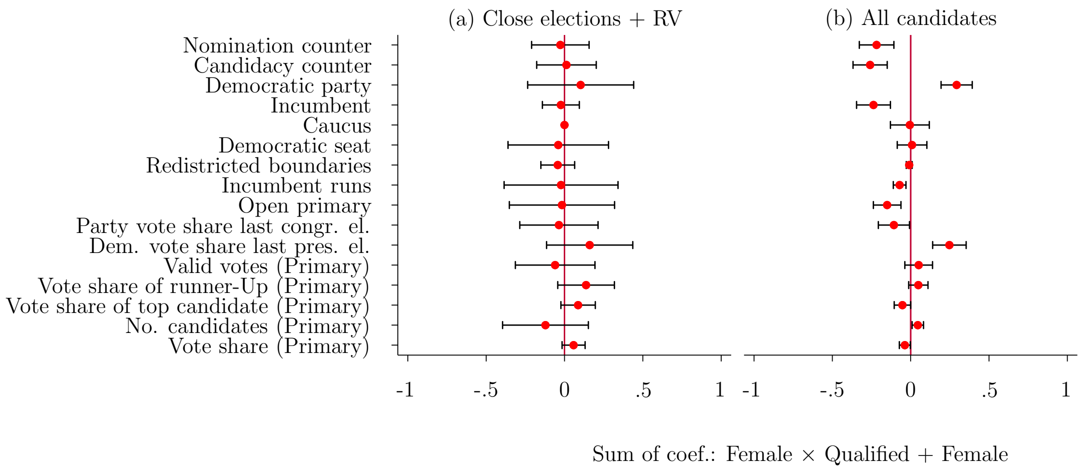

Figure A8: Characteristics of male and female US Primary winners

Notes: The figure reports estimated coefficients and confidence intervals for $\beta_1 + \beta_3$ in equation (8), where each $Y_{i,t}$ reported in the figure has been previously standardized to have mean equal to 0 and variance equal to 1. In Panel (a), the sample is restricted to candidates within the Calonico et al. (2014) optimal bandwidth, while in Panel (b), all candidates are included and the control for $f(F_i, h_{i,t})$ is omitted. 95% confidence intervals are based on standard errors robust to clustering at the congressional district-by-party level.

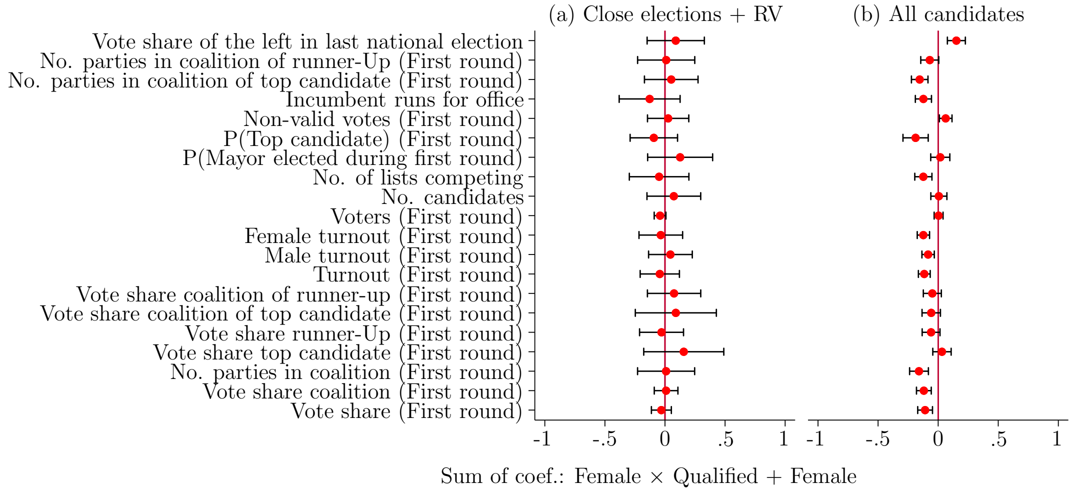

Figure A9: Characteristics of male and female candidates admitted to second round (Italian municipalities)

Notes: The figure reports estimated coefficients and confidence intervals for $\beta_1+\beta_3$ in equation (8), where each $Y_{i,t}$ reported in the figure has been previously standardized to have mean equal to 0 and variance equal to 1. In Panel (a), the sample is restricted to candidates within the Calonico et al. (2014) optimal bandwidth, while in Panel (b), all candidates are included and the control for $f(F_i, h_{i,t})$ is omitted. 95% confidence intervals are based on standard errors robust to clustering at the municipality level.

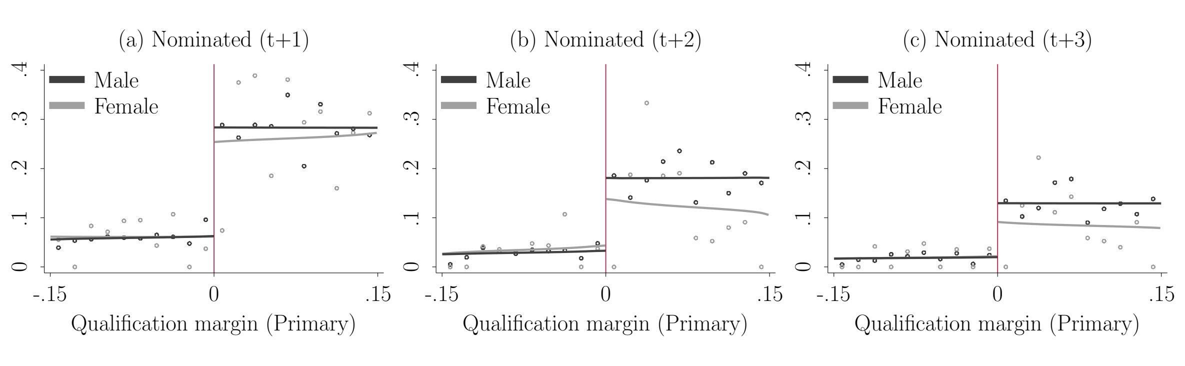

Figure A10: US Primary victory and future nominations, by gender

Notes: The dependent variable is an indicator equal to 1 if the candidate is nominated by her party to compete in the general elections held after two, four, and six years, respectively. In all panels, the solid lines represent non-parametric smoothers of the qualification margin, separately estimated on either side of the qualification threshold and for male and female candidates. Light gray lines and markers refer to female candidates, while dark gray lines and markers refer to male candidates. Markers represent, for each gender, sample averages within bins of the running variable $h_{i,t}$ equal to 0.15.

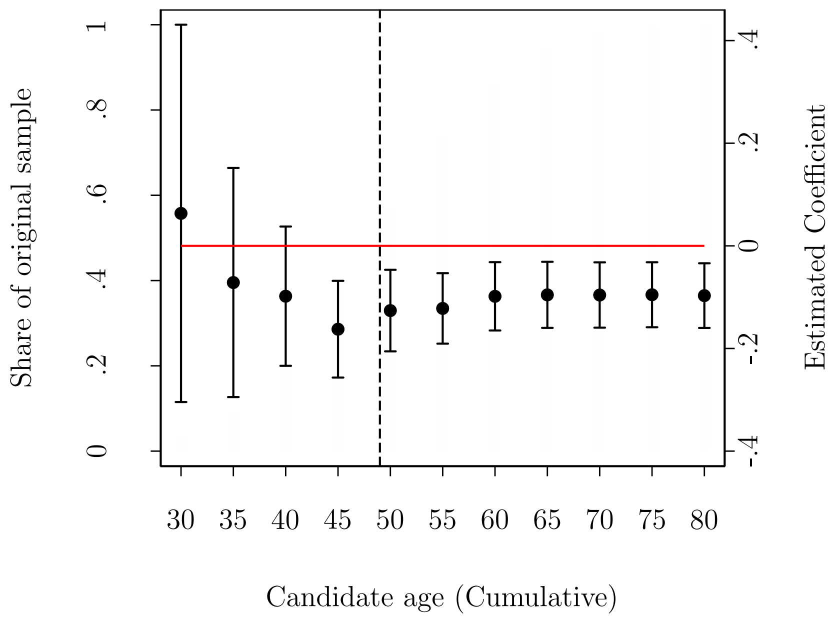

Figure A11: Estimated coefficients by age groups (Italian municipalities)

Notes: This figure reports estimated coefficients and 95% confidence intervals for versions of equation (8) in which the sample is restricted to candidate below the age specified on the horizontal axis. Bars, as measured in the left axis, represent the share of the original sample included in each regression. The dependent variable is a dummy equal to 1 if the candidate is elected as mayor during the second round. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. 95% Confidence intervals are based on standard errors robust to clustering at the municipality level.

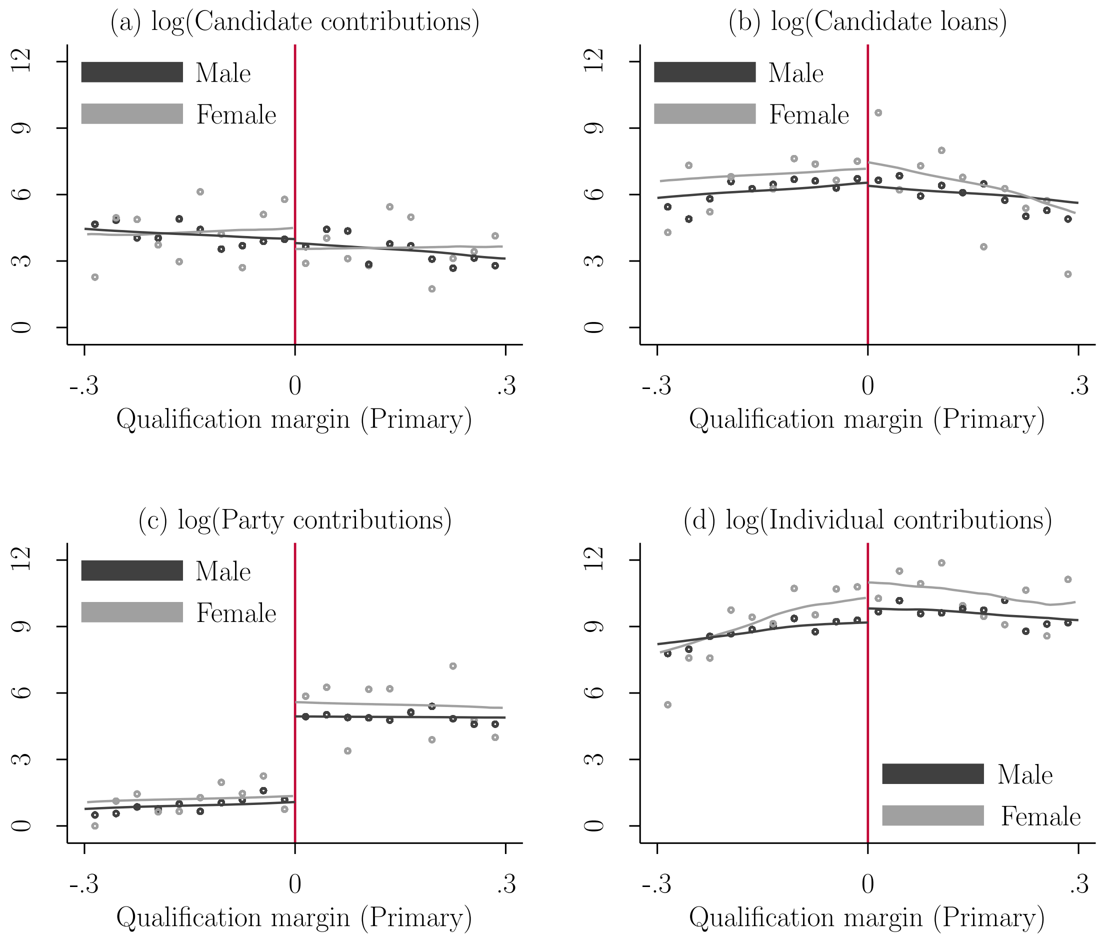

Figure A12: US Primary victory and composition of campaign financing, by gender

Notes: In Panel (a), the dependent variable is the natural logarithm of the total contribution made by the candidate to her campaign; in Panel (b), the dependent variable is the natural logarithm of the total loans subscribed by the candidate to finance her campaign; in Panel (c), the dependent variable is the natural logarithm of the total amount that the candidate's campaign received by her political party; in Panel (d), the dependent variable is the natural logarithm of the total amount that the candidate's campaign received by individual donors. In all panels, the solid lines represent non-parametric smoothers of the qualification margin, separately estimated on either side of the qualification threshold and for male and female candidates. Light gray lines and markers refer to female candidates, while dark gray lines and markers refer to male candidates. Markers represent, for each gender, sample averages within bins of the running variable $h_{i,t}$ equal to 0.03.

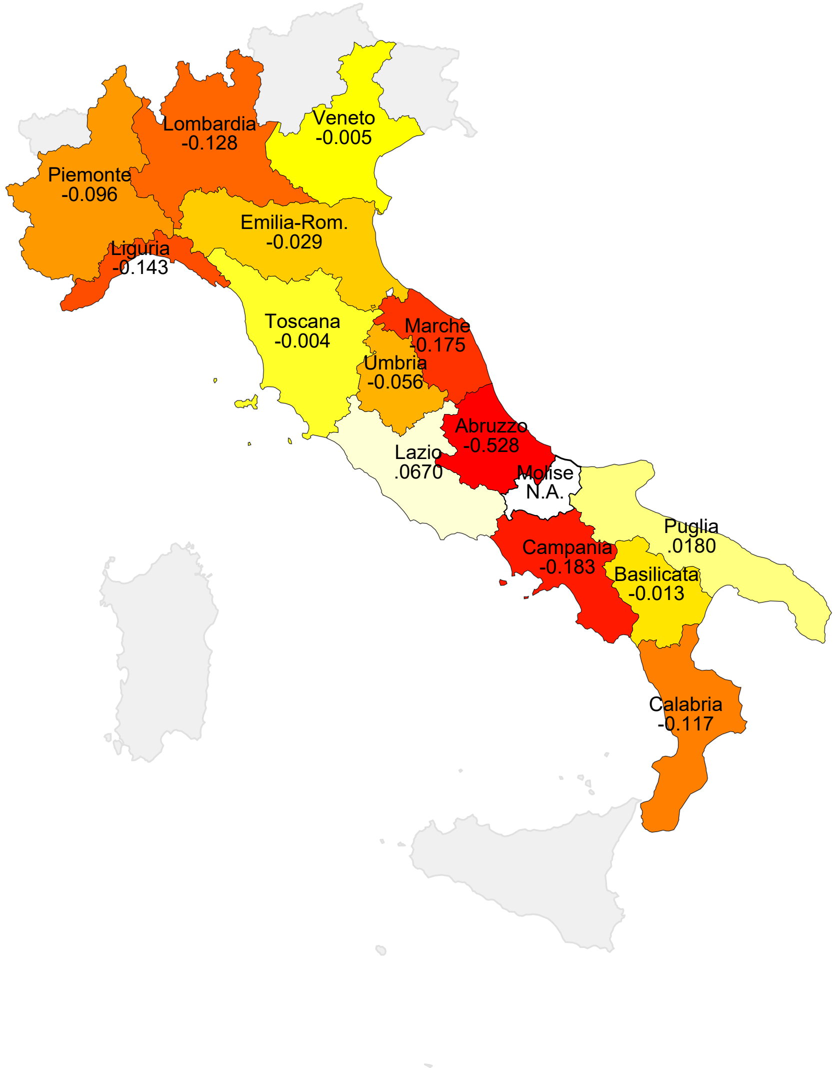

Figure A13: Gender gap in the probability of winning the election, estimated by region

Notes: The map reports estimated coefficients for β_1 from estimating equation (8) separately for each of the fifteen ordinary regions of Italy. The dependent variable is an indicator equal to 1 if the candidate is elected to the position of mayor at the end of the second round. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth.

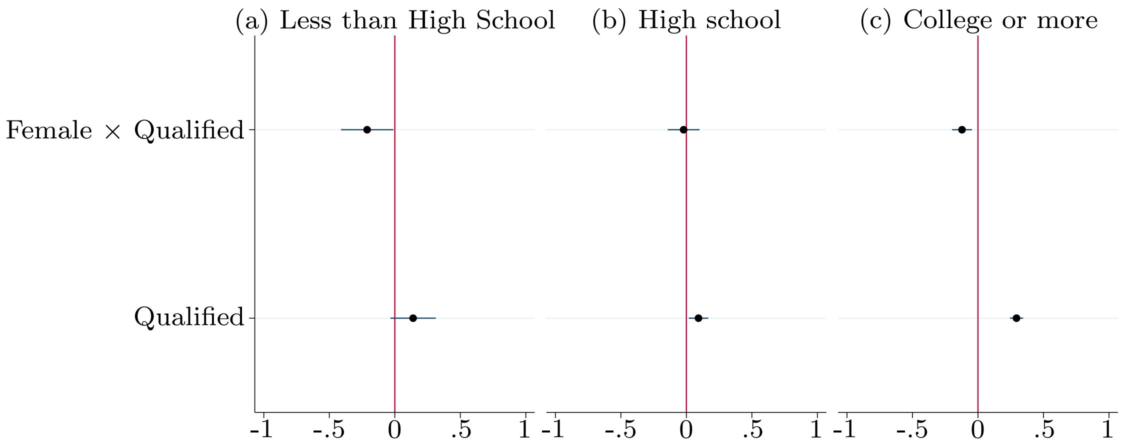

Figure A14: Heterogeneous effects by education level of the candidate

Notes: The dependent variable is a dummy equal to 1 if the candidate is elected as mayor during the second round. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. In each panel, the sample is restricted to candidates who, at the time of the election, hold the educational degree specified in the title. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. 95% confidence intervals are based on standard errors robust to clustering at the municipality level.

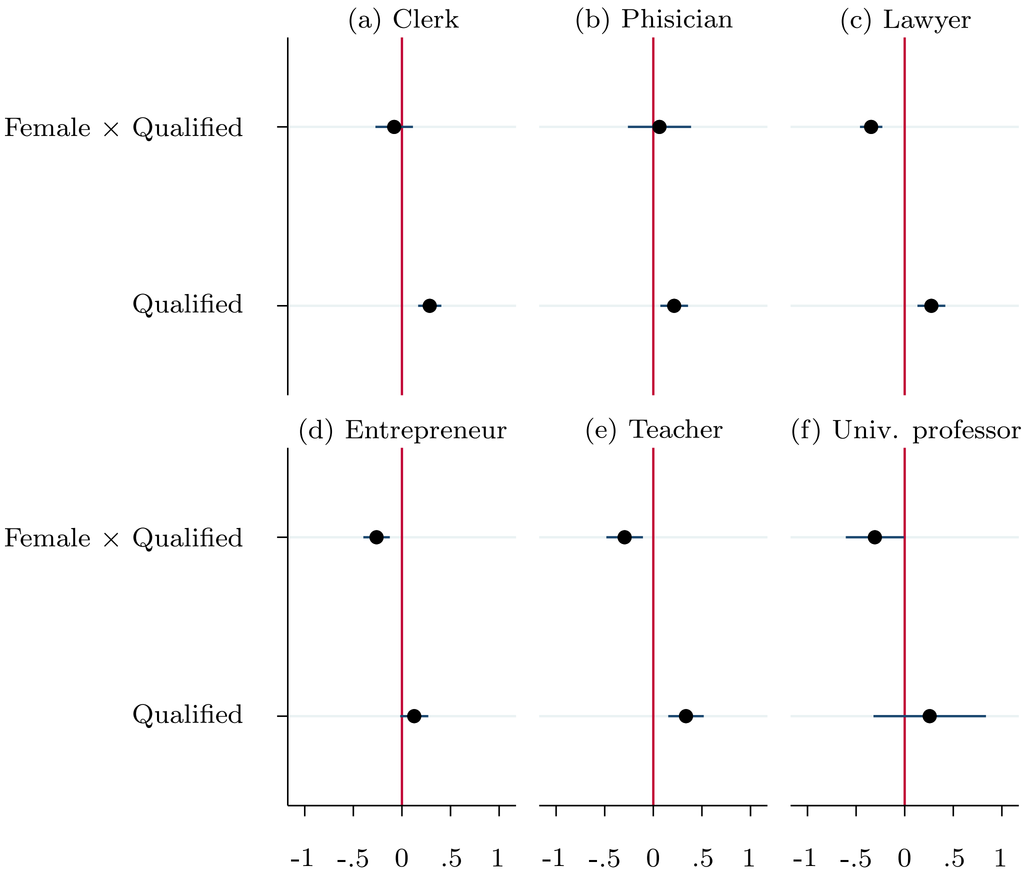

Figure A15: Heterogeneous effects by profession of the candidate

Notes: The dependent variable is a dummy equal to 1 if the candidate is elected as mayor during the second round. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. In each panel, the sample is restricted to candidates who, at the time of the election, were employed in the profession specified in the title of the panel. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. 95% confidence intervals are based on standard errors robust to clustering at the municipality level.

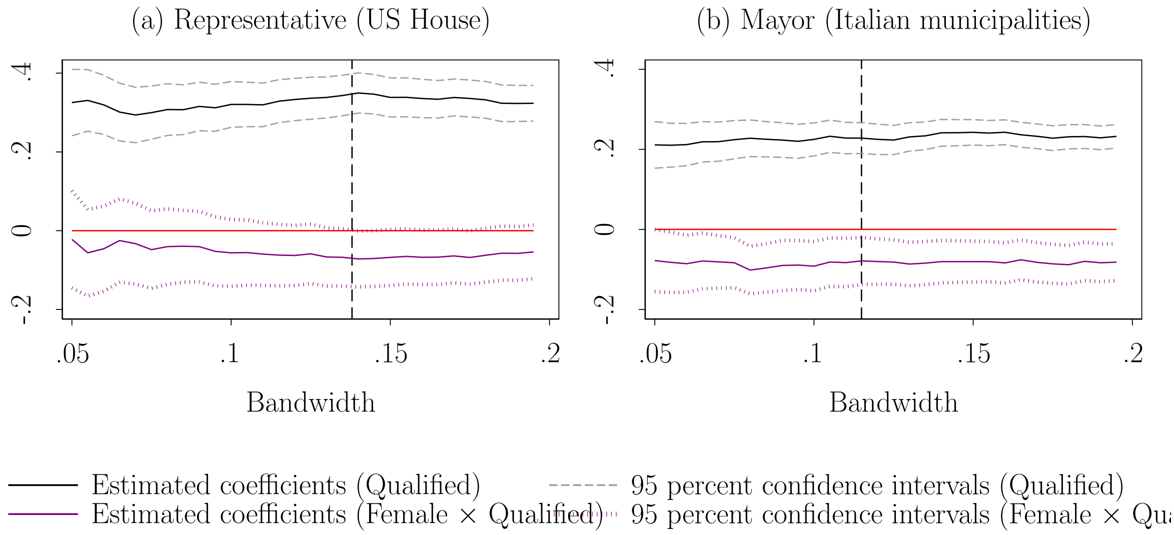

Figure A16: Qualification to final round and probability of winning the election (Bandwidth sensitivity)

Notes: Panel (a): The dependent variable is an indicator equal to 1 if the candidate is elected as member of the US House of Representatives during the general election. Panel (b): The dependent variable is a dummy equal to 1 if the candidate is elected as mayor during the second round. In both panels, the horizontal axis represents the bandwidths used to fit the local linear regression. The black solid line represents the estimated coefficients for β_2 as a function of the chosen bandwidth, while the purple solid line represents the estimated coefficients for β_1 as a functon of the chosen bandwidth. Dashed lines represent the 95 percent confidence intervals of each coefficient. Estimation method: local linear regression with uniform kernel as in equation (8), and bandwidths ranging from $h_{i,t} = 0.05$ to $h_{i,t} = 0.20$. The vertical line represents the Calonico et al. (2014) optimal bandwidth. All specifications include the interaction terms between the runoff dummy and the assignment variable and election year dummies. Each estimation concerns a variation of bandwidth equal to 0.005. 95% confidence intervals are based on standard errors robust to clustering at the congressional district-by-party level (Panel (a)) and municipality level (Panel (b)).

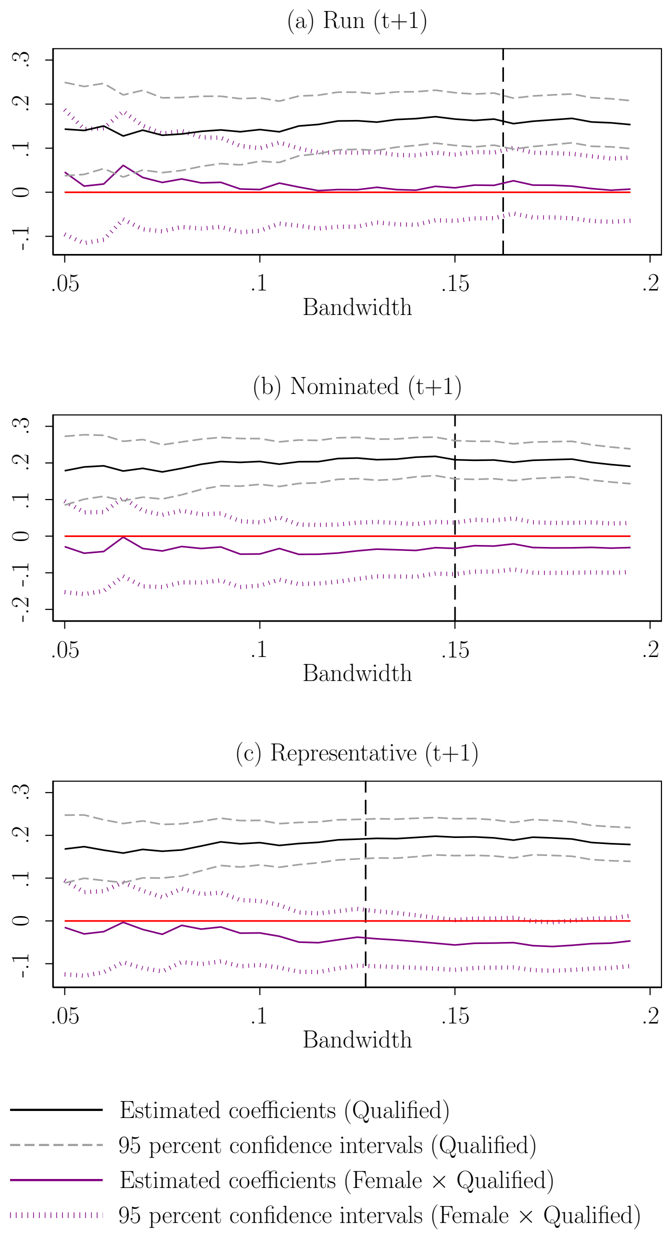

Figure A17: US Primary victory and future career outcomes (Bandwidth sensitivity)

Notes: In Panel (a), the dependent variable is an indicator equal to 1 if the candidate runs in the primary election for the position of member of the US House of Representatives during the elections held after two years; in Panel (b), the dependent variable is an indicator equal to 1 if the candidate is nominated by her party to compete in the general elections held after two years; in Panel (c), the dependent variable is an indicator equal to 1 if the candidate is elected in the US House of Representatives during the next elections held after two years. In all panels, the horizontal axis represents the bandwidths used to fit the local linear regression. The black solid line represents the estimated coefficients for β_2 as a function of the chosen bandwidth, while the purple solid line represents the estimated coefficients for β_1 as a functon of the chosen bandwidth. Dashed lines represent the 95 percent confidence intervals of each coefficient. Estimation method: local linear regression with uniform kernel as in equation (8), and bandwidths ranging from $h_{i,t} = 0.05$ to $h_{i,t} = 0.20$. The vertical line represents the Calonico et al. (2014) optimal bandwidth. All specifications include the interaction terms between the Qualified dummy and the assignment variable and election year dummies. Each estimation concerns a variation of bandwidth equal to 0.005. 95% confidence intervals are based on standard errors robust to clustering at the congressional district-by-party level.

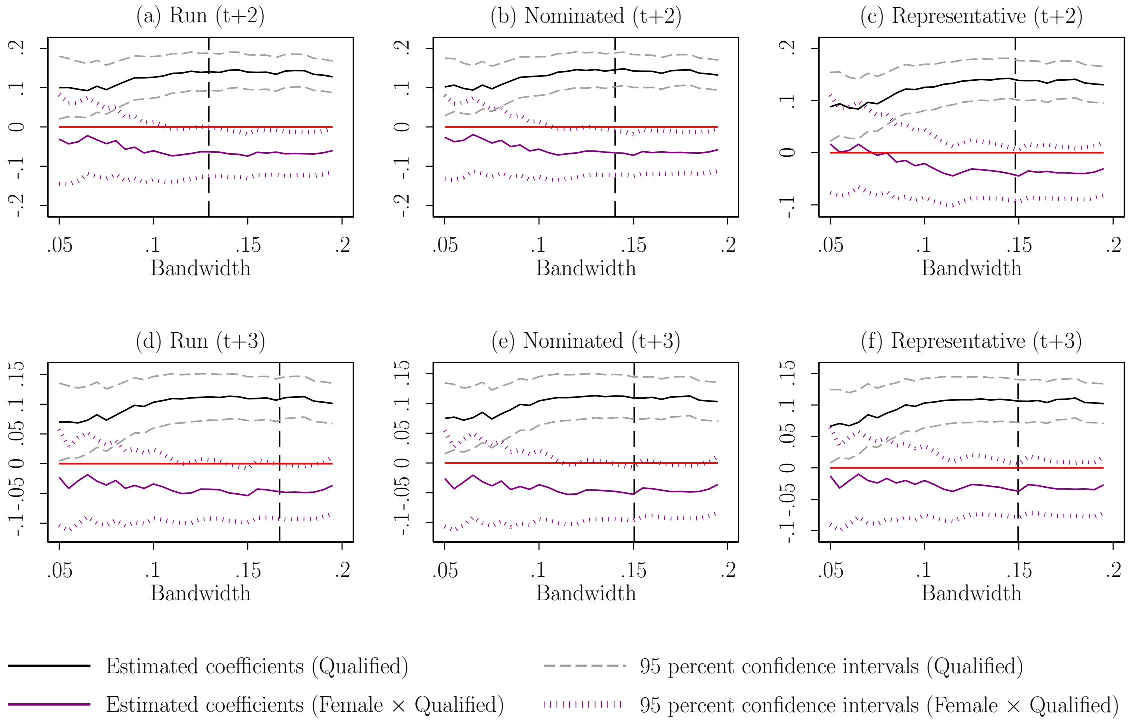

Figure A18: US Primary victory and future career outcomes (Bandwidth sensitivity)

Notes: In Panels (a) and (d), the dependent variable is an indicator equal to 1 if the candidate runs in the primary election for the position of member of the US House of Representatives during the elections held after four and six years, respectively; in Panels (b) and (e), the dependent variable is an indicator equal to 1 if the candidate is nominated by her party to compete in the general elections held after four and six years, respectively; in Panels (c) and (f), the dependent variable is an indicator equal to 1 if the candidate is elected in the US House of Representatives during the elections held after four and six years, respectively. In all panels, the horizontal axis represents the bandwidths used to fit the local linear regression. In all panels, the horizontal axis represents the bandwidths used to fit the local linear regression. The black solid line represents the estimated coefficients for β_2 as a function of the chosen bandwidth, while the purple solid line represents the estimated coefficients for β_1 as a functon of the chosen bandwidth. Dashed lines represent the 95 percent confidence intervals of each coefficient. Estimation method: local linear regression with uniform kernel as in equation (8), and bandwidths ranging from $h_{i,t} = 0.05$ to $h_{i,t} = 0.20$. The vertical line represents the Calonico et al. (2014) optimal bandwidth. All specifications include the interaction terms between the Qualified dummy and the assignment variable and election year dummies. Each estimation concerns a variation of bandwidth equal to 0.005. 95% confidence intervals are based on standard errors robust to clustering at the congressional district-by-party level.

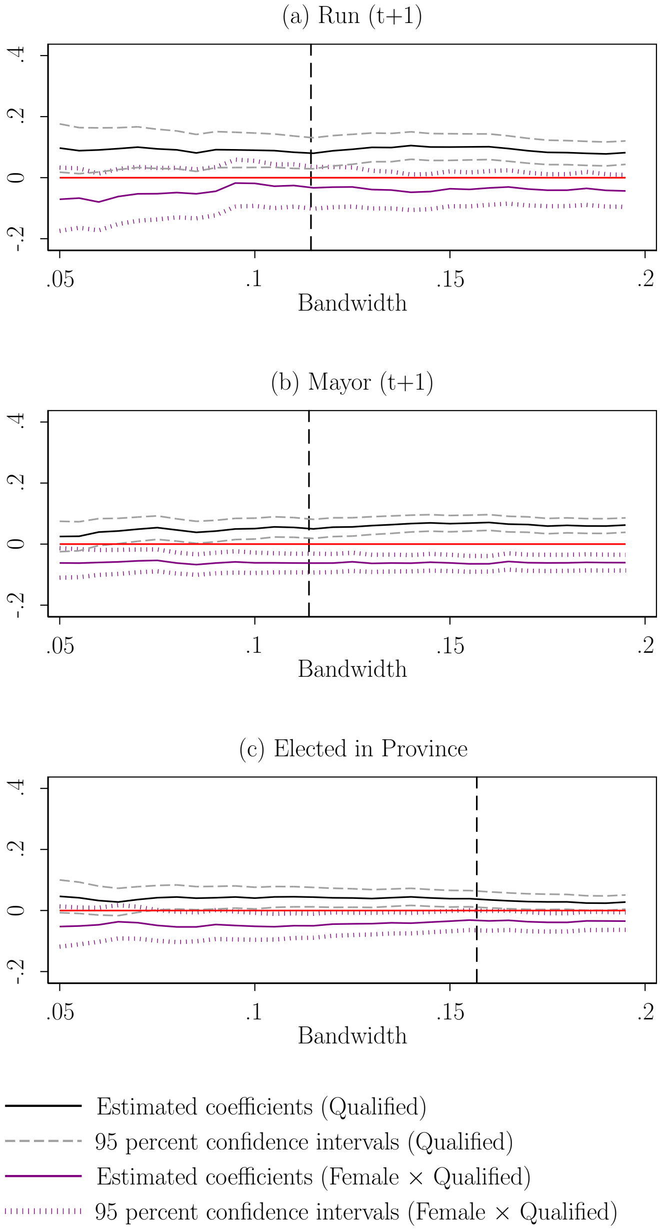

Figure A19: Qualification to second round and future career outcomes (Italian municipalities)

Notes: In Panel (a), the dependent variable is a dummy equal to 1 if the candidate runs for office in the occasion of elections held in term $t+1$. In Panel (b), the dependent variable is a dummy equal to 1 if the candidate is elected as mayor during elections held in term $t+1$. In Panel (c), the dependent variable is a dummy equal to 1 if the candidate is in the future elected as a member of the Provincial Council or to the position of President of the Province. In all panels, the horizontal axis represents the bandwidths used to fit the local linear regression. The black solid line represents the estimated coefficients for β_2 as a function of the chosen bandwidth, while the purple solid line represents the estimated coefficients for β_1 as a functon of the chosen bandwidth. Dashed lines represent the 95 percent confidence intervals of each coefficient. Estimation method: local linear regression with uniform kernel as in equation (8), and bandwidths ranging from $h_{i,t} = 0.05$ to $h_{i,t} = 0.20$. The vertical line represents the Calonico et al. (2014) optimal bandwidth. All specifications include the interaction terms between the Qualified dummy and the assignment variable and election year dummies. Each estimation concerns a variation of bandwidth equal to 0.005. 95% confidence intervals are based on standard errors robust to clustering at the municipality level.

B Tables

Table B1: Descriptive statistics (US House of Representatives)

	(1)	(2)	(3)
Variable	Mean	St. Dev.	Obs.
	(a) Career Indicators		
Representative	0.310	0.462	24,107
Run (t+1)	0.304	0.460	24,107
Nominated (t+1)	0.259	0.438	24,107
Representative (t+1)	0.214	0.410	24,107
Run (t+2)	0.174	0.379	24,107
Nominated (t+2)	0.162	0.369	24,107
Representative (t+2)	0.149	0.356	24,107
Run (t+3)	0.113	0.317	24,107
Nominated (t+3)	0.109	0.312	24,107
Representative (t+3)	0.103	0.304	24,107
	(b) Individual Characteristics		
Female	0.113	0.317	24,107
Incumbent	0.279	0.448	24,107
Democratic party	0.517	0.500	24,107
Candidacy counter	1.476	2.615	24,107
Nomination counter	1.407	2.566	24,105
	(c) Election returns		
No. candidates (Primary)	2.810	2.382	24,107
Vote share of top candidate (Primary)	0.726	0.252	24,107
Vote share of runner-Up (Primary)	0.173	0.153	24,107
Valid votes (Primary)	34,846	32,005	24,107
Vote share (Primary)	0.571	0.376	24,107
Voters (General)	190,185	59,824	24,107
Valid votes (General)	183,830	58,625	24,107
Party vote Share (General)	0.535	0.204	24,107
Vote share of top candidate (General)	0.665	0.126	24,107
Vote share of runner-up (General)	0.335	0.126	24,107
Vote share of dem. candidate (General)	0.547	0.189	23,546
	(d) District characteristics		
Dem. vote share last pres. el.	0.505	0.132	23,627
Party vote share last congr. el.	0.537	0.216	22,199
Open primary	0.369	0.483	24,107
Unopposed primary	0.338	0.473	24,107
Caucus	0.0114	0.106	24,107
Incumbent runs	0.804	0.397	24,107
Redistricted boundaries	0.207	0.405	24,107
Democratic seat	0.563	0.496	24,107
	(e) Campaign activity		
Campaign contributions (USD × 1,000)	665.3	1,829	8,382
Campaign spending (USD × 1,000)	646.1	1,834	8,382
Candidate contributions (USD × 1,000)	11.23	136.9	8,381
Candidate loans (USD × 1,000)	50.73	354.7	8,380
Party contributions (USD × 1,000)	5.083	19.87	8,382
Individual contributions (USD × 1,000)	340.7	1,083	8,380

Notes: All candidates competing in primary elections for the US House of Representatives between 1976 and 2010 and included in Pettigrew et al. (2014) dataset. Information in Panel (e) refers to candidate merged based on surname, congressional district, party and year with campaign contribution data released by the US Federal Election Commission (1979–2010).

Table B2: Descriptive statistics (Italian municipalities)

	(1)	(2)	(3)
Variable	Mean	St. Dev.	Obs.
(a) Career Indicators			
Mayor	0.188	0.390	19,037
Mayor (t+1)	0.0684	0.253	19,037
Run (t+1)	0.188	0.391	19,037
Elected in Province	0.0596	0.237	19,037
(b) Individual Characteristics			
Female candidate	0.124	0.330	18,115
College	0.644	0.479	11,779
High-skill job	0.0336	0.180	11,622
Age	49.26	10.60	17,377
Incumbent	0.0865	0.281	19,037
Candidacy counter	0.260	0.591	19,037
Run again	0.194	0.396	19,037
Left-wing	0.211	0.408	19,037
Right-wing	0.175	0.380	19,037
Independent	0.633	0.482	19,037
(c) Election returns			
No. candidates	6.067	2.249	19,037
No. of lists competing	14.05	5.427	19,037
Vote share top candidate (First round)	0.458	0.111	19,037
Vote share runner-Up (First round)	0.282	0.0780	19,036
Voters (First round)	42,696	132,556	19,037
Turnout (First round)	0.749	0.0852	19,037
Male turnout (First round)	0.759	0.0966	14,917
Female turnout (First round)	0.745	0.0969	14,917
Non-valid votes (First round)	0.0433	0.0164	19,037
Vote share coalition of top candidate (First round)	0.433	0.116	19,037
Vote share coalition of runner-up (First round)	0.262	0.0778	19,036
Vote share (First round)	0.188	0.179	19,037
No. seats	1.548	3.105	19,019
No. parties in coalition (First round)	2.475	2.032	19,037
Vote share coalition (First round)	0.175	0.172	19,037
Proportion of disjoint votes	0.0128	0.0272	19,037
Turnout (Second round)	0.614	0.105	11,417
Vote share (Second round)	0.500	0.0911	4,266
No. parties in coalition (Second round)	4.282	2.191	4,266
No. endorsements	0.294	0.838	4,266
Vote share coalition (Second round)	0.411	0.122	4,266
Male turnout (Second round)	0.617	0.131	9,817
Female turnout (Second round)	0.607	0.127	9,817
Vote share coalition of top candidate (Second round)	0.456	0.120	12,394
Vote share coalition of runner-up (Second round)	0.353	0.108	12,394
(d) Municipality characteristics			
Census population	70,095	222,073	19,037
Surface (km^2)	96.61	141.6	18,983
Elderly (%)	0.140	0.0427	18,983
Migrants (%)	0.00506	0.00391	18,983
Students (%)	0.0791	0.0140	18,983
Unemployment (%)	0.0277	0.0128	18,983
South	0.322	0.467	19,037
Vote share of the left in last national election	0.242	0.104	18,510

Notes: All mayoral candidates in the municipalities belonging to the ordinary regions running for election between 1993 and 2019 under the runoff voting system.

Table B3: Balance of pre-determined characteristics (US House of Representatives)

	(1)	(2)	(3)
Variable	Coefficient	St. Error	Obs.
Female	0.033	(0.059)	5,219
Incumbent	0.035	(0.032)	3,037
Democratic party	0.002	(0.022)	4,025
Candidacy counter	-0.039	(0.038)	3,549
Nomination counter	-0.007	(0.040)	3,823
Vote share (Primary)	0.013	(0.012)	3,865
Party vote Share (General)	-0.018	(0.016)	3,245
No. candidates (Primary)	-0.052	(0.041)	4,124
Vote share of top candidate (Primary)	0.020	(0.018)	3,865
Valid votes (Primary)	0.002	(0.022)	3,733
Voters (General)	0.002	(0.012)	5,803
Valid votes (General)	-0.003	(0.012)	5,900
Vote share of top candidate (General)	0.014	(0.016)	3,331
Vote share of runner-up (General)	-0.012	(0.015)	3,312
Vote share of dem. candidate (General)	0.008	(0.019)	3,742
Dem. vote share last pres. el.	0.008	(0.019)	5,015
Party vote share last congr. el.	-0.016	(0.018)	3,285
Open primary	0.017	(0.018)	5,154
Incumbent runs	0.021	(0.026)	4,339
Redistricted boundaries	0.002	(0.009)	4,526
Democratic seat	0.023	(0.022)	4,347

Notes: The dependent variable is specified in each row. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) where β_1 and β_3 are restricted to be equal 0 with uniform kernel and Calonico et al. (2014) optimal bandwidth. Each variable reported in the table has been previously standardized to have mean equal to 0 and variance equal to 1. Standard errors robust to clustering at the party-by-electoral district level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B4: Balance of pre-determined characteristics (Italian municipalities)

	(1)	(2)	(3)
Variable	Coefficient	St. Error	Obs.
Female candidate	-0.008	(0.054)	5,417
College	0.024	(0.074)	2,446
High-skill job	-0.053	(0.070)	2,960
Age	0.056	(0.061)	4,074
Incumbent	-0.020	(0.054)	3,748
Candidacy counter	-0.062	(0.062)	4,499
Run again	-0.070	(0.057)	4,302
Vote share (First round)	-0.013	(0.009)	2,657
Vote share coalition (First round)	0.005	(0.015)	2,874
No. parties in coalition (First round)	0.032	(0.037)	4,137
Vote share top candidate (First round)	0.004	(0.027)	2,535
Vote share runner-Up (First round)	-0.003	(0.019)	2,548
Vote share coalition of top candidate (First round)	0.016	(0.026)	2,534
Vote share coalition of runner-up (First round)	-0.005	(0.019)	2,670
Turnout (First round)	0.006	(0.015)	3,378
Male turnout (First round)	0.013	(0.018)	3,029
Female turnout (First round)	0.007	(0.014)	2,709
Voters (First round)	0.006	(0.005)	2,139
No. candidates	0.024	(0.020)	3,031
No. of lists competing	0.011	(0.015)	3,779
Non-valid votes (First round)	0.001	(0.016)	4,617
Incumbent runs for office	0.005	(0.021)	4,322
No. parties in coalition of top candidate (First round)	0.013	(0.018)	3,473
No. parties in coalition of runner-Up (First round)	0.030**	(0.015)	3,035
Census population	0.002	(0.005)	2,278
Surface (km^2)	-0.004	(0.013)	4,319
Elderly (%)	-0.004	(0.026)	4,244
Migrants (%)	-0.001	(0.023)	4,293
Students (%)	0.010	(0.028)	4,532
Unemployment (%)	0.001	(0.033)	6,610
South	-0.016	(0.023)	3,908
Vote share of the left in last national election	-0.012	(0.019)	4,159
P(Mayor elected during first round)	0.014	(0.018)	2,765

Notes: The dependent variable is specified in each row. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8), where β_1 and β_3 are restricted to be equal 0 with uniform kernel and Calonico et al. (2014) optimal bandwidth. Each variable reported in the table has been previously standardized to have mean equal to 0 and variance equal to 1. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B5: Prediction of general election outcome from US primary returns

	LPM	Probit	Logit
	(1)	(2)	(3)
	Representative	Representative	Representative
Vote share (Primary)	0.318***	0.510***	0.507***
	(0.008)	(0.032)	(0.032)
No. candidates (Primary)	0.010***	0.023***	0.022***
	(0.001)	(0.004)	(0.004)
Vote share of runner-Up (Primary)	0.161***	0.332***	0.325***
	(0.015)	(0.030)	(0.030)
Valid votes (Primary)	0.000***	0.000***	0.000***
	(0.000)	(0.000)	(0.000)
Dem. vote share last pres. el.	0.068**	0.068**	0.066**
	(0.026)	(0.028)	(0.030)
Dem. vote share last congr. el.	-0.093***	-0.085***	-0.082***
	(0.017)	(0.017)	(0.017)
Open primary	-0.002	0.002	0.001
	(0.007)	(0.007)	(0.007)
Incumbent runs	-0.092***	-0.076***	-0.077***
	(0.005)	(0.006)	(0.006)
Redistricted boundaries	0.005	0.005	0.005
	(0.010)	(0.009)	(0.010)
Democratic seat	0.012**	0.013**	0.012**
	(0.006)	(0.006)	(0.006)
Caucus	-0.039	-0.028*	-0.023
	(0.025)	(0.015)	(0.015)
Incumbent	0.716***	0.189***	0.167***
	(0.006)	(0.005)	(0.005)
Democratic party	-0.023***	-0.022***	-0.021***
	(0.003)	(0.003)	(0.003)
Candidacy counter	0.002*	0.001	0.001
	(0.001)	(0.001)	(0.001)
Nomination counter	0.002	0.002	0.002
	(0.001)	(0.001)	(0.001)
Party vote share last congr. el.	0.281***	0.239***	0.227***
	(0.012)	(0.013)	(0.013)
Observations	21,232	21,229	21,229
R^2	0.754		
Pseudo R^2		0.728	0.729
Individual Char.	✓	✓	✓
District Fixed Effects	✓	✓	✓
Election Year Effects	✓	✓	✓
Mean Dep. Var.	0.298	0.298	0.298

Notes: Estimation of the probability of winning the election to the position of member of the House of Representatives using primary returns as in equation (10). The dependent variable is an indicator equal to 1 if the candidate is elected during the general election. In Column (1), the predicted probability is estimated using a Linear Probability Model; in Column (2), the predicted probability is estimated using a Probit Model; in Column (3), the predicted probability is estimated using a Logit Model. In Columns (2) and (3), the table reports marginal effects at the mean. Standard errors robust to clustering at the congressional district-by-party level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B6: Prediction of second round outcome from first-round returns (Italian municipalities)

	LPM	Probit	Logit
	(1)	(2)	(3)
	Mayor	Mayor	Mayor
Vote share top candidate (First round)	-0.109	-1.024***	-1.069***
	(0.087)	(0.041)	(0.044)
Vote share runner-Up (First round)	-0.137	-0.979***	-0.983***
	(0.102)	(0.045)	(0.045)
Vote share (First round)	2.503***	2.332***	2.378***
	(0.086)	(0.078)	(0.079)
Vote share coalition (First round)	-0.510***	-0.985***	-1.025***
	(0.091)	(0.082)	(0.084)
Vote share coalition of top candidate (First round)	0.038	0.452***	0.474***
	(0.085)	(0.041)	(0.044)
Vote share coalition of runner-up (First round)	0.073	0.403***	0.413***
	(0.102)	(0.046)	(0.046)
No. parties in coalition (First round)	-0.035***	-0.009***	-0.008***
	(0.002)	(0.002)	(0.002)
Turnout (First round)	0.303	1.106	1.182*
	(3.933)	(0.706)	(0.712)
Male turnout (First round)	-0.103	-0.542	-0.577*
	(1.905)	(0.342)	(0.345)
Female turnout (First round)	-0.116	-0.576	-0.615*
	(2.026)	(0.363)	(0.366)
Voters (First round)	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
No. candidates	0.013***	-0.001	-0.001
	(0.002)	(0.001)	(0.001)
No. of lists competing	0.005***	0.003***	0.002***
	(0.001)	(0.001)	(0.001)
Non-valid votes (First round)	-0.159	-0.007	-0.011
	(0.289)	(0.036)	(0.037)
Incumbent runs for office	-0.001	-0.000	-0.001
	(0.005)	(0.001)	(0.001)
Observations	14,916	14,916	14,916
R^2	0.566		
Pseudo R^2		0.733	0.732
Individual Char.	✓	✓	✓
Municipality Fixed Effects	✓	✓	✓
Election Year Effects	✓	✓	✓
Mean Dep. Var.	0.184	0.184	0.184

Notes: Estimation of the probability of winning the election to the position of mayor using first-round returns as in equation (10). The dependent variable is an indicator equal to 1 if the candidate is elected as mayor during the second round. In Column (1), the predicted probability is estimated using a Linear Probability Model; in Column (2), the predicted probability is estimated using a Probit Model; in Column (3), the predicted probability is estimated using a Logit Model. In Columns (2) and (3), the table reports marginal effects at the mean. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B7: Gender of the candidate and predicted probability of winning the election (US House of Representatives)

	(1)	(2)	(3)	(4)	(5)	(6)
	Predicted repres. (LPM)	Predicted repres. (LPM)	Predicted repres. (Probit)	Predicted repres. (Probit)	Predicted repres. (Logit)	Predicted repres. (Logit)
Female × Qualified	-0.024		-0.023		-0.026	
	(0.019)		(0.023)		(0.023)	
Qualified	0.012	0.009	0.017	0.013	0.018	0.014
	(0.012)	(0.011)	(0.012)	(0.011)	(0.012)	(0.011)
Observations	2,832	2,832	2,661	2,661	2,686	2,686
R^2	0.037	0.037	0.049	0.048	0.050	0.050
Election Year Effects	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.114	0.114	0.115	0.115	0.114	0.114
Bandwidth	0.130	0.130	0.122	0.122	0.123	0.123

Notes: In Columns (1) and (2), the dependent variable is the predicted probability of winning the election to the position of member of the House of Representatives estimated using equation (10) and a Linear Probability Model; in Columns (3) and (4), the dependent variable is the predicted probability of winning the election to the position of member of the House of Representatives estimated using equation (10) and a Probit Model; in Columns (5) and (6), the dependent variable is the predicted probability of winning the election to the position of member of the House of Representatives estimated using equation (10) and a Logit Model. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. In Columns (2), (4) and (6), the coefficients β_1 and β_3 are restricted to be equal zero. Standard errors robust to clustering at the congressional district-by-party level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B8: Gender of the candidate and predicted probability of winning the election (Italian municipalities)

	(1)	(2)	(3)	(4)	(5)	(6)
	Predicted mayor (LPM)	Predicted mayor (LPM)	Predicted mayor (Probit)	Predicted mayor (Probit)	Predicted mayor (Logit)	Predicted mayor (Logit)
Female × Qualified	-0.003		-0.027		-0.029	
	(0.011)		(0.018)		(0.018)	
Qualified	0.003	0.004	0.007	0.004	0.004	0.001
	(0.005)	(0.004)	(0.010)	(0.010)	(0.011)	(0.010)
Observations	2,652	2,664	1,944	1,954	1,918	1,927
R^2	0.369	0.368	0.184	0.178	0.179	0.174
Election Year Effects	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.201	0.201	0.101	0.100	0.0997	0.0995
Bandwidth	0.0884	0.0884	0.0657	0.0657	0.0649	0.0649

Notes: In Columns (1) and (2), the dependent variable is the predicted probability of winning the election to the position of mayor estimated using equation (10) and a Linear Probability Model; in Columns (3) and (4), the dependent variable is the predicted probability of winning the election to the position of mayor estimated using equation (10) and a Probit Model; in Columns (5) and (6), the dependent variable is the predicted probability of winning the election to the position of mayor estimated using equation (10) and a Logit Model. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. In Columns (2), (4) and (6), the coefficients β_1 and β_3 are restricted to be equal zero. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B9: US primary victory and future career outcomes, by gender

	(1)	(2)	(3)	(4)
	Representative	Run (t+1)	Nominated (t+1)	Representative (t+1)
Female × Qualified	-0.071**	0.016	-0.033	-0.037
	(0.036)	(0.038)	(0.036)	(0.033)
Qualified	0.347***	0.162***	0.209***	0.195***
	(0.026)	(0.030)	(0.027)	(0.023)
Observations	3,406	3,986	3,710	3,144
R^2	0.212	0.087	0.116	0.103
Election Year Effects	✓	✓	✓	✓
Mean Dep. Var.	0.130	0.208	0.151	0.102
Bandwidth	0.138	0.162	0.150	0.127

Notes: In Column (1), the dependent variable is an indicator equal to 1 if the candidate is elected as member of the US House of Representatives during the general election, as in Table 1; in Column (2), the dependent variable is an indicator equal to 1 if the candidate runs in the primary election for the position of member of the US House of Representatives during the next election, held after two years; in Column (3), the dependent variable is an indicator equal to 1 if the candidate is nominated by her party to compete in the general election held after two years; in Column (4), the dependent variable is an indicator equal to 1 if the candidate is elected in the US House of Representatives during the next election, held after two years. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the congressional district-by-party level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B10: US primary victory and future career outcomes, by gender

	(1)	(2)	(3)	(4)	(5)	(6)
	Run (t+2)	Nominated (t+2)	Representative (t+2)	Run (t+3)	Nominated (t+3)	Representative (t+3)
Female × Qualified	-0.063**	-0.066**	-0.042*	-0.049**	-0.052**	-0.037*
	(0.032)	(0.029)	(0.025)	(0.023)	(0.023)	(0.021)
Qualified	0.140***	0.146***	0.142***	0.109***	0.109***	0.106***
	(0.024)	(0.022)	(0.019)	(0.018)	(0.018)	(0.017)
Observations	3,197	3,453	3,665	4,085	3,718	3,702
R^2	0.073	0.084	0.083	0.068	0.071	0.072
Election Year Effects	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.103	0.0898	0.0729	0.0636	0.0616	0.0559
Bandwidth	0.129	0.140	0.148	0.167	0.151	0.150

Notes: In Columns (1) and (2), the dependent variables are indicators equal to 1 if the candidate runs in the primary election for the position of member of the US House of Representatives during the elections held after four and six years, respectively. In Columns (3) and (4), the dependent variables are indicators equal to 1 if the candidate wins her party's primary election for the position of member of the US House of Representatives during the elections held after four and six years, respectively. In Columns (5) and (6), the dependent variables are indicators equal to 1 if the candidate is elected position of member of the US House of Representatives in the occasion of the elections held after four and six years, respectively. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the congressional district-by-party level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B11: Qualification to second round and future career outcomes, by gender (Italian municipalities)

	(1)	(2)	(3)	(4)
	Mayor	Run (t+1)	Mayor (t+1)	Elected in Province
Female × Qualified	-0.097***	-0.028	-0.066***	-0.049***
	(0.032)	(0.044)	(0.019)	(0.019)
Qualified	0.232***	0.106***	0.065***	0.035**
	(0.021)	(0.030)	(0.019)	(0.015)
Observations	3,582	3,505	3,554	5,491
R^2	0.200	0.079	0.060	0.022
Election Year Effects	✓	✓	✓	✓
Mean Dep. Var.	0.129	0.243	0.0802	0.0750
Bandwidth	0.115	0.113	0.115	0.175

Notes: In Column (1), the dependent variable is an indicator equal to 1 if the candidate is elected as mayor during the second round, as in Table 1. In Column (2), the dependent variable is a dummy equal to 1 if the candidate runs for office in the occasion of elections held in term $t+1$. In Column (3), the dependent variable is a dummy equal to 1 if the candidate is elected as mayor during elections held in term $t+1$. In Column (4), the dependent variable is a dummy equal to 1 if the candidate is in the future elected as a member of the Provincial Council or to the position of President of the Province. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B12: Electoral success and future career outcomes, by gender

	(1)	(2)	(3)	(4)	(5)	(6)
	Run (t+1)	Mayor (t+1)	Elected in Province	Run (t+1)	Mayor (t+1)	Elected in Province
Female × Mayor				0.052***	-0.003	-0.001
				(0.018)	(0.015)	(0.008)
Mayor	0.187***	0.171***	0.030***	0.178***	0.170***	0.029***
	(0.013)	(0.012)	(0.005)	(0.013)	(0.012)	(0.005)
Observations	19,624	23,186	34,933	19,505	23,037	34,681
R^2	0.189	0.127	0.024	0.190	0.127	0.024
Election Year Effects	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.398	0.230	0.0529	0.400	0.231	0.0533
Bandwidth	0.0840	0.100	0.158	0.0840	0.100	0.158

Notes: In Columns (1) and (4), the dependent variable is a dummy equal to 1 if the candidate runs for office in the occasion of elections held in term $t+1$. In Columns (2) and (5), the dependent variable is a dummy equal to 1 if the candidate is elected as mayor during elections held in term $t+1$. In Columns (3) and (6), the dependent variable is a dummy equal to 1 if the candidate is in the future elected as a member of the Provincial Council or to the position of President of the Province. All specifications include the interaction terms between the Mayor dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (11) with uniform kernel and Calonico et al. (2014) optimal bandwidth. In Columns (1)–(3), the coefficients γ_1 and γ_3 are restricted to zero. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B13: US primary victory and campaign financing, by gender

	(1)	(2)	(3)	(4)	(5)	(6)
	log(Campaign contributions)	log(Campaign spending)	log(Candidate contributions)	log(Candidate loans)	log(Party contributions)	log(Individual contributions)
Female × Qualified	-0.054	-0.043	-0.373	0.191	0.290	0.206
	(0.260)	(0.260)	(0.592)	(0.618)	(0.395)	(0.389)
Qualified	0.696***	0.639***	0.233	0.307	3.636***	0.645***
	(0.171)	(0.169)	(0.350)	(0.375)	(0.265)	(0.227)
Observations	1,598	1,593	2,211	2,336	2,605	1,667
R^2	0.351	0.303	0.130	0.209	0.407	0.695
Election Year Effects	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	11.67	11.66	3.921	6.445	2.649	9.512
Bandwidth	0.156	0.155	0.225	0.236	0.265	0.165

Notes: In Column (1), the dependent variable is the natural logarithm of the campaign contributions by the candidate; in Column (2), the dependent variable is the natural logarithm of the amount of resources invested by the candidate to finance her campaign; in Column (3), the dependent variable is the natural logarithm of the total contribution made by the candidate to her campaign; in Column (4), the dependent variable is the natural logarithm of the total loans subscribed by the candidate to finance her campaign; in Column (5), the dependent variable is the natural logarithm of the total amount that the candidate's campaign received by her political party; in Column (6), the dependent variable is the natural logarithm of the total amount that the candidate's campaign received by individual donors. All specifications include the interaction terms between the Qualified dummy dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the congressional district-by-party level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B14: Heterogeneity by election year

	(a) US House of Representatives Election year		(b) Italian municipalities Election year	
	Before 1992	From 1992	Before 2006	From 2006
	(1)	(2)	(1)	(2)
	Representative	Representative	Mayor	Mayor
Female × Qualified	-0.168***	-0.038	-0.120**	-0.067
	(0.054)	(0.046)	(0.048)	(0.043)
Qualified	0.318***	0.366***	0.230***	0.237***
	(0.040)	(0.034)	(0.028)	(0.030)
Observations	1,494	1,910	1,931	1,629
R^2	0.204	0.220	0.210	0.185
Election Year Effects	✓	✓	✓	✓
Mean Dep. Var.	0.128	0.131	0.133	0.122
Bandwidth	0.138	0.138	0.115	0.115

Notes: Panel (a): The dependent variable is a dummy equal to 1 if the candidate is elected as member of the House of Representatives during the general election. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Panel (b): The dependent variable is a dummy equal to 1 if the candidate is elected as mayor during the second round. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. In Panel (a), the sample is restricted to elections held between 1974 and 1990 in Column (1) and between 1992 and 2010 in Column (2). In Panel (b), the sample is restricted to elections held between 1993 and 2005 in Column (1) and between 2006 and 2019 in Column (2). Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the congressional district-by-party level (Panel (a)) and municipality level (Panel (b)) are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B15: Heterogeneity by party affiliation of the candidate

	(a) US House of Representatives Party affiliation		(b) Italian municipalities Party affiliation		
	Democratic	Republican	Right	Left	Independent
	(1)	(2)	(1)	(2)	(3)
	Representative	Representative	Mayor	Mayor	Mayor
Female × Qualified	-0.053	-0.095*	-0.089	-0.084	-0.121**
	(0.047)	(0.055)	(0.065)	(0.051)	(0.051)
Qualified	0.355***	0.340***	0.283***	0.207***	0.214***
	(0.038)	(0.036)	(0.044)	(0.037)	(0.030)
Observations	1,563	1,841	918	985	1,744
R^2	0.233	0.212	0.268	0.182	0.211
Election Year Effects	✓	✓	✓	✓	✓
Mean Dep. Var.	0.131	0.128	0.190	0.130	0.0975
Bandwidth	0.138	0.138	0.115	0.115	0.115

Notes: Panel (a): The dependent variable is a dummy equal to 1 if the candidate is elected as member of the US House of Representatives during the general election. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Panel (b): The dependent variable is a dummy equal to 1 if the candidate is elected as mayor during the second round. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. In Panel (a), the sample is restricted to candidates competing in the Democratic primary in Column (1), and to candidates competing in the Republican primary in Column (2). In panel (b), the sample is restricted to candidates affiliated with the centre-left coalition in Column (1), with the centre-right coalition in Column (2), and to candidates not affiliated with any of the two coalitions in Column (3). Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the congressional district-by-party level (Panel (a)) and municipality level (Panel (b)) are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B16: Quadratic and cubic specifications (US House of Representatives)

(a) Quadratic fit

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Representative	Run (t+1)	Nominated (t+1)	Representative (t+1)	Run (t+2)	Nominated (t+2)	Representative (t+2)	Run (t+3)	Nominated (t+3)	Representative (t+3)
Female × Qualified	-0.031	0.014	0.003	-0.020	-0.051**	-0.054**	-0.028	-0.038**	-0.037**	-0.031*
	(0.031)	(0.030)	(0.029)	(0.027)	(0.026)	(0.026)	(0.025)	(0.018)	(0.019)	(0.018)
Qualified	0.324***	0.133***	0.190***	0.192***	0.147***	0.139***	0.140***	0.094***	0.101***	0.101***
	(0.027)	(0.032)	(0.029)	(0.026)	(0.027)	(0.025)	(0.023)	(0.019)	(0.019)	(0.018)
Observations	6,901	7,559	6,812	6,157	5,865	5,644	5,488	7,939	7,508	7,374
R^2	0.234	0.109	0.157	0.130	0.096	0.101	0.094	0.079	0.084	0.082
Election Year Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.116	0.183	0.132	0.0898	0.0916	0.0808	0.0676	0.0534	0.0514	0.0475
Bandwidth	0.281	0.310	0.277	0.249	0.237	0.230	0.224	0.327	0.307	0.302

(b) Cubic fit

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Representative	Run (t+1)	Nominated (t+1)	Representative (t+1)	Run (t+2)	Nominated (t+2)	Representative (t+2)	Run (t+3)	Nominated (t+3)	Representative (t+3)
Female × Qualified	-0.050*	0.015	0.003	-0.021	-0.053**	-0.054**	-0.029	-0.038**	-0.037**	-0.031*
	(0.028)	(0.030)	(0.029)	(0.027)	(0.026)	(0.026)	(0.025)	(0.018)	(0.019)	(0.018)
Qualified	0.331***	0.173***	0.216***	0.185***	0.111***	0.116***	0.097***	0.113***	0.106***	0.105***
	(0.031)	(0.042)	(0.039)	(0.034)	(0.034)	(0.032)	(0.030)	(0.025)	(0.024)	(0.024)
Observations	9,071	7,559	6,812	6,157	5,865	5,644	5,488	7,939	7,508	7,374
R^2	0.248	0.109	0.157	0.130	0.097	0.102	0.096	0.080	0.084	0.082
Election Year Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.110	0.183	0.132	0.0898	0.0916	0.0808	0.0676	0.0534	0.0514	0.0475
Bandwidth	0.381	0.310	0.277	0.249	0.237	0.230	0.224	0.327	0.307	0.302

Notes: In Column (1), the dependent variable is an indicator equal to 1 if the candidate is elected as member of the US House of Representatives during the general election. In Columns (2), (5), and (8), the dependent variable is an indicator equal to 1 if the candidate runs for office in her party's primary elections held after 2, 4, and 6 years, respectively. In Columns (3), (6), and (9), the dependent variable is an indicator equal to 1 if the candidate wins her party's primary elections held after 2, 4, and 6 years, respectively. In Columns (4), (7), and (10), the dependent variable is an indicator equal to 1 if the candidate is elected to the US House of Representatives in the elections held after 2, 4, and 6 years, respectively. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-quadratic regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth in Panel (a); local-cubic regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth in Panel (b). Standard errors robust to clustering at the congressional district-by-party level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B17: Quadratic and cubic specifications (Italian municipalities)

(a) Quadratic fit

	(1)	(2)	(3)	(4)
	Mayor	Run (t+1)	Mayor (t+1)	Elected in Province
Female × Qualified	-0.094***	-0.050	-0.059***	-0.051**
	(0.027)	(0.036)	(0.016)	(0.021)
Qualified	0.242***	0.093**	0.069***	0.043
	(0.024)	(0.039)	(0.021)	(0.027)
Observations	6,020	4,652	6,061	4,448
R^2	0.231	0.077	0.058	0.021
Election Year Effects	✓	✓	✓	✓
Mean Dep. Var.	0.132	0.233	0.0734	0.0764
Bandwidth	0.188	0.144	0.190	0.138

(b) Cubic fit

	(1)	(2)	(3)	(4)
	Mayor	Run (t+1)	Mayor (t+1)	Elected in Province
Female × Qualified	-0.078***	-0.047	-0.060***	-0.040*
	(0.025)	(0.036)	(0.017)	(0.021)
Qualified	0.225***	0.102**	0.030	0.045
	(0.030)	(0.052)	(0.028)	(0.036)
Observations	7,529	5,006	6,109	4,783
R^2	0.255	0.075	0.059	0.021
Election Year Effects	✓	✓	✓	✓
Mean Dep. Var.	0.143	0.233	0.0732	0.0746
Bandwidth	0.239	0.155	0.191	0.148

Notes: In Column (1), the dependent variable is an indicator equal to 1 if the candidate is elected as mayor during the second round. In Column (2), the dependent variable is a dummy equal to 1 if the candidate runs for office in the occasion of elections held in term $t+1$. In Column (3), the dependent variable is a dummy equal to 1 if the candidate is elected as mayor during elections held in term $t+1$. In Column (4), the dependent variable is a dummy equal to 1 if the candidate is in the future elected as a member of the Provincial Council or to the position of President of the Province. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-quadratic regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth in panel (a); local-cubic regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth in Panel (b). Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B18: Alternative running variables (US House of Representatives)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Representative	Run (t+1)	Nominated (t+1)	Representative (t+1)	Run (t+2)	Nominated (t+2)	Representative (t+2)	Run (t+3)	Nominated (t+3)	Representative (t+3)
(a) Baseline running variable										
Female × Qualified	-0.071**	0.016	-0.033	-0.037	-0.063**	-0.066**	-0.042	-0.049**	-0.052**	-0.037*
	(0.036)	(0.038)	(0.036)	(0.033)	(0.032)	(0.029)	(0.025)	(0.024)	(0.023)	(0.022)
Qualified	0.347***	0.162***	0.209***	0.195***	0.140***	0.146***	0.142***	0.109***	0.109***	0.106***
	(0.026)	(0.030)	(0.027)	(0.023)	(0.024)	(0.022)	(0.019)	(0.018)	(0.018)	(0.017)
Bandwidth	0.138	0.162	0.150	0.127	0.129	0.140	0.148	0.167	0.151	0.150
(b) Only first and second recipient										
Female × Qualified	-0.072*	0.009	-0.029	-0.039	-0.067**	-0.074**	-0.038	-0.049*	-0.046*	-0.034
	(0.037)	(0.042)	(0.038)	(0.034)	(0.033)	(0.030)	(0.027)	(0.026)	(0.025)	(0.023)
Qualified	0.345***	0.171***	0.211***	0.197***	0.144***	0.142***	0.140***	0.110***	0.107***	0.106***
	(0.026)	(0.032)	(0.028)	(0.024)	(0.025)	(0.022)	(0.020)	(0.019)	(0.018)	(0.018)
Bandwidth	0.137	0.153	0.150	0.127	0.132	0.146	0.141	0.155	0.155	0.142
(c) Assign extra votes to turnout										
Female × Qualified	-0.065**	0.009	-0.039	-0.045	-0.067**	-0.054*	-0.036	-0.055**	-0.045*	-0.037*
	(0.028)	(0.046)	(0.039)	(0.035)	(0.032)	(0.030)	(0.026)	(0.026)	(0.025)	(0.022)
Qualified	0.288***	0.158***	0.193***	0.173***	0.132***	0.137***	0.136***	0.103***	0.106***	0.106***
	(0.015)	(0.034)	(0.028)	(0.024)	(0.024)	(0.022)	(0.020)	(0.020)	(0.019)	(0.016)
Bandwidth	1.998	0.112	0.128	0.128	0.128	0.139	0.150	0.128	0.143	0.196
(d) Weight by $1/ncandidates$										
Female × Qualified	-0.065**	0.005	-0.042	-0.046	-0.075**	-0.065**	-0.039*	-0.057***	-0.044**	-0.032
	(0.029)	(0.050)	(0.040)	(0.029)	(0.030)	(0.027)	(0.022)	(0.021)	(0.022)	(0.020)
Qualified	0.227***	0.117***	0.162***	0.123***	0.099***	0.098***	0.094***	0.073***	0.072***	0.075***
	(0.014)	(0.036)	(0.029)	(0.021)	(0.025)	(0.022)	(0.018)	(0.019)	(0.017)	(0.015)
Bandwidth	1.998	0.112	0.128	0.128	0.128	0.139	0.150	0.128	0.143	0.196
(e) Assign extra votes to other candidates										
Female × Qualified	-0.054	0.017	-0.024	-0.041	-0.053*	-0.051*	-0.021	-0.045**	-0.038*	-0.027
	(0.034)	(0.037)	(0.035)	(0.029)	(0.028)	(0.028)	(0.026)	(0.023)	(0.023)	(0.022)
Qualified	0.319***	0.150***	0.191***	0.180***	0.125***	0.134***	0.130***	0.104***	0.106***	0.105***
	(0.023)	(0.027)	(0.025)	(0.020)	(0.020)	(0.020)	(0.017)	(0.016)	(0.016)	(0.015)
Bandwidth	0.112	0.114	0.106	0.112	0.115	0.104	0.115	0.123	0.112	0.116

Notes: In Column (1), the dependent variable is an indicator equal to 1 if the candidate is elected as member of the US House of Representatives during the general election. In Columns (2), (5), and (8), the dependent variable is an indicator equal to 1 if the candidate runs for office in her party's primary elections held after 2, 4, and 6 years, respectively. In Columns (3), (6), and (9), the dependent variable is an indicator equal to 1 if the candidate wins her party's primary elections held after 2, 4, and 6 years, respectively. In Columns (4), (7), and (10), the dependent variable is an indicator equal to 1 if the candidate is elected to the US House of Representatives in the elections held after 2,4, and 6 years, respectively. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. In Panel (a), $h_{i,t}$ is the vote share distance between the candidate and the minimal vote share required to be eligible to compete in the second round, as in the baseline estimation. In Panel (b), the running variable $h_{i,t}$ is computed as in Panel (a) and the sample is restricted to candidates who either win the primary or rank as the runner-up. In Panel (c), $h_{i,t}$ is the distance between each candidate and the minimal vote share required to be eligible to compete in the second round, computed by adding to (removing from) each candidate a number of votes from (to) the total number of valid votes in the spirit of Fiva et al. (2018). See Appendix E for details. In Panel (d), $h_{i,t}$ is defined as in Panel (c), and each individual observation is weighted by $1/ncandidates$. In Panel (e), $h_{i,t}$ is the distance between each candidate and the minimal vote share required to be eligible to compete in the second round, computed by adding to (removing from) each candidate a number of votes from (to) other candidates, proportionally to their original vote share, in the spirit of Folke (2014) and Fiva et al. (2018). See Appendix E for details. Standard errors robust to clustering at the congressional district-by-party level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B19: Alternative running variables (Italian municipalities)

	(1)	(2)	(3)	(4)
	Mayor	Run (t+1)	Mayor (t+1)	Elected in Province
(a) Baseline running variable				
Female × Qualified	-0.096***	-0.036	-0.066***	-0.047**
	(0.032)	(0.043)	(0.019)	(0.019)
Qualified	0.232***	0.099***	0.061***	0.031**
	(0.021)	(0.030)	(0.018)	(0.015)
Bandwidth	0.115	0.113	0.115	0.175
(b) Only second and third recipient				
Female × Qualified	-0.081**	-0.041	-0.041**	-0.059*
	(0.033)	(0.052)	(0.019)	(0.033)
Qualified	0.245***	0.121***	0.063***	0.046*
	(0.023)	(0.033)	(0.018)	(0.024)
Bandwidth	0.100	0.102	0.132	0.0857
(c) Assign extra votes to turnout				
Female × Qualified	-0.090***	-0.030	-0.074***	-0.067**
	(0.026)	(0.045)	(0.020)	(0.027)
Qualified	0.255***	0.089***	0.055***	0.056***
	(0.015)	(0.031)	(0.020)	(0.021)
Bandwidth	0.271	0.102	0.0883	0.0919
(d) Weight by $1/ncandidates$				
Female × Qualified	-0.090***	-0.018	-0.067***	-0.056**
	(0.026)	(0.047)	(0.021)	(0.025)
Qualified	0.247***	0.068**	0.032	0.045**
	(0.015)	(0.033)	(0.020)	(0.021)
Bandwidth	0.271	0.102	0.0883	0.0919
(e) Assign extra votes to other candidates				
Female × Qualified	-0.100***	-0.041	-0.072***	-0.054**
	(0.031)	(0.044)	(0.018)	(0.023)
Qualified	0.252***	0.103***	0.071***	0.045**
	(0.019)	(0.030)	(0.019)	(0.018)
Bandwidth	0.106	0.0901	0.0927	0.0998

Notes: In Column (1), the dependent variable is an indicator equal to 1 if the candidate is elected as mayor during the second round. In Column (2), the dependent variable is a dummy equal to 1 if the candidate runs for office in the occasion of elections held in term $t+1$. In Column (3), the dependent variable is a dummy equal to 1 if the candidate is elected as mayor during elections held in term $t+1$. In Column (4), the dependent variable is a dummy equal to 1 if the candidate is in the future elected as a member of the Provincial Council or to the position of President of the Province. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. In Panel (a), $h_{i,t}$ is the vote share distance between the candidate and the minimal vote share required to be eligible to compete in the second round, as in the baseline estimation. In Panel (b), the running variable $h_{i,t}$ is computed as in Panel (a) and the sample is restricted to candidates who either rank as the runner-up or as the third-largest recipient during the second round. In Panel (c), $h_{i,t}$ is the distance between each candidate and the minimal vote share required to be eligible to compete in the second round, computed by adding to (removing from) each candidate a number of votes from (to) the total number of valid votes in the spirit of Fiva et al. (2018). See Appendix E for details. In Panel (d), $h_{i,t}$ is defined as in Panel (c), and each individual observation is weighted by $1/ncandidates$. In Panel (e), $h_{i,t}$ is the distance between each candidate and the minimal vote share required to be eligible to compete in the second round, computed by adding to (removing from) each candidate a number of votes from (to) other candidates, proportionally to their original vote share, in the spirit of Folke (2014) and Fiva et al. (2018). See Appendix E for details. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B20: Only mixed-gender elections (US House of Representatives)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Representative	Run	Nominated	Representative	Run	Nominated	Representative	Run	Nominated	Representative
		(t+1)	(t+1)	(t+1)	(t+2)	(t+2)	(t+2)	(t+3)	(t+3)	(t+3)
Female × Qualified	-0.170***	-0.013	-0.042	-0.073*	-0.084**	-0.086**	-0.067**	-0.057*	-0.059**	-0.053*
	(0.046)	(0.043)	(0.045)	(0.041)	(0.036)	(0.034)	(0.032)	(0.030)	(0.029)	(0.028)
Qualified	0.491***	0.205***	0.238***	0.252***	0.186***	0.190***	0.178***	0.114***	0.123***	0.122***
	(0.047)	(0.051)	(0.051)	(0.045)	(0.041)	(0.039)	(0.037)	(0.032)	(0.031)	(0.031)
Observations	1,228	1,380	1,140	1,190	1,280	1,467	1,577	1,773	1,878	1,909
R^2	0.295	0.129	0.136	0.129	0.093	0.108	0.104	0.087	0.090	0.093
Election Year Effects	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.145	0.203	0.160	0.111	0.103	0.0893	0.0761	0.0626	0.0596	0.0561
Bandwidth	0.139	0.157	0.129	0.135	0.145	0.169	0.182	0.203	0.215	0.219

Notes: In Column (1), the dependent variable is an indicator equal to 1 if the candidate is elected as member of the US House of Representatives during the general election. In Columns (2), (5), and (8), the dependent variable is an indicator equal to 1 if the candidate runs for office in her party's primary elections held after 2, 4, and 6 years, respectively. In Columns (3), (6), and (9), the dependent variable is an indicator equal to 1 if the candidate wins her party's primary elections held after 2, 4, and 6 years, respectively. In Columns (4), (7), and (10), the dependent variable is an indicator equal to 1 if the candidate is elected to the US House of Representatives in the elections held after 2,4, and 6 years, respectively. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. The sample is restricted to cases in which at least one female candidate and one male candidate compete in the same primary election. Standard errors robust to clustering at the congressional district-by-party level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B21: Only mixed-gender elections (Italian municipalities)

	(1)	(2)	(3)	(4)
	Mayor	Run (t+1)	Mayor (t+1)	Elected in Province
Female × Qualified	-0.102***	-0.042	-0.050**	-0.048**
	(0.031)	(0.039)	(0.020)	(0.023)
Qualified	0.238***	0.137***	0.070***	0.053**
	(0.024)	(0.034)	(0.021)	(0.024)
Observations	3,013	2,496	2,682	2,264
R^2	0.243	0.094	0.059	0.022
Election Year Effects	✓	✓	✓	✓
Mean Dep. Var.	0.123	0.220	0.0611	0.0680
Bandwidth	0.190	0.155	0.168	0.141

Notes: In Column (1), the dependent variable is an indicator equal to 1 if the candidate is elected as mayor during the second round. In Column (2), the dependent variable is a dummy equal to 1 if the candidate runs for office in the occasion of elections held in term $t + 1$. In Column (3), the dependent variable is a dummy equal to 1 if the candidate is elected as mayor during elections held in term $t + 1$. In Column (4), the dependent variable is a dummy equal to 1 if the candidate is in the future elected as a member of the Provincial Council or to the position of President of the Province. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. The sample is restricted to cases in which at least one female candidate and one male candidate compete in the same first-round election. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B22: Qualification to final round and probability of winning the election, by gender of the excluded candidate

(a) US House of Representatives

	(1)	(2)	(3)	(4)	(5)
	Representative	Representative	Representative	Representative	Representative
Female Excluded × Qualified	-0.001	0.007	-0.001	0.007	0.007
	(0.037)	(0.036)	(0.036)	(0.035)	(0.036)
Qualified	0.337***	0.333***	0.339***	0.335***	0.337***
	(0.026)	(0.026)	(0.026)	(0.026)	(0.028)
Observations	3,404	3,403	3,404	3,403	3,403
R^2	0.210	0.255	0.223	0.273	0.360
Individual Char.		✓		✓	✓
Primary Returns			✓	✓	✓
District Fixed Effects					✓
Election Year Effects	✓	✓	✓	✓	✓
Mean Dep. Var.	0.130	0.130	0.130	0.130	0.130
Bandwidth	0.138	0.138	0.138	0.138	0.138

(b) Italian municipalities

	(1)	(2)	(3)	(4)	(5)
	Mayor	Mayor	Mayor	Mayor	Mayor
Female Excluded × Qualified	-0.013	-0.013	-0.015	-0.015	-0.048
	(0.030)	(0.030)	(0.027)	(0.027)	(0.045)
Qualified	0.224***	0.223***	0.224***	0.224***	0.240***
	(0.021)	(0.021)	(0.020)	(0.020)	(0.027)
Observations	3,560	3,560	3,560	3,560	3,560
R^2	0.200	0.208	0.275	0.283	0.504
Individual Char.		✓		✓	✓
First round Returns			✓	✓	✓
Election Fixed Effects					✓
Election Year Effects	✓	✓	✓	✓	
Mean Dep. Var.	0.129	0.129	0.129	0.129	0.129
Bandwidth	0.115	0.115	0.115	0.115	0.115

Notes: Panel (a): The dependent variable is a dummy equal to 1 if the candidate is elected as member of the House of Representatives during the general election. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable and election year dummies. Column (2) includes controls for *Incumbent, Candidacy counter, Nomination counter,* and *Democratic party*. Column (3) includes controls for *Vote share (Primary)*. Column (4) includes the sets of controls in (2) and (3). Column (5) includes the set of controls in (4) and congressional district fixed effects. Panel (b): The dependent variable is a dummy equal to 1 if the candidate is elected as mayor during the second round. All specifications include the interaction terms between the Qualified dummy and the function of the assignment variable. Column (2) includes controls for *Age, College, High-Skill job, Incumbent, Candidacy counter, Left-wing, Right-wing*, and *Independent*. Column (3) includes controls for *Vote share (First round), Vote share coalition (First round),* and *No. parties in coalition (First round)*. Columns (4) and (5) jointly include the sets of controls in (2) and (3). Columns (1)–(5) include election year dummies, while Column (6) includes individual election fixed effects. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth, where $F_i \times Qualified_{i,t}$ has been replaced by an indicator equal to 1 if the closest candidate who does not qualify to the final round is a female, and zero otherwise. Standard errors robust to clustering at the congressional district-by-party level (Panel (a)) and municipality level (Panel (b)) are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B23: Alternative RD specification

(a) US House of Representatives

	(1)	(2)	(3)	(4)
	Female's party wins general election	Female's party wins general election	Female's party wins general election	Female's party wins general election
Female Qualified	-0.122	-0.101	-0.120	-0.189*
	(0.090)	(0.094)	(0.090)	(0.100)
Observations	449	449	449	449
R^2	0.064	0.180	0.065	0.303
Election Year Effects	✓	✓	✓	✓
State Effects		✓		
Party Effects			✓	
State × Party Effects				✓
Mean Dep. Var.	0.307	0.307	0.307	0.307
Bandwidth	0.0815	0.0815	0.0815	0.0815

(b) Italian municipalities

	(1)	(2)	(3)	(4)
	Runner-up wins second round	Runner-up wins second round	Runner-up wins second round	Runner-up wins second round
Female Qualified	-0.129	-0.127	-0.233*	-0.246*
	(0.109)	(0.107)	(0.131)	(0.138)
Observations	311	310	311	310
R^2	0.096	0.116	0.279	0.294
Province Effects			✓	✓
Election Year Effects	✓	✓	✓	✓
Covariates		✓		✓
Mean Dep. Var.	0.264	0.265	0.264	0.265
Bandwidth	0.129	0.129	0.129	0.129

Notes: Panel (a): The dependent variable is a dummy equal to 1 if the female candidate's party wins the seat in the US House of Representatives. All specifications include the interaction terms between the Female Qualified dummy and the function of the assignment variable and election year dummies. Column (2) includes state fixed effects. Column (3) includes party fixed effects. Column (4) includes state-by-party fixed effects. Panel (b): The dependent variable is a dummy equal to 1 if the runner-up is elected as mayor during the second round. All specifications include the interaction terms between the runoff dummy and the function of the assignment variable. Column (2) includes controls for the individual characteristics of the incumbent mayor (*Female, College, Age, Occupation*). Column (3) includes province fixed effects. Column (4) includes province fixed effects and the set of covariates included in Column (2). Estimation methods: local-linear regression as in equation (12) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the congressional district-by-party level (Panel (a)) and municipality level (Panel (b)) are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table B24: Placebo final round and future career outcomes, by gender

	(a) Plurality			(b) Top candidate above 50%		
	(1)	(2)	(3)	(1)	(2)	(3)
	Run (t+1)	Mayor (t+1)	Elected in Province	Run (t+1)	Mayor (t+1)	Elected in Province
Female × Placebo	-0.030	-0.018	-0.004	0.028	-0.000	-0.020
	(0.019)	(0.011)	(0.007)	(0.052)	(0.009)	(0.025)
Placebo	0.030**	-0.008	0.005	0.006	-0.010	-0.010
	(0.015)	(0.009)	(0.006)	(0.044)	(0.015)	(0.023)
Observations	12,230	11,781	14,607	1,488	1,214	1,713
R^2	0.063	0.034	0.009	0.054	0.033	0.023
Election Year Effects	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	0.222	0.0708	0.0304	0.145	0.0148	0.0432
Bandwidth	0.0855	0.0824	0.102	0.135	0.117	0.150

Notes: All mayoral candidates in the municipalities belonging to the ordinary regions running for election between 1993 and 2019 under the plurality system (Columns (1)–(3)) and under the runoff system (Columns (4)–(6)). In Columns (4)–(6), the sample is restricted to cases in which the candidate with largest vote share has been elected mayor during the first round. Data are at the term level. In Columns (1) and (4), the dependent variable is a dummy equal to 1 if the candidate runs for office in the occasion of elections held in term $t + 1$. In Columns (2) and (5), the dependent variable is a dummy equal to 1 if the candidate is elected as mayor during elections held in term $t + 1$. In Columns (3) and (6), the dependent variable is a dummy equal to 1 if the candidate is in the future elected as a member of the Provincial Council or to the position of President of the Province. All specifications include the interaction terms between the Placebo dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

C Simulation

In Section 2, I introduce how two-stages elections allow us to account for sample selection bias in the estimation of gender gaps in returns from running in an election. In this section, I perform a simulation exercise to highlight this approach's improvements compared to standard linear models using a self-selected sample of individuals. Consistent with Section 2, I consider the case of returns from participating in an electoral competition even if the results can be generalized to other circumstances in which selection is based on a threshold of an observed variable.

C.1 Probability of winning the election

Suppose equation (1) and equation (2) take the form

$$Y_i = -0.5F_i + u_i, \tag{C1}$$

and

$$z_i = \mathbf{1}\{\alpha F_i + v_i > 0\}, \tag{C2}$$

where F_i is a random uniform integer equal to 0 or to 1, $u_i = v_i + \varepsilon_i$, $v_i \sim \mathcal{N}(0, 1)$, $\varepsilon_i \sim \mathcal{N}(0, 1)$, and α is random one-decimal number in the interval $[0, 1]$.

In each replication, I generate a dataset of N=100,000 observations of this form, and estimate equation (3) in the sample of observations for which $z_i = 1$ as well as equation (4). The exercise is repeated for a total of 100,000 replications, in each of which α varies at random.

Figure C1 summarizes the results. The figure reports the distribution of estimated gaps according to equation (3) and equation (4) for different values of α. The black and white histograms summarize the distribution of gaps estimated using equation (4), while the gray histogram estimates the gender gap in returns using equation (3). Black and white histograms are centered around $\gamma = -0.5$ for any value of α. In contrast, the gray histogram identifies $\gamma = -0.5$ only when $\alpha = 0$ (i.e., only when the selection into a profession is equal for individuals of both genders in the entire population of potential candidates). In male-dominated environments, where $\alpha < 0$, not accounting for selection would hide the presence of gaps in returns, as the estimated gap from equation (3) approaches zero as α decreases.

The focus on candidates marginally admitted to the final round, who cannot perfectly predict that they will be eligible for it, also helps if selection into the first round is not random. In Figure C2, I show that the gender gap in returns estimated using equation (4) is a consistent estimator of γ despite a reduction of precision as long as the unobserved selection process is uncorrelated with the gender of the candidate in a narrow window around $h_i = 0$ (i.e., as long as Assumption I holds). In contrary, the estimated gender gap based on equation (3) would be even more attenuated than in Figure C1 for any $\alpha < 0$.

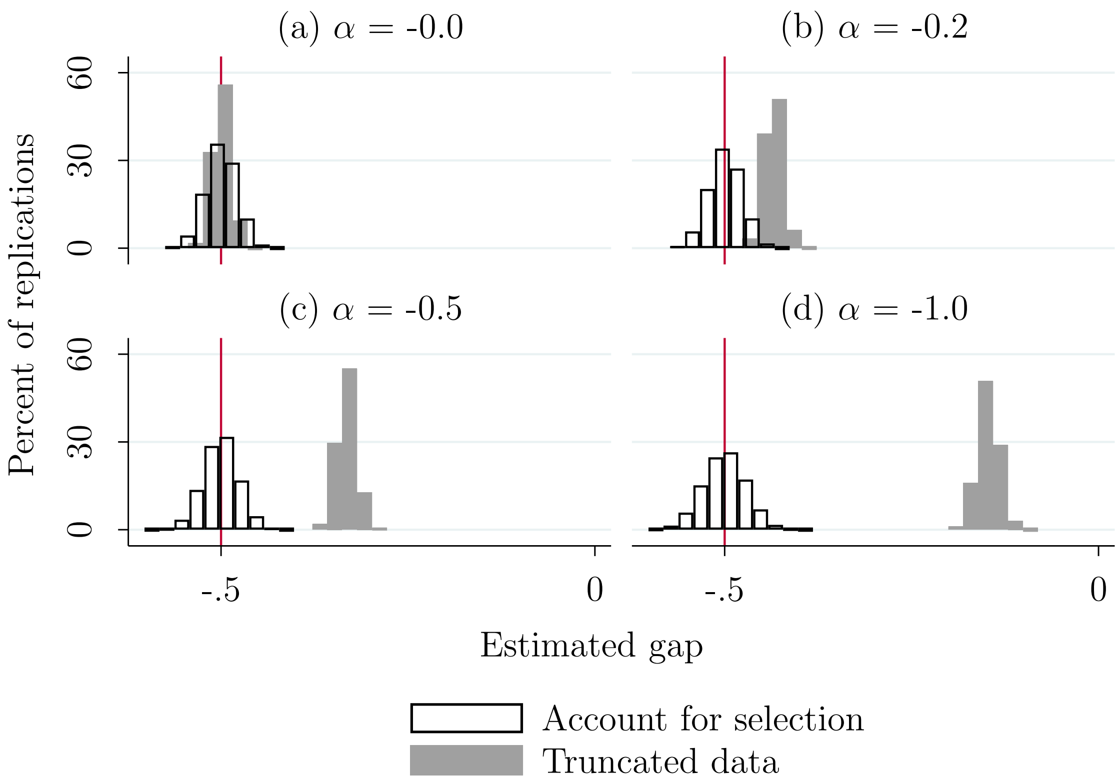

Figure C1: Distribution of estimated gender gaps using equation (3) and equation (4)

Notes: Distribution of estimated gender gaps from estimating equation (3) (gray shaded histogram) and equation (4) (black and white histogram) based on 100,000 replications for different values of α in equation (2). In each replication, N=100,000 observations are generated according to equation (C1) and equation (C2).

C.2 Future career outcomes

Suppose equation (5) takes the form

$$Y_i = \alpha F_i + \mathbf{1}\{\alpha F_i + v_i > 0\}\big(-0.5F_i\big) + u_i, \qquad \text{(C3)}$$

while z_i follows equation (C2). Consistently with equation (C1), F_i is a random uniform integer equal to 0 or to 1, $v_i \sim \mathcal{N}(0,1)$, $\varepsilon_i \sim \mathcal{N}(0,1)$, and α is a random one-decimal number in the interval $[0,1]$.

In each replication, I generate a dataset of N=100,000 observations of this form, and estimate equation (6) in the sample of observations for which $z_i = 1$ and the Difference-in-discontinuities in equation (7). For a sake of comparison, I also show the results from estimating equation (6) in the case in which selection is assigned based on $z_i = \mathbf{1}\{h_i > 0\}$. The exercise is repeated for a total of 100,000 replications, in each of which α varies at random.

Figure C3 summarizes the results. The figure reports the distribution of estimated gaps estimated according to equations (3), (4) and (7) for different values of α. The dark gray histograms summarize the distribution of gaps estimated using equation (4), while the light gray histograms estimate the gender gap in returns using equation (3). Lastly,

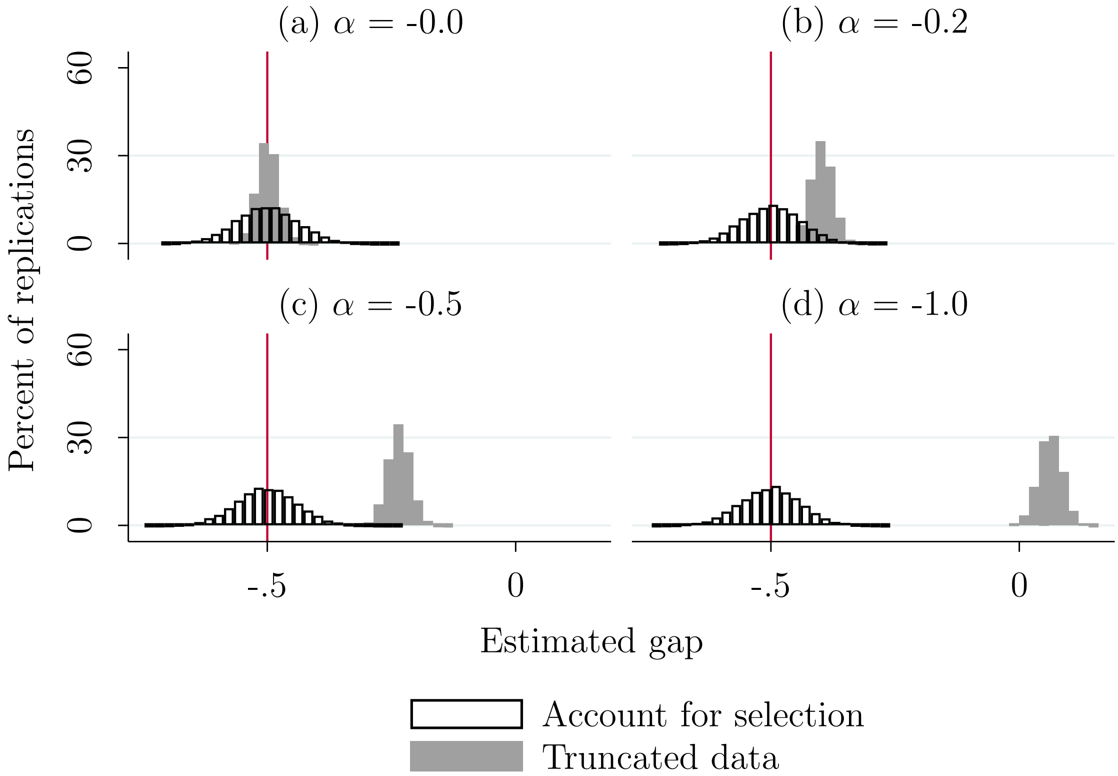

Figure C2: Distribution of estimated gender gaps using equation (3) and equation (4) when selection into the first round is unobserved and locally uncorrelated with gender

Notes: Distribution of estimated gender gaps from estimating equation (3) (gray shaded histogram) and equation (4) (black and white histogram) based on 100,000 replications for different values of α in equation (2). In each replication, N=100,000 observations are generated according to (C1) and (C2), and the coefficients are estimated only a subsample of candidates based on expected returns but not on gender.

the black and white histograms show the distribution of gaps according to equation (7). Black and white histograms are centered around $\gamma^c = -0.5$ for any value of α, while the gray histograms identify $\gamma^c = -0.5$ only if $\alpha = 0$. In male-dominated environments, sample selection bias leads to an overestimation of gender gaps in future career outcomes since α enters both z_i and Y_i. Accounting for z_i by exploiting two-stages elections but not correcting for the direct impact of α on Y_i in equation (C3) would result in a substantial overestimation of career gender gaps.

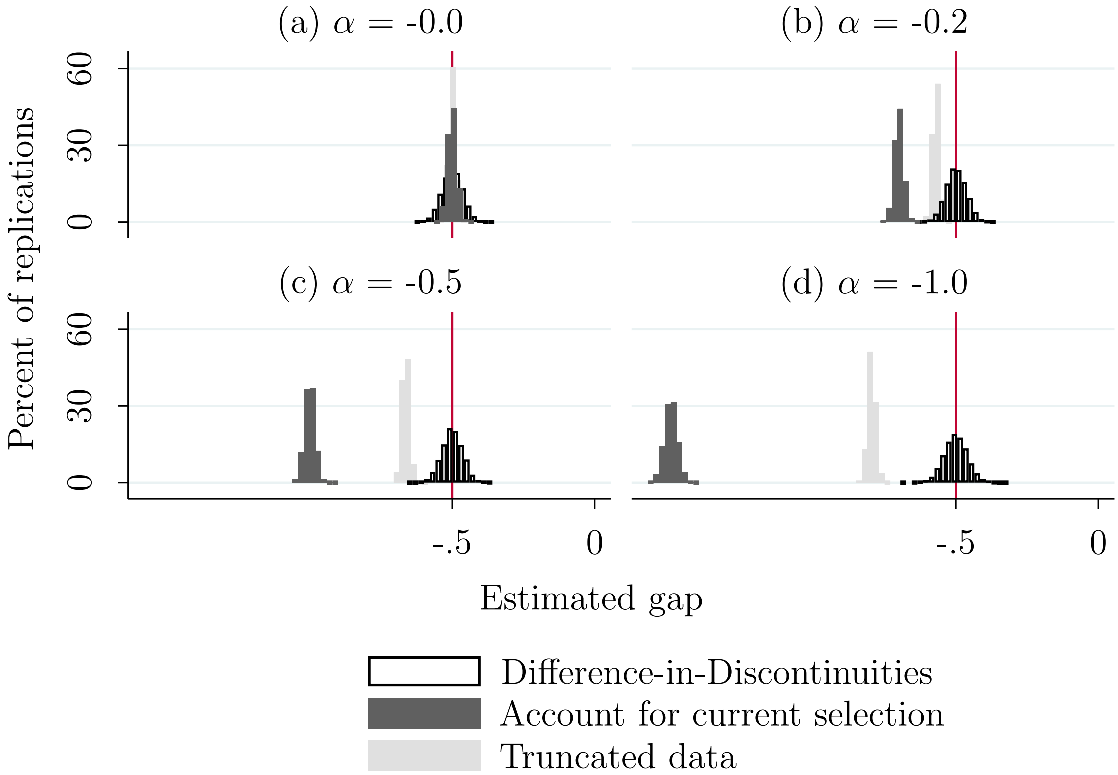

Figure C3: Distribution of estimated career gender gaps using equation (6) and equation (7)

Notes: Distribution of estimated gender gaps from estimating equation (6) (dark gray shaded histogram) and equation (7) (black and white histogram) based on 100,000 replications for different values of α in equation (2). The light gray shaded histogram reports the distribution of estimated gaps by using a version of equation (6) in which $z_i = \mathbf{1}\{h_i > 0\}$. In each replication, N=100,000 observations are generated according to equation (C3) and equation (C2).

D Data Appendix

D.1 List of data sources

Primary election returns for the US House of Representatives (Pettigrew et al., 2014). The dataset covers 38,559 winning and losing candidates running in primary elections for the US House of Representatives between 1956 and 2010 and includes information on the gender of the candidate. Figure D1 shows the geographical distribution of primaries covered in the dataset, highlighting that some districts are included in the data more regularly than others. The coverage is satisfactory, as the data refers to more than 95 percent of the seats assigned between 1976 and 2010. I restrict the sample to candidates competing in primary elections organized by either the Democratic party or by the Republican party. For each congressional seat, I observe two separate primary elections. The data allow tracking the same individual candidate over time after changes in boundary and labeling of a congressional district.

General election returns for the US House of Representatives. The Massachussets Institute of Technology *Election Data + Science Lab* covers the general election returns of all congressional elections held between 1956 and 2018. In total, it contains 29,636 observations from 14,818 congressional races. Data from elections held after 2010 are used to track individual candidates over time. I drop 210 uncontested elections.

Federal Election Commission data. The dataset covers information on campaigning financing and expenditures of winning and losing candidates running for office between 1980 and 2010. I use individual-level data on total contributions and spending and contributions made by the candidate to her campaign, loans subscribed by the candidate, and contributions from individual donors and the candidate's political party. Information is available both for primary winners and losers.

Italian municipalities election returns. This dataset, released by the Italian Ministry of Internal Affairs and publicly available, covers candidate-level returns of all municipality elections held since the introduction of the mayors' direct election in 1993. For municipalities in which a second round takes place, the dataset includes data on endorsements between the two rounds and returns of the second round.

Anagrafe degli Amministratori Locali. This dataset, released by the Italian Ministry of Internal Affairs and publicly available, contains all elected officials' individual-level background characteristics at the local level since 1986. I observe both elected mayors and losing candidates who are elected into the municipality council. If a candidate is not included in the dataset after a given election year, I can retrieve her background characteristics based on the latest available information.

Dati rilevazione semestrale del corpo elettorale. This dataset, released by the Italian Ministry of Internal Affairs and publicly available, contains biannual information on the number of men and women eligible to vote in each municipality.

ISTAT Data. I add municipality-level yearly socioeconomic and demographic covariates from the Italian Institute of Statistics (ISTAT) *Atlante Statistico dei Comuni*. I also ob-

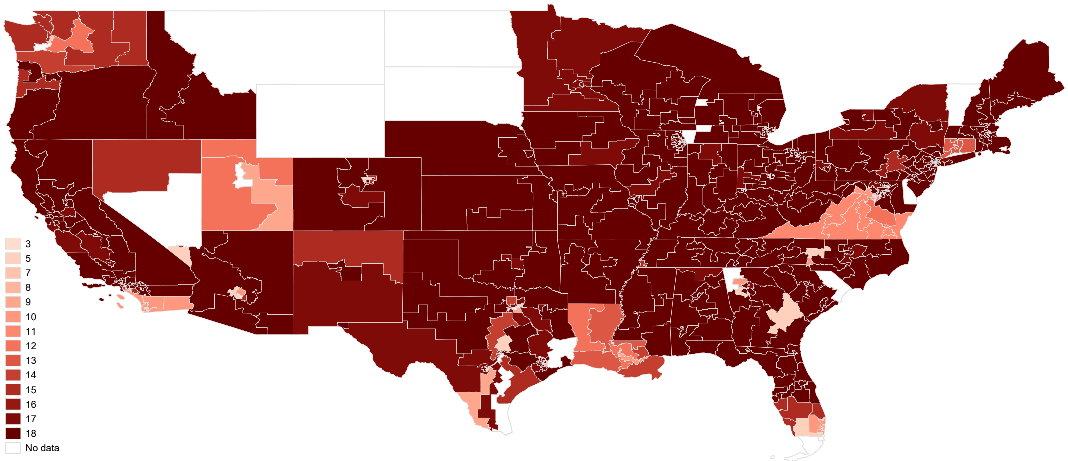

Figure D1: US House of Representatives (1976–2010) data availability

This map reports the number of primary elections included in Pettigrew et al. (2014) dataset per each congressional district. Darker area represent the district observed more often, while lighter areas represent districts observed fewer times. White areas represent districts never included in the dataset. Source: author's elaboration on Pettigrew et al. (2014) dataset and on the US Census Bureau geodata with reference to the 116th Congress.

serve population censuses held in 1991, 2001, and 2011 to retrieve the official population used to assign municipalities to either voting system.

FACTIVA Data. I hand-collect on the Factiva portal the number of mentions on local newspapers covered in the archive for each candidate who marginally qualify to the final round of a runoff election. For simplicity, I focus on candidates who lie within the Calonico et al. (2014) optimal bandwidth for the main dependent variable $Mayor_{i,t}$. I restrict the search to articles published between January 1st of the election year and the date in which the second round took place and looked for the candidate's name and surname in all articles written in Italian.

D.2 Final datasets

D.2.1 US House of Representatives

The final dataset covers 24,080 candidates competing in 13,774 primaries held between 1976 and 2010 for 7,478 seats in the US House of Representatives, merged based on congressional district, party, and election year with general election returns. I can merge based on surname, congressional district, party, and election-year 8,379 winning and losing primary candidates competing between 1980 and 2010 with the Federal Election Commission data on campaign financing and expenditures.

D.2.2 Italian municipalities

The final dataset covers 19,037 candidates running for the position of mayors in municipalities with more than 15,000 residents according to the latest available population census. I retrieve background characteristics of winning and losing candidates which include education level, occupation, age, and party affiliation, based on a perfect match on the town, election year, position in the administration, name, and surname. Moreover, the match allows following career paths in the municipality or higher administrative hierarchy levels.[41] I construct a measure of voter turnout by gender combining the gender distribution of the voting population with information from the gender distribution of the population eligible to vote. I add additional municipality-level covariates from the population censuses held in 1991, 2001, and 2011 and from the *Atlante statistico dei comuni*, released by the ISTAT.

[41] I observe these additional covariates only for 80 percent of all candidates, positively selected in terms of vote share. Mismatches are primarily caused by mayoral candidates receiving less than 3 percent of the votes, who do not secure a seat into the municipality council.

E Alternative running variables

In this paper's research design, the running variable is the distance between the vote share received by a candidate and the minimum share of total votes required to qualify for the final round. Both for Italian municipalities, where the identifying variation comes from races in which at least three candidates compete for office, and the US House of Representatives, it is likely that more than two candidates compete in the first round of voting.

When more than two candidates compete for office, the minimum share of total votes required to be eligible to compete in the second round is unobservable. It depends on the number of candidates and the identity of candidates from (to) whom extra votes would be removed (added) in a counterfactual election. Consistently with Carozzi et al. (2020) and the Online Appendix therein, I estimate all the results in this paper using three definitions of $h_{i,t}$.

E.1 Vote Share distance

The first approach is the most intuitive. Yet it assumes that, in a counterfactual election, votes would only move between candidate i and the closest candidate to the threshold on the opposite side of it.

According to this approach, the running variable is defined as follows:

$$h_{i,t} = \begin{cases} VoteShare_{i,t} - VoteShare_{A,t} < 0 & \text{if } VoteShare_{i,t} < VoteShare_{A,t} \\ VoteShare_{i,t} - VoteShare_{B,t} > 0 & \text{if } VoteShare_{i,t} > VoteShare_{B,t} \end{cases}, \tag{E4}$$

where $VoteShare_{A,t}$ is the share of valid votes of the candidate closest to the threshold and on the right of it, and $VoteShare_{B,t}$ is the share of valid votes of the candidate closest to the threshold and on its left.

The same running variable can also be applied to cases in which the focus is restricted to only the two closest candidates to the threshold, one on its left and one on its right.[42] This is the definition of $h_{i,t}$ employed to estimate the main results, as described in Section 4.

E.2 Assignment of extra (missing) votes to (from) changes in turnout

The second approach assumes that, in a counterfactual election, the aggregate number of valid votes would increase or decrease by an amount equal to the number of votes that candidate i needs to gain or to lose to jump on the opposite side of the threshold.

Consider this example. Suppose candidates receive respectively 1,400, 800, and 500 votes. The total number of valid votes is thus 2,700. The aim is to calculate the distance between the candidate with fewer votes and the runner-up in a counterfactual election in which the candidate receives 300 extra votes. The (counterfactual) vote distribution is 1,400, 800, 800, while the total number of valid votes is 3,000. The running variable for the candidate with fewer votes would be equal to $-\frac{300}{3000} = -0.1$. In this case, by construction,

[42] This approach has been implemented in cases of runoff system in Pons and Tricaud (2019). Folke et al. (2016) restrict the sample to candidates around the threshold and elections in which only two candidates are close to the threshold.

two adjacent candidates on opposite sides of the threshold would have opposite values of the running variable.

Formally,

$$h_{i,t} = \frac{Votes_{i,t} + \Delta}{TotalVotes_{i,t} + \Delta}, \tag{E5}$$

where $TotalVotes_t$ is the total number of valid votes cast in the election, and Δ is defined as follows:

$$\Delta = \begin{cases} Votes_{i,t} - Votes_{A,t} < 0 & \text{if } VoteShare_{i,t} < VoteShare_{A,t} \\ Votes_{i,t} - Votes_{B,t} > 0 & \text{if } VoteShare_{i,t} > VoteShare_{B,t} \end{cases}. \tag{E6}$$

Consistently with the definition presented in equation (E4), $Votes_{A,t}$ and $VoteShare_{A,t}$ are the number of votes and the share of valid votes received by the candidate closest to the threshold and on the right of it while $Votes_{B,t}$ and $VoteShare_{B,t}$ are the number of votes and the share of valid votes received by the candidate closest to the threshold and on its left.

As compared to equation (E4), the running variable defined in equation (E5) takes into account that candidates in small constituencies have a higher propensity to lie far from the threshold and that this propensity is attenuated if relatively more candidates compete in large constituencies.[43]

E.3 Assignment of extra (missing) votes to (from) other candidates

The third approach assumes that a candidate below the threshold would receive extra votes from other candidates proportionally to their share of votes. Conversely, a candidate above the threshold would lose votes to each other candidates' advantage, proportionally to her share of valid votes.

In this case, the number of votes to re-assign in a counterfactual election is lower than 300. Indeed, the actual runner-up would lose her votes as long as the candidate with fewer votes receive more. I use an iterative numerical approximation to calculate the running variable according to this method. In each iteration, I assign 0.001 percent of valid votes from the reference candidate to each other or vice-versa. Then, I define the threshold as the number of re-assigned votes necessary to obtain a qualification status change. Defining such threshold, in terms of share of valid votes, as $\widetilde{\Delta}$, the running variable takes the following form:

$$h_{i,t} = \begin{cases} VoteShare_{i,t} - \widetilde{\Delta} < 0 & \text{if } VoteShare_{i,t} < \widetilde{\Delta} \\ VoteShare_{i,t} - \widetilde{\Delta} > 0 & \text{if } VoteShare_{i,t} > \widetilde{\Delta} \end{cases}. \tag{E7}$$

Equation E7 defines the most general running variable, as it relaxes rank dependency. More specifically, the counterfactual vote allocation ranking can vary compared to the observed vote allocation for candidates away from the threshold. The definition of $\widetilde{\Delta}$ im-

[43] In the extreme case of a constituency of three voters where two candidates run, the running variable calculated according to equation (E5) would take either the value 0.2 or -0.2, as compared to 0.33 or -0.33 in the case the running variable is calculated according to equation (E4).

plies that the domain of $h_{i,t}$ is half the domains of running variables calculated with other methods since the votes re-assigned in the counterfactual allocation are simultaneously added to some candidates and removed from others.

E.4 Cross-validation

Figures E1, E2, and E3 show the correlation between the three running variables for each of the main specifications used in the paper, as in equation (8) and equation (11). More specifically, the horizontal axes represent the values of $h_{i,t}$ used in the main analysis of this paper and described in Section E.1, while each vertical axis reports the corresponding values of an alternative running variable calculated according to the approaches described in Section E.2 and E.3. The figures confirm that the definition of the running variable is crucial to ensure the fulfillment of the continuity of potential outcomes and the conditional randomization of the treatment around the threshold. Table E1 shows that the correlation coefficient between alternative running variables is always above 0.9 even if the running variable calculated according to equation (E5) follows a non-linear relationship with respect to the one defined in equation (E4), and the running variable defined in equation (E7) is noisier than the other two alternatives.

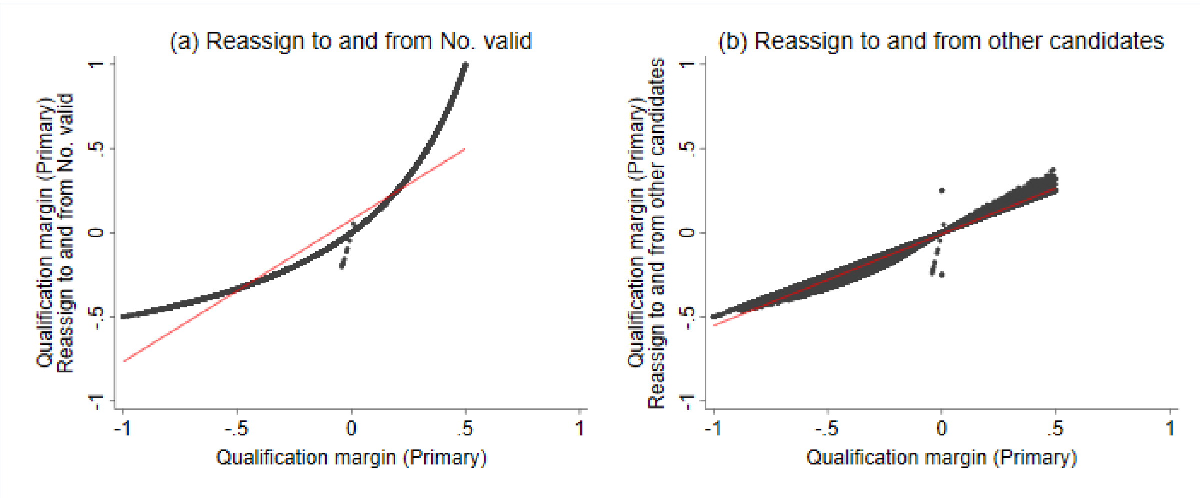

Figure E1: Correlation between different running variables (US House of Representatives)

Notes: Panel (a) reports the correlation between the running variable defined in equation (E4) (horizontal axis) and the running variable defined in equation (E5) (vertical axis) used to estimate equation (8). Panel (b) reports the correlation between the running variable defined in equation (E4) (horizontal axis) and the running variable defined in equation (E5) (vertical axis) used to estimate equation (8). Each scatter represents a unique observation. The red line represents the best linear fit. All domains have been restricted to take values between −1 and 1.

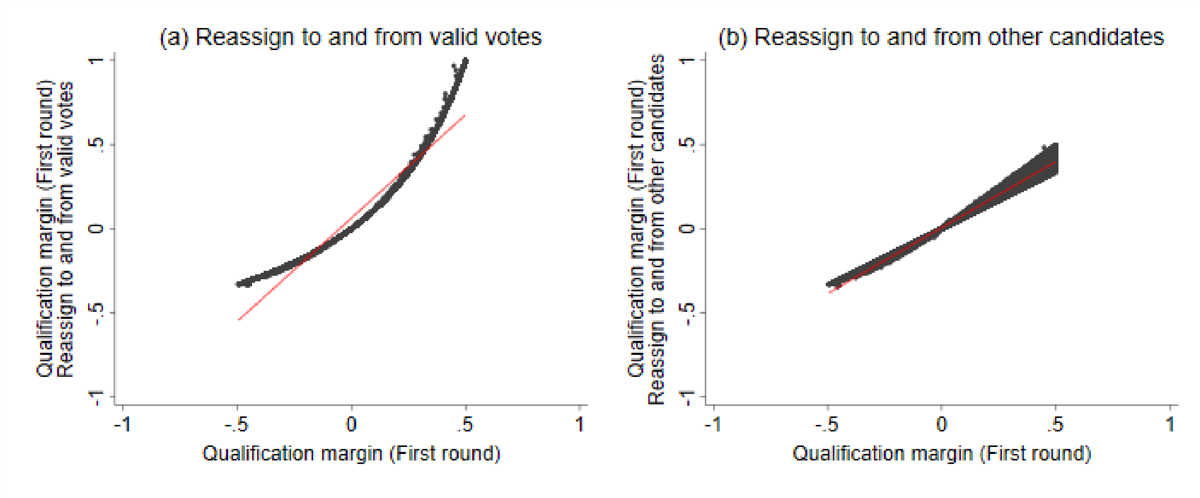

Figure E2: Correlation between different running variables (Italian municipalities)

Notes: Panel (a) reports the correlation between the running variable defined in equation (E4) (horizontal axis) and the running variable defined in equation (E5) (vertical axis) used to estimate equation (8). Panel (b) reports the correlation between the running variable defined in equation (E4) (horizontal axis) and the running variable defined in equation (E5) (vertical axis) used to estimate equation (8). Each scatter represents a unique observation. The red line represents the best linear fit. All domains have been restricted to take values between -1 and 1.

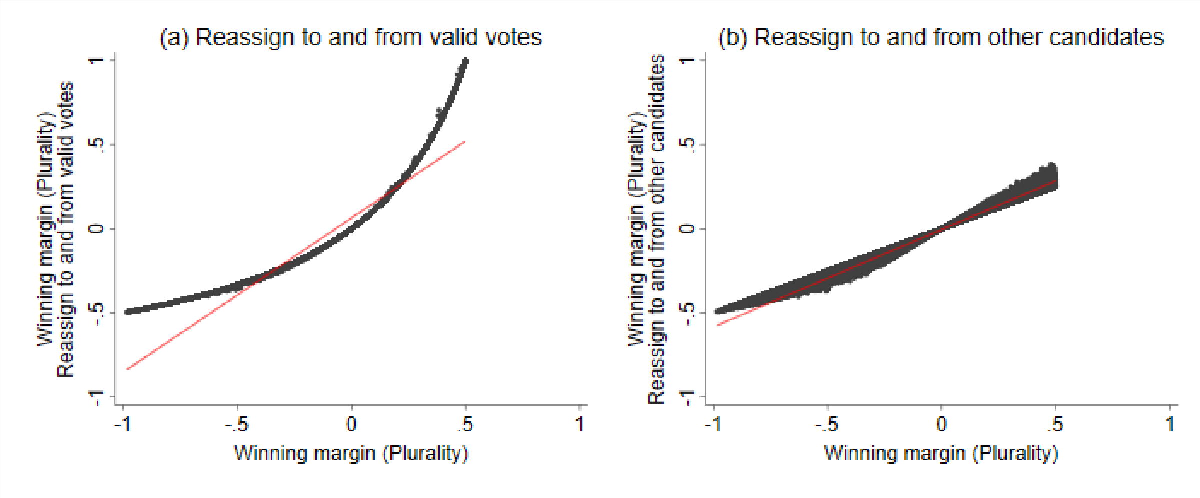

Figure E3: Correlation between different running variables (Plurality)

Notes: Panel (a) reports the correlation between the running variable defined in equation (E4) (horizontal axis) and the running variable defined in equation (E5) (vertical axis) used to estimate equation (8). Panel (b) reports the correlation between the running variable defined in equation (E4) (horizontal axis) and the running variable defined in equation (E5) (vertical axis) used to estimate equation (8). Each scatter represents a unique observation. The red line represents the best linear fit. All domains have been restricted to take values between −1 and 1.

Table E1: Correlation matrix of running variables

(a) US House of Representatives

	Baseline running variable	Reassign to and from valid votes	Reassign to and from other candidates
Baseline running variable	1.00	0.94	0.99
Reassign to and from valid votes	0.94	1.00	0.94
Reassign to and from other candidates	0.99	0.94	1.00

(b) Italian municipalities (Runoff)

	Baseline running variable	Reassign to and from valid votes	Reassign to and from other candidates
Baseline running variable	1.00	0.96	1.00
Reassign to and from valid votes	0.96	1.00	0.97
Reassign to and from other candidates	1.00	0.97	1.00

(c) Italian municipalities (Plurality)

	Baseline running variable	Reassign to and from valid votes	Reassign to and from other candidates
Baseline running variable	1.00	0.95	0.99
Reassign to and from valid votes	0.95	1.00	0.94
Reassign to and from other candidates	0.99	0.94	1.00

Notes: Panel (a) reports the correlation coefficients between each of the alternative running variables introduced in Section E and used to estimate equation (8). Sample: US House of Representatives. Panel (b) reports the correlation coefficients between each of the alternative running variables introduced in Section E and used to estimate equation (8). Sample: Italian municipalities under the runoff system. Panel (c) reports the correlation coefficients between each of the alternative running variables introduced in Section E and used to estimate equation (11). Sample: Italian municipality under the plurality system. All domains have been restricted to take values between -1 and 1.

F Additional empirical results

F.1 RD diagnostics for equation (11)

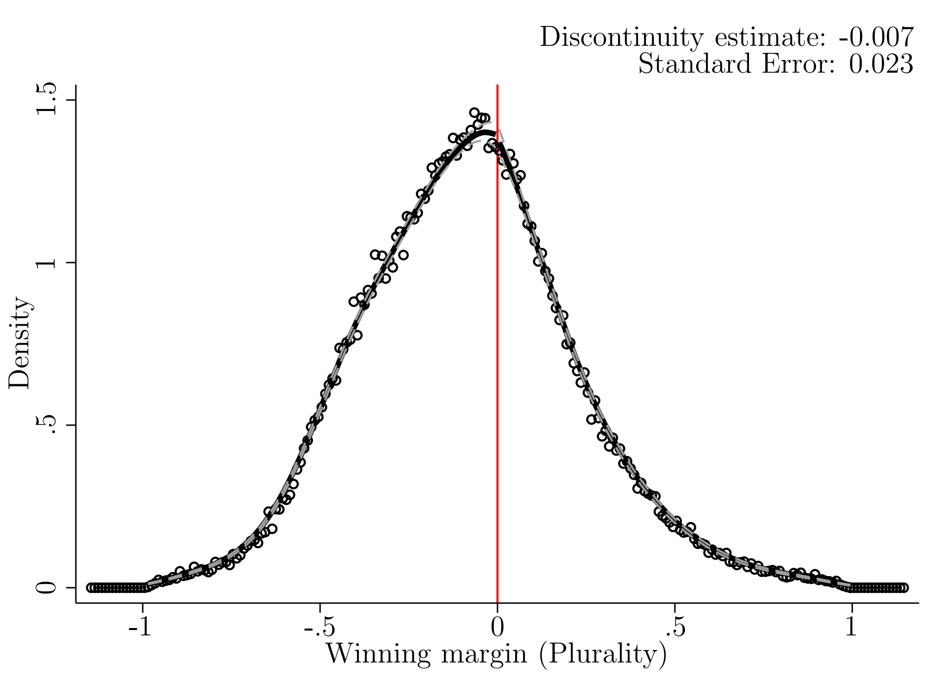

Figure F1: McCrary (2008) test

Notes: Log-density discontinuity and standard errors are computed performing a formal McCrary (2008) with optimal bandwidth. Markers represent sample averages within bins of the running variable $WinningMargin_{i,t}$ equal to 0.01.

Table F1: Balance of pre-determined characteristics

	(1)	(2)	(3)
Variable	Coefficient	St. Error	Obs.
Female candidate	-0.006	(0.018)	41,688
College	0.040*	(0.023)	27,736
High-skill job	0.016	(0.022)	30,497
Age	0.002	(0.020)	34,332
Incumbent	-0.032	(0.022)	34,074
Candidacy counter	0.008	(0.021)	32,184
Run again	0.013	(0.020)	34,982
Vote share (First round)	0.000	(0.003)	22,461
Vote share top candidate (First round)	0.000	(0.004)	22,461
Vote share runner-Up (First round)	0.014***	(0.005)	19,160
Turnout (First round)	-0.003	(0.003)	29,353
Male turnout (First round)	-0.006	(0.004)	22,617
Female turnout (First round)	-0.005	(0.004)	23,374
Voters (First round)	-0.005	(0.005)	25,156
No. candidates	-0.002	(0.006)	22,253
No. of lists competing	-0.002	(0.006)	22,253
Non-valid votes (First round)	-0.001	(0.002)	31,869
Incumbent runs for office	-0.001	(0.004)	32,125
Census population	-0.004	(0.005)	25,290
Surface (km^2)	-0.006	(0.004)	28,570
Elderly (%)	-0.000	(0.004)	26,374
Migrants (%)	0.003	(0.003)	29,748
Students (%)	0.001	(0.004)	27,881
Unemployment (%)	0.002	(0.004)	41,511
South	-0.003	(0.005)	36,696
Vote share of the left in last national election	-0.003	(0.003)	25,680
P(Mayor elected during first round)	-0.005	(0.004)	18,417

Notes: The dependent variable is specified in each row. All specifications include the interaction terms between the Mayor dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (11), where γ_1 and γ_3 are restricted to be equal 0 with uniform kernel and Calonico et al. (2014) optimal bandwidth. Each variable reported in the table has been previously standardized to have mean equal to 0 and variance equal to 1. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

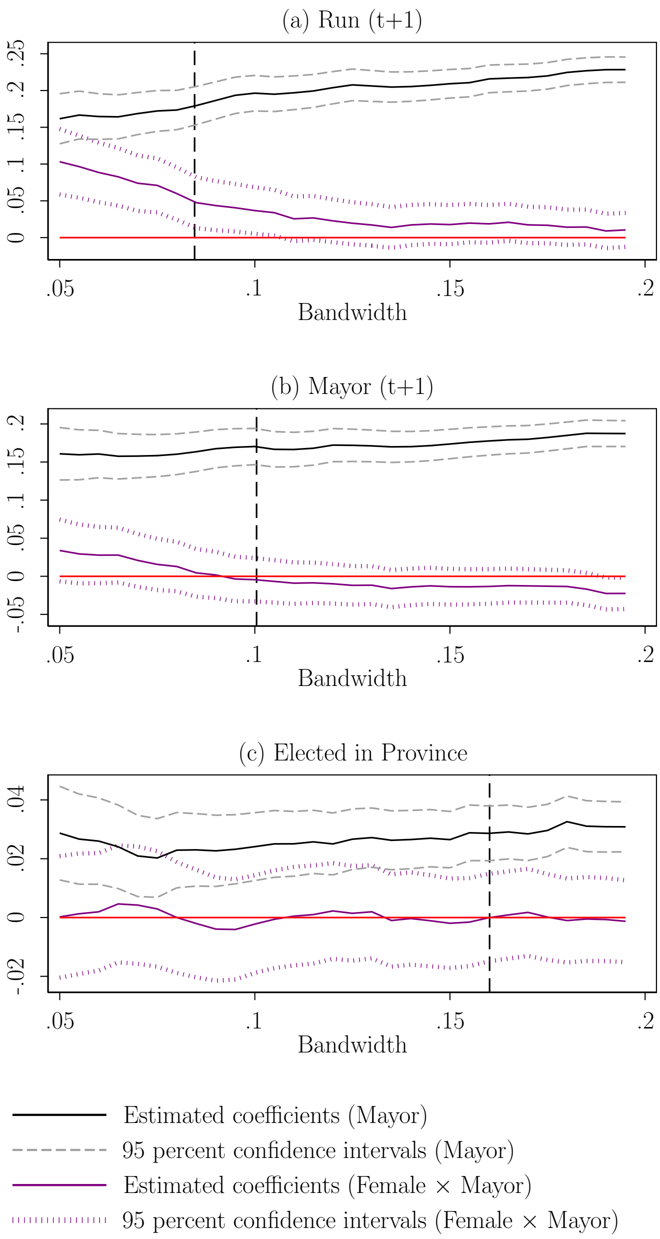

Figure F2: Bandwidth sensitivity

Notes: In Panel (a), the dependent variable is a dummy equal to 1 if the candidate runs for office in the occasion of elections held in term $t+1$. In Panel (b), the dependent variable is a dummy equal to 1 if the candidate is elected as mayor during elections held in term $t+1$. In Panel (c), the dependent variable is a dummy equal to 1 if the candidate is in the future elected as a member of the Provincial Council or to the position of President of the Province. The horizontal axis represents the bandwidths used to fit the local linear regression. The black solid line represents the estimated coefficients for γ_2 as a function of the chosen bandwidth, while the purple solid line represents the estimated coefficients for γ_1 as a function of the chosen bandwidth. Dashed lines represent the 95 percent confidence intervals of each coefficient. Estimation method: local linear regression with uniform kernel as in equation (11), and bandwidths ranging from $WinningMargin_{i,t} = 0.05$ to $WinningMargin_{i,t} = 0.20$. The vertical line represents the Calonico et al. (2014) optimal bandwidth. All specifications include the interaction terms between the Mayor dummy and the assignment variable and election year dummies. Each estimation concerns a variation of bandwidth equal to 0.005. 95% confidence intervals are based on standard errors robust to clustering at the municipality level.

Table F2: Alternative running variables

	(1)	(2)	(3)
	Run (t+1)	Mayor (t+1)	Elected in Province
(a) Baseline running variable			
Female × Mayor	0.083***	0.013	0.002
	(0.020)	(0.016)	(0.009)
Mayor	0.188***	0.175***	0.029***
	(0.014)	(0.013)	(0.005)
Bandwidth	0.0850	0.103	0.156
(b) Only first and second recipient			
Female × Mayor	0.081***	0.014	0.004
	(0.020)	(0.017)	(0.009)
Mayor	0.195***	0.177***	0.028***
	(0.014)	(0.013)	(0.005)
Bandwidth	0.0890	0.102	0.167
(c) Assign extra votes to turnout			
Female × Mayor	0.111***	0.027	-0.002
	(0.021)	(0.018)	(0.010)
Mayor	0.175***	0.171***	0.025***
	(0.015)	(0.014)	(0.007)
Bandwidth	0.0744	0.0876	0.0959
(d) Weight by $1/ncandidates$			
Female × Mayor	0.119***	0.039**	0.001
	(0.022)	(0.019)	(0.010)
Mayor	0.166***	0.167***	0.020***
	(0.016)	(0.015)	(0.007)
Bandwidth	0.0744	0.0876	0.0959
(e) Assign extra votes to other candidates			
Female × Mayor	0.108***	0.019	0.002
	(0.021)	(0.017)	(0.009)
Mayor	0.177***	0.171***	0.025***
	(0.015)	(0.014)	(0.005)
Bandwidth	0.0431	0.0523	0.0822

Notes: In Column (1), the dependent variable is an indicator equal to 1 if the candidate is elected as mayor during the second round. In Column (2), the dependent variable is a dummy equal to 1 if the candidate runs for office in the occasion of elections held in term $t+1$. In Column (3), the dependent variable is a dummy equal to 1 if the candidate is elected as mayor during elections held in term $t+1$. In Column (4), the dependent variable is a dummy equal to 1 if the candidate is in the future elected as a member of the Provincial Council or to the position of President of the Province. All specifications include the interaction terms between the Mayor dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (11) with uniform kernel and Calonico et al. (2014) optimal bandwidth. In Panel (a), $WinningMargin_{i,t}$ is the vote share distance between the candidate and the minimal vote share required to be eligible to compete in the second round. In Panel (b), the running variable $WinningMargin_{i,t}$ is computed as in Panel (a) and the sample is restricted to candidates who either rank as the runner-up or as the third-largest recipient during the second round. In Panel (c), $WinningMargin_{i,t}$ is the distance between each candidate and the minimal vote share required to be eligible to compete in the second round, computed by adding to (removing from) each candidate a number of votes from (to) the total number of valid votes in the spirit of Fiva et al. (2018). See Appendix E for details. In Panel (d), $WinningMargin_{i,t}$ is defined as in Panel (c), and each individual observation is weighted by $1/ncandidates$. In Panel (e), $WinningMargin_{i,t}$ is the distance between each candidate and the minimal vote share required to be eligible to compete in the second round, computed by adding to (removing from) each candidate a number of votes from (to) other candidates, proportionally to their original vote share, in the spirit of Folke (2014) and Fiva et al. (2018). See Appendix E for details. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

F.2 RD diagnostics for equation (12)

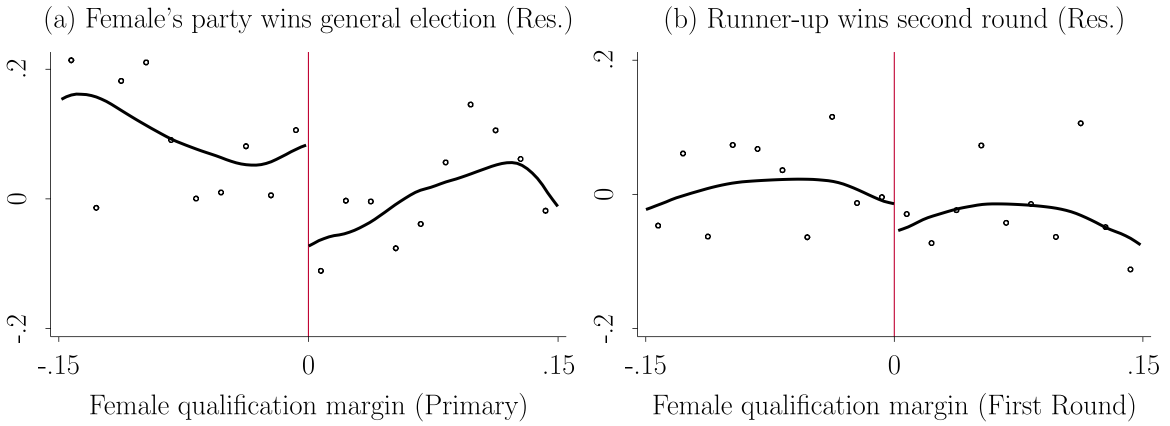

Figure F3: RD Plots

Notes: The dependent variables are the residual of a regression of the form $\mathbf{1}(Elected = MargQualified)_{e,t} = \mathbf{X}_{e,t}\beta + u_{e,t}$, where the controls in $\mathbf{X}_{e,t}$ are the variables included in Column (4) of Table B23. In both panels, the solid lines represent non-parametric smoothers of the qualification margin, separately estimated on either side of the qualification threshold. Markers represent sample averages within bins of the running variable $h^F_{e,t}$ equal to 0.015.

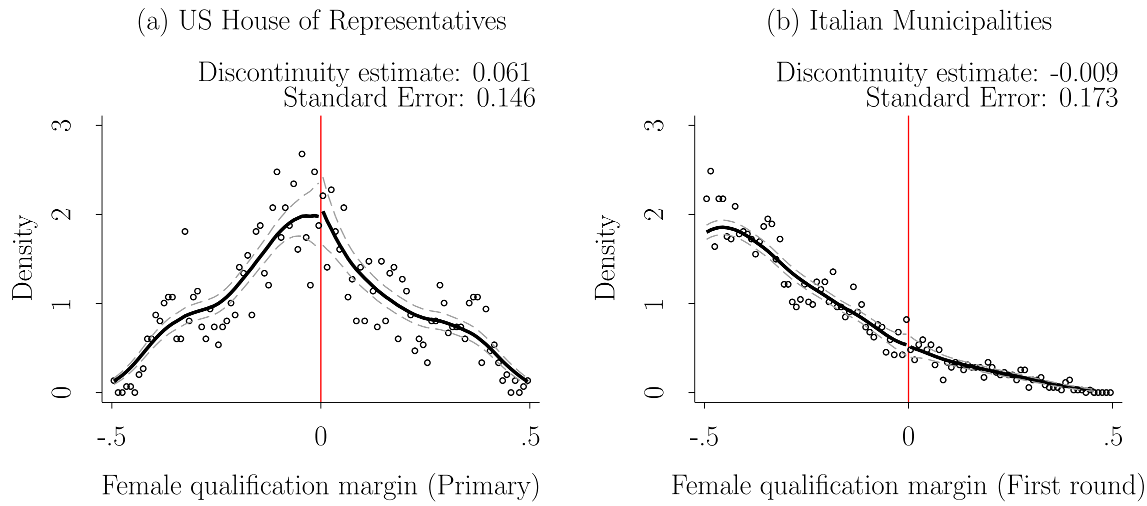

Figure F4: McCrary (2008) test

Notes: Panel (a): In both panels, log-density discontinuity and standard errors are computed performing a formal McCrary (2008) with optimal bandwidth. Markers represent sample averages within bins of the running variable $h^F_{e,t}$ equal to 0.01.

Table F3: Balance of pre-determined characteristics (US House of Representatives)

	(1)	(2)	(3)
Variable	Coefficient	St. Error	Obs.
No. candidates (Primary)	-0.194	(0.242)	679
Vote share of top candidate (Primary)	0.040	(0.074)	587
Vote share of runner-Up (Primary)	0.155	(0.109)	499
Valid votes (Primary)	-0.221	(0.162)	489
Dem. vote share last pres. el.	-0.086	(0.139)	531
Open primary	0.125	(0.159)	614
Incumbent runs	0.238	(0.267)	604
Redistricted boundaries	-0.136	(0.084)	451
Democratic seat	-0.114	(0.186)	489
Democratic party	-0.074	(0.177)	537
Incumbent	0.052	(0.035)	325
P(Experience!=0)	-0.300	(0.227)	282

Notes: In each column, the dependent variable is specified in the column header. All specifications include the interaction terms between the Female Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (12) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the congressional district-by-party level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

Table F4: Balance of pre-determined characteristics (Italian municipalities)

	(1)	(2)	(3)
Variable	Coefficient	St. Error	Obs.
Vote share top candidate (First round)	-0.226	(0.206)	403
Vote share runner-Up (First round)	0.169	(0.106)	590
Vote share coalition of top candidate (First round)	-0.162	(0.214)	375
Vote share coalition of runner-up (First round)	0.221*	(0.121)	550
Turnout (First round)	0.092	(0.134)	325
Male turnout (First round)	0.237	(0.165)	332
Female turnout (First round)	0.010	(0.165)	313
Voters (First round)	-0.126	(0.109)	482
No. candidates	-0.009	(0.220)	266
No. of lists competing	0.170	(0.210)	394
Non-valid votes (First round)	0.065	(0.135)	363
Incumbent runs for office	-0.071	(0.178)	438
No. parties in coalition of top candidate (First round)	0.177	(0.178)	399
No. parties in coalition of runner-Up (First round)	0.160	(0.163)	467
Census population	-0.145	(0.130)	550
Surface (km^2)	-0.025	(0.162)	379
Elderly (%)	-0.757***	(0.260)	265
Migrants (%)	0.262*	(0.148)	567
Students (%)	0.044	(0.203)	326
Unemployment (%)	0.006	(0.194)	332
South	0.292	(0.204)	278
Vote share of the left in last national election	-0.457**	(0.212)	310
P(Mayor elected during first round)	-0.285	(0.213)	352

Notes: The dependent variable is specified in each row. All specifications include the interaction terms between the Female Qualified dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (12) with uniform kernel and Calonico et al. (2014) optimal bandwidth. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

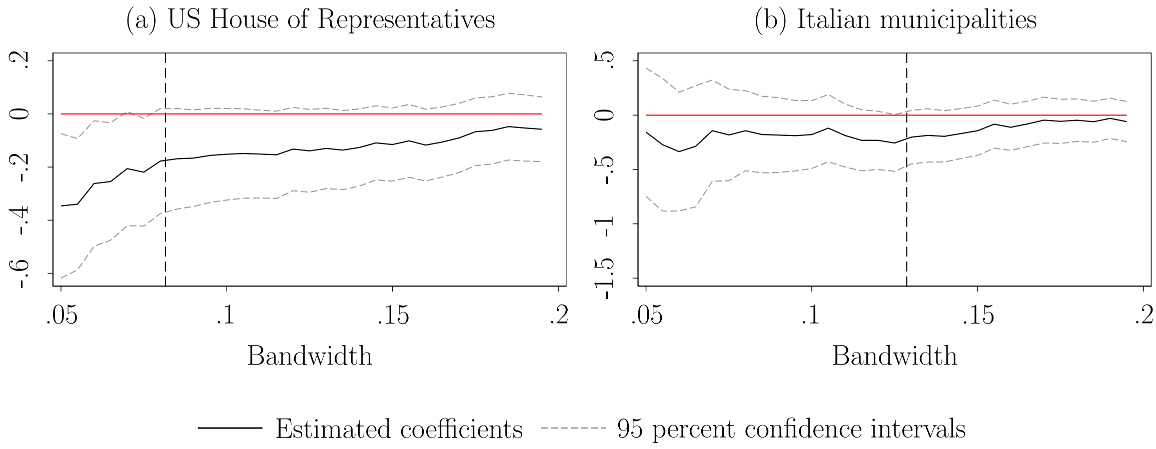

Figure F5: Bandwidth sensitivity

Notes: Panel (a): The dependent variable is a dummy equal to 1 if the party wins the seat in the US House of Representatives. The estimated equation includes state-by-party fixed effects and election year dummies. Panel (b): The dependent variable is a dummy equal to 1 if the runner-up is elected as mayor during the second round. The estimated equation includes province fixed effects, election year dummies, and the covariates included in Columns (2) and (4) of Table B23. In all panels, the horizontal axis represents the bandwidths used to fit the local linear regression. The solid line represents the estimated coefficients for κ as a functon of the chosen bandwidth. Dashed lines represent the 95 percent confidence intervals of each coefficient. Estimation method: local linear regression with uniform kernel as in equation (12), and bandwidths ranging from $h^F_{e,t} = 0.05$ to $h^F_{e,t} = 0.20$. The vertical line represents the Calonico et al. (2014) optimal bandwidth. All specifications include the interaction terms between the Female Qualified dummy and the assignment variable. Each estimation concerns a variation of bandwidth equal to 0.005. 95% confidence intervals are based on standard errors robust to clustering at the congressional district-by-party level (Panel (a)) and municipality level (Panel (b)).

F.3 RD diagnostics for equation (8) in municipalities subject to the plurality system

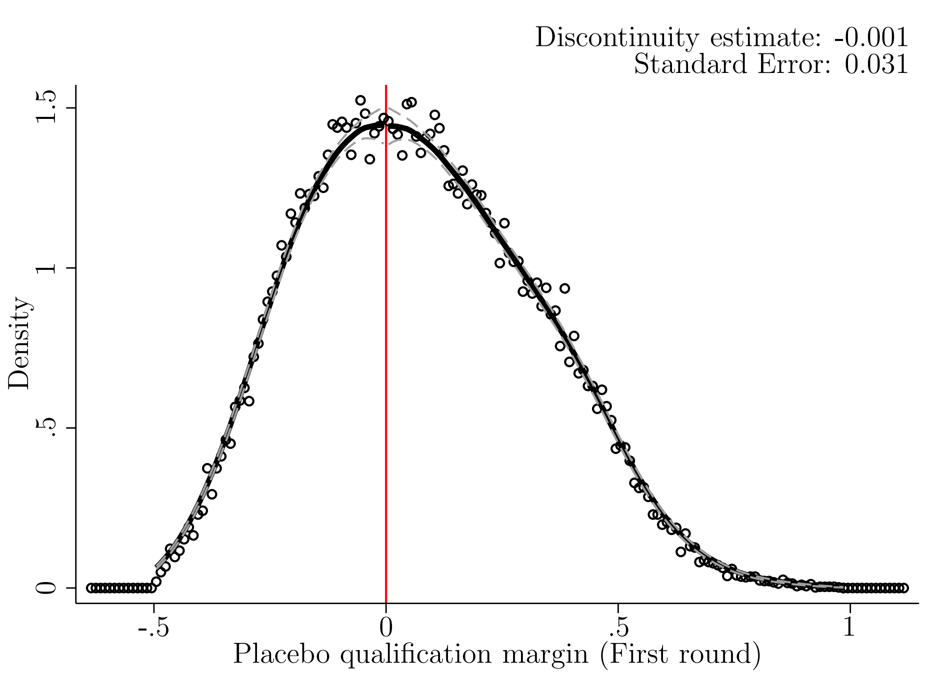

Figure F6: McCrary (2008) test

Notes: Log-density discontinuity and standard errors are computed performing a formal McCrary (2008) with optimal bandwidth. Markers represent sample averages within bins of the running variable $h_{i,t}$ equal to 0.01.

Table F5: Balance of pre-determined Characteristics

	(1)	(2)	(3)
Variable	Coefficient	St. Error	Obs.
Female candidate	-0.078**	(0.037)	11,939
College	-0.031	(0.037)	11,342
High-skill job	-0.007	(0.037)	11,690
Age	0.039	(0.034)	12,995
Incumbent	0.011	(0.023)	13,715
Candidacy counter	-0.031	(0.028)	14,964
Run again	-0.013	(0.028)	14,674
Vote share (First round)	0.000	(0.003)	11,070
Vote share top candidate (First round)	-0.014	(0.010)	9,846
Vote share runner-Up (First round)	0.015**	(0.007)	9,820
Turnout (First round)	-0.009	(0.006)	15,415
Male turnout (First round)	-0.003	(0.007)	13,924
Female turnout (First round)	0.001	(0.007)	15,446
Voters (First round)	0.012	(0.010)	12,639
No. candidates	0.013	(0.010)	10,791
No. of lists competing	0.013	(0.010)	10,791
Non-valid votes (First round)	-0.003	(0.004)	15,018
Incumbent runs for office	-0.003	(0.008)	19,584
Census population	0.015	(0.010)	12,823
Surface (km^2)	0.012	(0.011)	12,575
Elderly (%)	-0.003	(0.007)	13,213
Migrants (%)	-0.003	(0.007)	14,148
Students (%)	0.008	(0.008)	16,271
Unemployment (%)	0.018**	(0.009)	15,034
South	0.015*	(0.009)	16,701
Vote share of the left in last national election	-0.011	(0.009)	13,415
P(Mayor elected during first round)	-0.010	(0.008)	13,634

Notes: The dependent variable is specified in the column header. All specifications include the interaction terms between the Placebo dummy and the function of the assignment variable and election year dummies. Estimation methods: local-linear regression as in equation (8), where β_1 and β_3 are restricted to be equal 0 with uniform kernel and Calonico et al. (2014) optimal bandwidth. Each variable reported in the table has been previously standardized to have mean equal to 0 and variance equal to 1. Standard errors robust to clustering at the municipality level are in parentheses. *,**,*** represent the 10%, 5%, 1% significance levels.

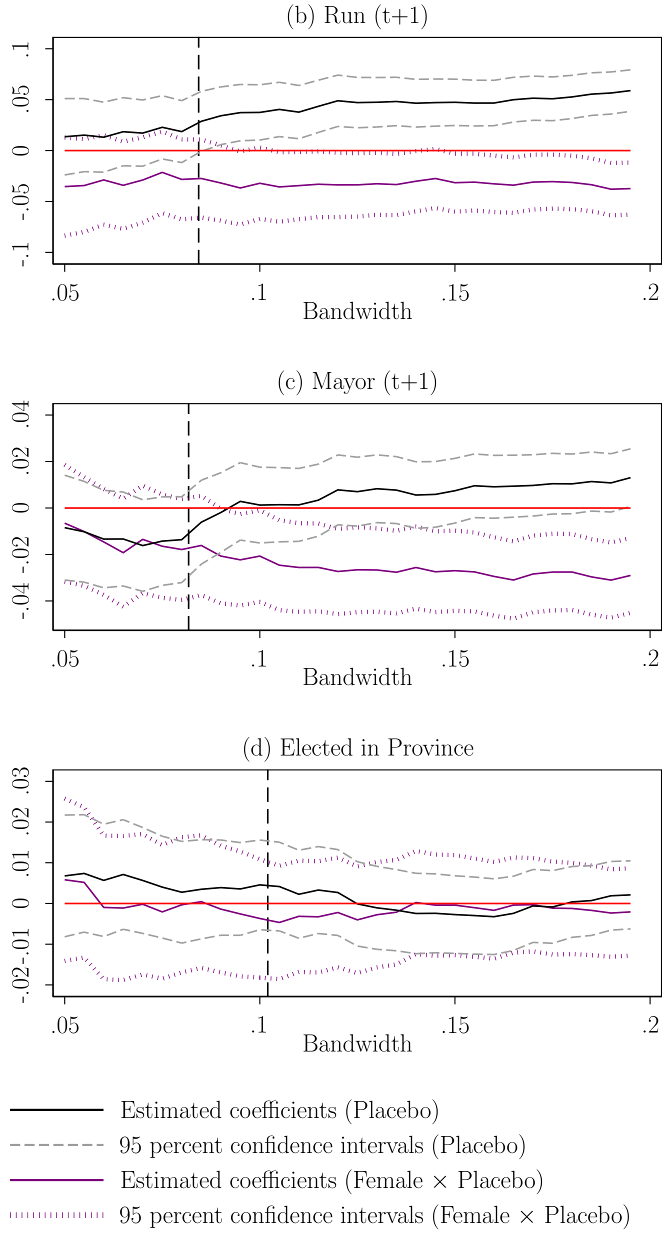

Figure F7: Entry to second round and future career paths (Placebo)

Notes: In Panel (a), the dependent variable is a dummy equal to 1 if the candidate runs for office in the occasion of elections held in term $t+1$. In Panel (b), the dependent variable is a dummy equal to 1 if the candidate is elected as mayor during elections held in term $t+1$. In Panel (c), the dependent variable is a dummy equal to 1 if the candidate is in the future elected as a member of the Provincial Council or to the position of President of the Province. The horizontal axis represents the bandwidths used to fit the local linear regression. The black solid line represents the estimated coefficients for β_2 as a function of the chosen bandwidth, while the purple solid line represents the estimated coefficients for β_1 as a function of the chosen bandwidth. Dashed lines represent the 95 percent confidence intervals of each coefficient. Estimation method: local linear regression with uniform kernel as in equation (8), and bandwidths ranging from $h_{i,t} = 0.05$ to $h_{i,t} = 0.20$. The vertical line represents the Calonico et al. (2014) optimal bandwidth. All specifications include the interaction terms between the Placebo dummy and the assignment variable. Each estimation concerns a variation of bandwidth equal to 0.005. 95% confidence intervals are based on standard errors robust to clustering at the municipality level. Standard errors are clustered at the municipality level.