UNIVERSITÀ DEGLI STUDI DI BERGAMO
Department of Economics

WORKING PAPERS

Turning bias into leverage: the case for gender quotas

Michela Cella, Elona Harka,
Elena Manzoni, Francesco Scervini

February 2026 - WP N. 38 Year 2026

Working papers – Department of Economics
n. 38

Turning bias into leverage: the case for gender quotas

Michela Cella, Elona Harka,
Elena Manzoni, Francesco Scervini

Università degli Studi di Bergamo
2026

Turning bias into leverage: the case of gender quotas / Michela Cella, Elona Harka, Elena Manzoni, Francesco Scervini.
Bergamo: Università degli Studi di Bergamo, 2026.
Working papers of Department of Economics, n. 38
ISSN: 2974-5586
DOI: 10.13122/WPEconomics_38

Progetto grafico: Servizi Editoriali – Università degli Studi di Bergamo
Università degli Studi di Bergamo
via Salvecchio, 19
24129 Bergamo
Cod. Fiscale 80004350163
P. IVA 01612800167

https://aisberg.unibg.it/handle/10446/321005

Turning bias into leverage: the case for gender quotas*

Michela Cella (University of Milan-Bicocca) Elona Harka (University of Bergamo)

Elena Manzoni (University of Bergamo) Francesco Scervini (University of Pavia)

February 16, 2026

Abstract

This paper explores how voter prejudice against women in politics can evolve over time through learning. Building on a Bayesian updating framework, we model voters as holding biased priors about female candidates' competence and updating their beliefs based on the performance of elected women. The model predicts that gender quotas can accelerate this learning process by increasing the visibility of competent female politicians, and that the effect is stronger, i.e., learning is faster, in more biased contexts. We test these predictions using two institutional reforms in Italian municipal elections. First, we exploit a short-lived gender quota reform (1993-1995) and find that quotas had a persistent effect on women's representation even after their removal. Their impact was stronger in municipalities with historically higher gender bias, proxied by referendum outcomes on abortion and divorce and the gender gap in education. Second, we examine the 2012 introduction of gender quotas and gender-conditioned double preference voting in municipalities with over 5,000 inhabitants. Using a sharp regression discontinuity design that allows for heterogeneous effects, we find that the reform increased women's representation in the first post-reform election and in the subsequent electoral cycle. Effects are larger in municipalities with higher

*We thanks participants to the SIEP Conference 2025 (Naples), and to the Workshop " BREAK THE BIAS: new data and methods to investigate the presence and impact of gender stereotypes on economic outcomes" (Bergamo) for useful comments. This work was funded by the European Union - NextGenerationEU, under the National Recovery and Resilience Plan (NRRP), Mission 4, Component 2, Investment 1.1, funding call PRIN 2022 D.D. 104 published on 2.2.2022 by the Italian Ministry of University and Research (Ministero dell'Università e della Ricerca), Project Title: BREAK THE BIAS: new data and methods to investigate the presence and impact of gender stereotypes on economic outcomes - CUP F53D23003020006.

pre-treatment prejudice, especially in the subsequent cycle. Overall, these results support our theoretical prediction that quota-driven exposure to female politicians reduces bias more quickly where initial prejudice is stronger.

JEL-Codes:
Keywords: gender bias, gender quotas, municipal council, abortion law.

1 Introduction

On November 8 2020 Vice President Elect Kamala Harris during her victory speech on election night said "While I may be the first woman in this office, I will not be the last", acknowledging that when women first access positions previously held by men they pave the road for other women to follow.

It is widely recognised that role models hugely contribute to reducing the gender gap in education and labour market (Gneezy et al., 2009; Porter and Serra, 2020) but being exposed to female leadership increases the performance of women also in the political arena (Campbell and Wolbrecht, 2006; Broockman, 2014; Bhalotra et al., 2018).

Greater female representation in office has been shown to increase both the likelihood and success of future female political participation, through mechanisms on both the supply and demand side. On the supply side, more women run for office despite substantial opportunity costs related to caregiving responsibilities inside the family (Fox et al., 2001), of their lower presence in the pipeline professions(Welch, 1977; Clark, 1991) and of their under-confidence (Fox and Lawless, 2011). On the demand side we see more voters willing to choose a woman on the ballot in spite of the still widespread belief that men are better suited as political officers (Huddy and Terkildsen, 1993; Dolan, 2004; Lawless, 2004; Falk and Kenski, 2006).

Reticence to vote for women is often explained as a more or less conscious gender-bias due to lack of information on part of the electorate. Sanbonmatsu (2002) shows how voters may use gender as a low-information short-cut when making decisions about which candidate to vote for. Gender is in fact the second heuristic after party affiliation that voters follow in case of low attention to politics and lack of sufficient information about the competence of the candidates.

Recently Bordalo et al. (2016) have formalized the formation of stereotypes building on the work of Tversky and Kahneman (1983). In their work, a stereotype is a probability distribution that over-(under-)estimates the likelihood of an event but it builds on a kernel of truth that is the “first thing that comes to mind” when making decisions. When dealing with gender stereotypes in politics though, this element of truth cannot be found anymore. The evidence is indeed consistent with the fact that women in office are on average better than the men in the same elected body. More precisely, there is evidence that women tend to have greater prior political experience (Pearson and McGhee, 2013), that they deliver more federal funds for their district (Anzia and Berry, 2011), that they put more bills through the legislative process (Volden et al., 2013) and that they deliver more speeches on the house floor (Pearson and Dancey, 2011).

The belief that voters hold about women being less competent although not corroborated by actual present data may be grounded in history. Indeed, in the not so distant past, it was true that women had a different distribution of political competence due to gaps in education and labour market participation. Nowadays, the gap in education has disappeared in most countries, and the gap in labour market participation is steadily shrinking. On the opposite, the gap in political performance and participation suffers from this additional information aspect.

If the bias is due to lack of information then institutional changes that allow for less costly acquisition of information may help reduce the bias. This is indeed the rationale for affirmative action policies like gender quotas in politics, which have proven to be effective in diminishing the gender gap among elected officials (Duflo, 2005; Beaman et al., 2009; De Paola et al., 2010; Franceschet et al., 2012).

Building on Cella and Manzoni (2023) we propose a way of studying the dynamics of the bias held by voters. We assume that voters believe that the competence of women candidates is drawn from a distribution that gives more weight to the lower values of competence and therefore elect them less often than men. This implies though that women that win elections are on average more competent than men.

We then use a Bayesian statistics method to study how this prejudice evolves over time. In our model voters update their beliefs using the observed competence of elected women, whenever the actual competence is above the mean of their perceived distribution their bias is reduced. This obviously is more likely to happen the lower is the perceived mean of women’s competence, therefore the stronger the bias.

In our framework, affirmative action policies, and gender quotas in particular, can work as an exogenous variation in the frequency of observations coming from female politicians. In other words any policy that increases exogenously the number of candidate or elected women will allow voters to acquire information on a larger set of women. In our environment this implies a faster reduction in the gender bias, as increasing the number of observations from the true distribution helps estimating the parameters faster.

One testable implication is that this effect would persist even in case these affirmative policies are then removed and that this is consistent with the empirical evidence (see Beaman et al., 2009; Bhavnani, 2009; De Paola et al., 2010).

A second testable implication is that the reduction in the bias would be faster where prejudice against women is stronger due to history and social norms. While the idea that affirmative action may be more effective where the disadvantage is deeper is present in empirical studies (see, for example, Bagde et al., 2016; Chattopadhyay and Duflo, 2004), the effect of quotas on gender bias has not been tested.

We test these predictions using Italian administrative data and exploit two institutional reforms affecting municipal elections. First, we study the temporary introduction of gender quotas under Law 81/1993, which were later declared unconstitutional and removed. This setting allows us to cleanly identify learning effects by examining electoral outcomes after the quotas ceased to apply. We find that municipalities exposed to quotas experienced a persistent increase in women's representation in subsequent elections, and that this effect is significantly stronger in areas characterized by higher pre-existing gender prejudice, proxied by historical referendum outcomes and educational gaps.

Second, we analyze the 2012 reform introducing permanent gender quotas and gender-conditioned double preference voting in municipalities with more than 5,000 inhabitants. Using a sharp regression discontinuity design that allows for heterogeneous treatment effects, we show that the reform substantially increased women's representation in both the first and second post-reform electoral cycles. While the immediate effects of the reform are large and largely homogeneous—reflecting the direct mechanical impact of the quota—the heterogeneous effects predicted by our model emerge more clearly over time. In particular, in the second electoral cycle, the increase in women's representation is significantly larger in municipalities with higher levels of pre-treatment prejudice. This dynamic pattern is consistent with a learning mechanism through which exposure to female politicians gradually reduces voter bias, especially where initial resistance to women's political participation is stronger.

Related literature based on Law 81/1993 and Law 215/2012 reforms. A growing literature studies the effects of the two gender quotas reform. Baltrunaite et al. (2014) show that the 1993 introduction of gender quotas on candidate lists increased the number and quality of elected women, with spillovers on political selection. Braga and

Scervini (2017) show how this in turn affects policy implementation and efficiency. De Paola et al. (2010)provide evidence that exposure to female politicians under the 1993 reform had persistent effects on women's political participation even after the abolition of the quotas.

Our analysis complements this literature by focusing explicitly on the role of pre-existing gender prejudice. Rather than estimating average treatment effects, we study how the impact of temporary quotas varies across municipalities with different initial levels of bias, and whether learning effects persist beyond the removal of the policy. This allows us to provide direct empirical evidence on the heterogeneous dynamics predicted by our theoretical framework.

For what concerns the 2012 reform, instead, Baltrunaite et al. (2019) and Andreoli et al. (2022) provide evidence that the introduction of quotas and double preferences by gender significantly increased women's representation. Our contribution differs in two key respects. First, we explicitly allow the effects of the reform to vary with pre-treatment gender prejudice. Second, we study the evolution of these heterogeneous effects across electoral cycles. This dynamic perspective is crucial in a setting where quotas are permanent and mechanical effects may coexist with belief updating.

The structure of the paper is as follows: Section 2 summarizes the key elements of the model in Cella and Manzoni (2023) and adapts them to the current needs; Section 3 studies the dynamic evolution of the bias; Section 4 describes the institutional setting of the reform, the data and our proxies for pre-existing gender bias; Section 5 discusses the methodology and the empirical findings on the impact of Law 81/1993; Section 6 presents the methodology and the empirical findings on the impact of Law 215/2012; and Section 7 concludes.

2 The model

We build on the model of gender bias due to prejudices on female competence developed by Cella and Manzoni (2023). In this section we briefly introduce the main assumptions of the model, and its main findings, as they are a building block for they analysis of the dynamics of Section 3.

We consider a two-period two-party election model: in each period, first two candidates (one from party L and one from party R) run for election, and then the winner implements a policy decision. There are two dimension of policy: ideology position of the policy, x and competence, v. Agents (both politicians and voters) have heterogeneous policy preferences but an homogeneous taste for higher competence.

Voters. Voter i's utility in period t is

$$u_t^i\left(y_t, v_t^P\right) = -\left(x^i - y_t\right)^2 + v_t^P,$$

where x_i is i's preferred ideological position, y_t is the implemented policy position, and v_t^P is the competence of the elected politician, whose party is $P = L, R$. Voters' preferred policies are distributed uniformly over $[-1, 1]$ so that the median voter bliss point is $x_m = 0$.

Politicians. Politicians are characterised by ideology x^k, and competence v^k, $k = L, R$, and they are policy oriented. The utility in period t of the candidate from party k is the same as the voters' one:

$$u_t^k\left(y_t, v_t^P\right) = -\left(x^k - y_t\right)^2 + v_t^P,$$

where x^k is the preferred policy of the politician from party k. Politicians' policy preferences are uniformly distributed on supports that differ according to their party: $x_L \sim U[-1, 0]$ and $x_R \sim U[0, 1]$. Every candidate's competence has the same distribution, i.e., $v^k \sim U[0, 1]$ for $k = L, R$. Both the candidate's policy preference and level of competence are his/her private information before the election, and they are observed when s/he is elected.

Gender and gender bias. We assume that candidates are randomly selected from a gender-balanced population, the first time in which they compete for election. Candidate's gender is observable, and it changes the voters' expectations on his/her competence. As a matter of fact, voters' are prejudiced on the females' distribution of level of competence. The competence of the candidates is $v_t^k \sim U[0, 1]$ regardless of the candidate's gender. Cella and Manzoni (2023) assume that voters' prior belief on female candidates is that there is a probability ϕ_t that they come from a worse distribution, specifically $v_t^k|F \sim U[0, V]$ where $V \in \left(\frac{1}{3}, 1\right)$, and from the true one $v_t^k|F \sim U[0, 1]$ with the complementary probability $(1-\phi_t)$.[1] In this paper, we express the prejudice with a different, but equivalent, formulation, in order to use the results by Craigmile and Tirrerington (1997) later to analyze the dynamics of the bias. We assume that the voters' perceive the distribution of

[1] The assumption $V > \frac{1}{3}$ prevents the trivial case in which each male candidate is perceived as more competent than any female one. See the original paper for a discussion on this.

female competences as a mixture density of the following form:

$$f(x|p, V, F) = p_t U[0, V] + (1 - p_t) U[V, 1],$$

i.e., as a mixture of a uniform over $[0, V]$ and a uniform over $[V, 1]$, with weight p_t. In this formulation, the true distribution of female competences has density $VU[0, V] + (1 - V)U[V, 1]$—i.e., a uniform over $[0, 1]$—while voters' belief in the presence of a prejudice is characterised by a higher weight on the uniform over $[0, V]$, i.e., $p_t > V$. Note that $p_t = \phi_t + (1 - \phi_t)V$, so that $p_t = V$ is equivalent to $\phi_t = 0$, i.e., to the true unbiased distribution. Therefore, the higher is p_t the higher is the gender bias. Period 1 belief p_1 is taken as given, while period 2 belief p_2 may be updated if a female politician is elected in period 1 and her competence is observed. We discuss the dynamics of p_t in Section 3.

Signal. When a candidate runs for the first time, voters observe a signal $\sigma_t^k \in \{\underline{v}, \overline{v}\}$ which perfectly reveals whether the competence of the candidate is below ($\underline{v}$) or above ($\overline{v}$) the median of its group. Hence the perceived expected competence given the signal σ_t^k differs for male and female candidates as follows:[2]

$$\mathbb{E}\left[v_t^k|\sigma_t^k, M\right] = \begin{cases} \frac{3}{4} & \text{if } \sigma_t^k = \overline{v} \\ \frac{1}{4} & \text{if } \sigma_t^k = \underline{v} \end{cases} \quad ; \quad \mathbb{E}\left[v_t^k|\sigma_t^k, F\right] = \begin{cases} \frac{3}{4}(1 + V - p_t) & \text{if } \sigma_t^k = \overline{v} \\ \frac{1}{4}(1 + V - p_t) & \text{if } \sigma_t^k = \underline{v} \end{cases} .$$

Recalling that $p_t > V$, the expected competence of a female candidate is lower than the one of a male candidate for any possible signal. Moreover $\mathbb{E}\left[v_t^k|\sigma_t^k, F\right]$ is decreasing in the bias p_t.[3]

2.1 Equilibrium analysis

The full equilibrium of this voting game, in terms of policy choices, voting decisions and possible re-candidacy decisions, is characterized in Proposition 1 of Cella and Manzoni (2023). Here we first summarize the main properties of this equilibrium, and then we

[2] The expected competences can be computed from the ones in Cella and Manzoni (2023) recalling that $\phi_t = \frac{p_t - V}{1 - V}$.

[3] As in the original model, given the assumption $V > \frac{1}{3}$, it is always the case that a woman with a high signal has a higher expected competence than a man with a low signal for any possible p_t.

highlight how the intensity of the prejudice p_t affects the winning probability of female candidates.[4]

Implemented policies. In this model, the implemented policy is always equal to the bliss point of the elected politicians. In the second period, this is due to the absence of re-election incentives, as no re-election is possible. In the first period, this is driven by the observability of the policy preference after election, which eliminates the incumbent's incentives to mimic a different policy preference, because voters will in any case forecast correctly his/her policy choice in case of re-election.

Voting on untried candidates. When candidates are ex-ante equal in terms of expected ideological distance from the median voter. Therefore the median voter chooses on the basis of the expected competence level, which is determined by the gender-signal pair. The median voter ranking of gender-signal pairs is

$$(M, \overline{v}) \succ (F, \overline{v}) \succ (M, \underline{v}) \succ (F, \underline{v}),$$

and he/she randomizes with equal probability when indifferent.

Elections with an incumbent. When an untried candidate runs for re-election against an incumbent, the incumbent's policy preference and competence are known, hence the incumbent's gender does not affect voting. The incumbent is more likely to be re-elected for higher competences and lower policy biases, and the lower is the expected competence of the challenger (i.e., his/her position in the gender-signal pairs ranking). The incumbent anticipates this, but does not know the gender-signal pair of the challenger before the re-election decision. Therefore his/her re-election choice is affected only by his/her level of competence and policy bias.

Effects of the gender bias on the female probability of winning. The presence of gender bias has two effects on the probability of winning of female candidates. First, when they run for an open seat—i.e., in an election with two untried candidates—they lose the electoral competition against males with the same signal. This reduces the female

[4] Cella and Manzoni (2023) also show that gender bias positively affect expected competence of elected women. We refer to their discussion on this relevant consequence of the model, which is not the focus of the current analysis.

candidate's probability of winning, but in a way that is not influenced by the specific value of p_t, given the structure of the signal that we consider.[5]

Second, the bias also affects the probability of winning when running against an incumbent. In equilibrium, an incumbent only runs for re-election if he wins against challengers with low signals. Therefore, the difference in the winning probability is driven by female challengers with high signals, who only win against incumbents with, i.e., those with $v^I < \left(x^I\right)^2 + \frac{5}{12} - \frac{3}{4}(p_2 - V)$. Hence, the higher p_t, the lower the probability that a female challenger with a high signal wins against an incumbent.

Female incumbents instead have the same probability of re-election as male incumbents, as competence is observed, and therefore no longer subject to prejudice.

Running for re-election. As voters' evaluation of an incumbent does not depend on his/her gender, re-election incentives are the same for male and female politicians. There are however two indirect effects of gender on the re-candidacy decision. First, as discussed in Cella and Manzoni (2023), elected women have higher competences and therefore they are are more likely to find re-candidacy optimal, so that on average they run for re-election more often. This effect, however, does not vary with p_t. A second effect is that higher p_t increases the re-election incentives for both males and females incumbent by making female challengers less competitive. As the probability of running for re-election is higher for females, this could benefit them.[6]

3 Dynamics of the gender bias

In this model, voters hold the prejudice that female politicians are on average less competent than men, that is they are drawn from a distribution which gives more weight to lower values of competence. In the Introduction, we argued that this prejudice may come from the past. As it used to be the case that women had lower or no access to education and to political experience, if we interpret competence as a trait that does not only depend on intrinsic qualities but also on learning, it is reasonable to think that their distribution of competence may have been different from the male one until the not so distant past. We are thus interested in understanding how a prejudice which has its origin in the past may evolve over time.

[5] If we lower the precision of the signal p_t has an effect also on this probability of winning.

[6] We believe this to be a second order effect due to the small share of female incumbents.

To understand the dynamics of the bias, we ask first what estimator can voters use to estimate the parameter of the distribution of valences. The true distribution is indeed a uniform on $[0, 1]$. However, as discussed in Section 2, voters believe that the distribution of female valences is the following mixture of two uniforms

$$f(x|p, V) = pU[0, V] + (1 - p)U[V, 1],$$

i.e., as a mixture of a uniform over $[0, V]$ and a uniform over $[V, 1]$, with weight p. Recall that voters' belief is prejudiced if $p > V$.

When V is known, a consistent and unbiased estimator of p is given by (Gupta and Miyawaki, 1978)

$$\tilde{p} = 1 + V - 2\hat{\mu},$$

where $\hat{\mu}$ is the sample average. The existence of a consistent (and unbiased) estimator of p implies that if voters disregarded their original prejudice and estimate p on the basis of observed competences only, their estimates are immediately unbiased and reflect the actual distribution of female competences. However, prejudices die hard. We therefore assume that there is some degree of inertia in the beliefs. Modelling the specific functional form of this assumed inertia is not the aim of this paper, yet, some relevant properties of the dynamics can be derived under general assumptions on the update process itself.

Let us call $\hat{p}_t$ the estimator of p at time $t \geq 1$. Note that we call $\hat{p}_t$ the estimator of p at the beginning of time t, before the election takes place. Therefore $\hat{p}_t$ does not include the information on the female politicians that will be elected at time t. We assume that at time $t = 1$, that is, before the election of period 1, the estimator has a known (non random) value p_1. To account for the inertia of the prejudice we assume that $\{p_t\}$ is an autocorrelated process, so that the convergence of the estimator to the (true) value V happens gradually and not instantaneously.

In this case, we can measure how much voters reduce their prejudice by looking at the expected square variation of $\hat{p}_t$ between period 1 and period 2:[7]

$$\mathbb{E}\left[(\hat{p}_2 - p_1)^2\right] = Var[\hat{p}_2] + (\mathbb{E}[\hat{p}_2] - p_1)^2 .$$

[7] We consider the expected square variation because we are not interested in the sign of the expected variation itself.

Note that the expected variation is greater the larger the difference between the original belief p_1 and the expected value of the estimator $\mathbb{E}[\hat{p}_2]$.[8] This difference depends both on the strength of the prejudice and on the inertia of the estimator.[9] To understand the effects of these two elements, let us consider, as an example, an easy autoregressive process such as:

$$\hat{p}_2 = \alpha p_1 + (1-\alpha)\tilde{p} = \alpha p_1 + (1-\alpha)(1+V-2\hat{\mu}).$$

In this case $\mathbb{E}[\hat{p}_2] = \alpha p_1 + (1-\alpha)V$, so that $(\mathbb{E}[\hat{p}_2] - p_1)^2 = (1-\alpha)^2(V-p_1)^2$. Moreover, $Var[\hat{p}_2] = (1-\alpha)^2 Var[\tilde{p}] = \frac{(1-\alpha)^2}{3n}$, which does not depend on V.[10] Overall, the expected square variation can be written as

$$\mathbb{E}\left[(\hat{p}_2 - p_1)^2\right] = \frac{(1-\alpha)^2}{3n} + (1-\alpha)^2(V-p_1)^2 = (1-\alpha)^2\left(\frac{1}{3n} + (V-p_1)^2\right).$$

From this we can easily note that (i) a higher persistence of the prejudice (higher α) reduces the expected variation, while (ii) a stronger bias (higher distance between V and p_1) and (iii) a larger sample (higher number of elected women n) increase it, thus favouring a larger reduction of the prejudice. This result holds regardless of the specific functional form of the process, provided it is positively autocorrelated.

Strength of the bias and the effects of V. As discussed a few lines above, a stronger bias will be reduced more easily than a weaker one. This persistence of weak biases is consistent with the empirical evidence in the literature that suggests that an increased exposure to female politicians has a stronger positive effect on the number of elected women in countries where the bias can be thought to be stronger (Beaman et al., 2009), or when women make their début in the political arena (Gilardi, 2015), than in mature western

[8] Even though our theoretical model has only two periods, we can derive the expected square variation between two generic periods t and $t+1$, which is

$$\mathbb{E}\left[(\hat{p}_{t+1} - \hat{p}_t)^2\right] = Var[\hat{p}_{t+1}] + Var[\hat{p}_t] + (\mathbb{E}[\hat{p}_{t+1}] - \mathbb{E}[\hat{p}_t])^2 - 2Cov[\hat{p}_{t+1}, \hat{p}_t].$$

We can note that the expected variation between two generic periods t and $t+1$ displays the same comparative statics discussed for the variation between periods 1 and 2. This formulation highlights the effect of autocorrelation on the process, through the covariance term. Higher covariance, which corresponds to greater inertia, induces lower expected square variation.

[9] To understand the two effects more clearly, we can decompose the difference $\mathbb{E}[\hat{p}_2] - p_1 = (\mathbb{E}[\hat{p}_2] - V) - (p_1 - V)$, where $(p_1 - V)$ is a measure the strength of the prejudice.

[10] On the computation of $Var[\tilde{p}]$ see Craigmile and Tirrerington (1997).

democracies where open discrimination should be a thing of the past (Broockman, 2014). As the prejudice becomes weaker it is harder to remove it.

In our model we take V as constant and exogenously given, and we model the intensity of the bias through the parameter p. However, V can also be considered a measure of the strength of the gender bias, and it is likely to be heterogeneous across countries, due to the influence of specific cultural characteristics or historical background. Specifically, if V is low, voters allow for the possibility that female politicians are characterised by valences that are much lower than the males' ones.

Possible effects of gender quotas and testable predictions. Affirmative action policies, such as gender quotas, can be interpreted in our model as an exogenous variation in the frequency of observations coming from female politicians. In other words any policy that increases exogenously the number of candidate or elected women will allow voters to acquire information on a larger set of women. In our environment this implies a faster reduction in the gender bias, as increasing the number of observations from the true distribution helps estimating the parameters faster.

A second prediction of our model is that the effect of gender quotas in reducing the bias is faster the stronger the initial prejudice. We test this assumption by looking at differential effects of the introduction of gender quotas at the municipal level, and by measuring the strength of the initial gender bias through divorce and abortion referendum results in the considered municipalities.

4 Institutional setting and data

Italy is administratively divided into approximately 8,000 municipalities, spread across 20 regions. Fifteen of these regions follow a common institutional and legislative framework — referred to as ordinary statute regions — while the remaining five (Valle d'Aosta, Trentino-Alto Adige/Südtirol, Friuli-Venezia Giulia, Sicily, and Sardinia) are classified as special statute regions, enjoying greater legislative autonomy under the Italian Constitution. Due to potential differences in local electoral laws, particularly concerning the timing and implementation of institutional reforms such as gender quotas or direct mayoral elections, we exclude municipalities from these five special statute regions from our analysis.

In the ordinary statute regions, municipal elections are governed by national legislation. Before 1993, the standard term for municipal councils was five years. However, mayors were not directly elected by citizens but were instead appointed by the municipal council. Their tenure depended on the continuing support of the council majority, and as such, local administrations were frequently subject to early termination due to political instability, mayoral resignation, or the collapse of governing coalitions. While the institutional framework was formally homogeneous, this political dynamic introduces variation in the timing of municipal elections across municipalities — an aspect that we leverage in our empirical strategy.

A major shift occurred with the 1993 reform (Law 81/1993), which introduced the direct election of mayors in Italy, replacing the previous system of appointment by the municipal council. This reform aimed to increase political accountability and enhance administrative stability by strengthening the mayor's individual legitimacy and visibility in the electoral process.

Gender quotas in Law 81/1993. Law 81/1993 also included provisions intended to promote gender balance in local political representation. Specifically, it established gender quotas on candidate lists: in municipalities with up to 15,000 inhabitants, no more than three-quarters of the candidates on any list could belong to the same gender, while in larger municipalities the threshold was set at two-thirds. However, these provisions were short-lived. In Decision No. 422/1995, the Italian Constitutional Court declared the gender quota rules unconstitutional, citing a violation of Articles 3 and 51 of the Constitution, which guarantee equal access to public office without discrimination. The Court argued that gender-based restrictions on candidate lists represented an undue constraint that privileged one group over another. As a result, the gender quota provisions were only in effect from March 25, 1993, until their annulment on September 12, 1995.

A second introduction of gender quotas: Law 215/2012. After nearly two decades without gender-based constraints on local electoral competition, gender quotas were reintroduced with Law 215/2012. The reform applies to municipalities with more than 5,000 inhabitants in ordinary statute regions and represents a substantial strengthening of policies aimed at promoting women's political representation at the local level.

The law introduced two main innovations. First, it imposed a gender quota on municipal council candidate lists, requiring that neither gender exceed two-thirds of the total number of candidates. Second, it established a system of gender-conditioned double preference voting. Voters may cast either one preference vote or two preference votes; however, if two preferences are expressed, they must be assigned to candidates of different genders. If both preferences refer to candidates of the same gender, the second preference is invalid. The enforcement mechanism is automatic and administrative: electoral commissions intervene ex ante by deleting candidates from the over-represented gender—starting from the bottom of the list—until the quota requirement is satisfied. Unlike the 1993 provisions, Law 215/2012 was explicitly grounded in the revised constitutional framework following the 2003 amendment to Article 51 of the Italian Constitution, which authorizes the adoption of measures promoting equal opportunities between women and men in access to elected office. The reform was first implemented in the 2013 municipal elections.

4.1 Data sources and sample construction

The data used in this study originate from three distinct administrative sources. The *Anagrafe degli Amministratori Locali* (Registry of Local Administrators), published annually by the Italian Ministry of the Interior, provides some information on members of municipal bodies (mayor, council, and executive), including their gender. From these raw data, we compute the gender of each mayor and the share of women in both the municipal executive and council. The same source also includes the election or appointment date for each member, allowing us to identify whether a municipality held elections under the 1993 gender quota regulations. The second data source concerns referendum results, which are publicly available from the Ministry of the Interior. These results are provided at various levels of aggregation, including the municipality level. Finally, socio-demographic and geographic characteristics of each municipality are obtained from ISTAT, the Italian National Institute of Statistics.

Over the years, several municipalities have undergone institutional changes, such as mergers, splits, name changes, or changes in provincial affiliation (due to the creation of new provinces or redrawing of provincial boundaries). Whenever appropriate, we track the same administrative entity over time. For instance, we treat a municipality that changed only its official name as the same unit, while we exclude from the sample three municipalities that merged into a single, larger one.

Since referendum results are time-invariant, we cannot exploit the panel dimension of the dataset (see the next section for further details). We therefore construct a cross-sectional dataset comprising approximately 6,000 municipalities.

For the first study, time-varying variables are measured at the time of the first municipal election following the abolition of gender quotas in 1995. Table 1 reports descriptive statistics for the variables used in this analysis. For the second study, outcome variables are measured at the first municipal election after the 2012 reform (2013–2017) and, in the subsequent electoral cycle. Predetermined municipal covariates, namely, altitude, coastal municipality and urbanization are measured in 2011. Table 2 reports the descriptive statistics by treatment status for the municipalities within the optimal bandwidth.

Table 1: Descriptive Statistics

	Mean	Std. dev.	Obs.
Share of women in Municipal Council (1st election post gender quota)	0.180	0.110	6025
Share of women in Municipal Executive (1st election post gender quota)	0.163	0.214	6019
Female Mayor (1st election post gender quota)	0.070	0.255	6014
Municipality ever held election under gender quota	0.958	0.200	6025
Prejudice(abortion)	0.377	0.125	6025
Prejudice(divorce)	0.497	0.153	6013
Vote share for Christian Democracy (DC) in 1992 (prejudice)	0.351	0.131	6022
Composite prejudice index (prejudice)	0.408	0.120	6025
Education gender gap in 1991 (prejudice)	1.106	0.243	6017
Non-urban population (%)	21.886	20.058	6025
Male-to-female ratio	96.536	6.347	6025
Old-age dependency ratio (%)	29.551	13.020	6025
Legally separated or divorced (% of 18+)	1.317	0.872	6025
Adults (25-64) with upper secondary or tertiary education (%)	19.802	6.618	6025
Mean household size	2.696	0.360	6025

Note:

4.2 Proxies of pre-existing gender bias

The aim of the empirical analysis is to assess whether the effect of gender quotas on the share of women elected to municipal councils is stronger in municipalities where prejudice against women in politics was more prevalent before the introduction of the quotas.

Since direct measures of such prejudice are generally unavailable in administrative data, we proxy this bias at the municipal level using four separate possible measures (i) results of the divorce referendum, (ii) results of the abortion referendum, (iii) education gap, and (iv) share of votes for Christian Democrats in 1992. We also build a comprehensive indicator, which takes into account all four proxies. We now briefly introduce the four proxies, and discuss why we believe that, considered separately or together, they are a good measure of the bias.

Divorce and abortion referendum. The first referendum, held in 1974, concerned the repeal of a 1970 law legalizing divorce. About 60% of voters opposed the repeal, allowing the law to remain in force. The second referendum, held in 1981, similarly addressed the repeal of a law legalizing abortion; 68% of voters opposed the repeal, thereby maintaining the law. Voter turnout was very high in both cases by today's standards: 87.72% in 1974 and 79.43% in 1981.

We consider therefore as proxies the share of votes in favour of the repeal in each municipality. We argue that this measure is a credible and accurate proxy for gender-related prejudice. First, the referendum results are strongly correlated (correlation coefficient = 0.78). Second, their geographic distribution aligns with sociological literature documenting territorial disparities in gender norms across Italy.

Share of DC votes in 1992. An alternative proxy of pre-existing gender bias that we use is the share of votes for Christian Democrats (DC) in 1992. This is a variable available at municipal level and correlates with levels of religiosity which, as Bozzano (2017) shows, in turn correlate with the presence of traditional gender roles.

Education gap. Finally, we use as proxy of pre-existing prejudice the education gap, i.e., male/female ratio in the share with at least upper secondary education. We measure it in 1991 for the analysis of the 1993 reform, and in 2011 for the analysis of the 2012 one.

Table 2: Descriptive Statistics RD sample

	Control			Treated		
	Mean	Std. dev.	Obs.	Mean	Std. dev.	Obs.
Share of women in Municipal Council (1st electoral cycle (2013-2017))	0.273	0.127	507	0.400	0.0944	342
Share of women in Municipal Council (2nd electoral cycle)	0.304	0.132	566	0.426	0.0819	366
Share of women in Municipal Executive (2nd electoral cycle)	0.465	0.129	289	0.489	0.110	224
Female Mayor (2nd electoral cycle)	0.149	0.357	770	0.187	0.391	459
Share of votes against abortion right (prejudice)	0.381	0.121	576	0.382	0.125	369
Share of votes against divorce right (prejudice)	0.498	0.144	575	0.500	0.154	367
Vote share for Christian Democracy (DC) in 1992 (prejudice)	0.337	0.112	576	0.327	0.112	369
Composite prejudice index (prejudice)	0.405	0.111	576	0.403	0.117	369
Education gender gap in 2011 (prejudice)	1.020	0.0806	576	1.018	0.0731	369
Altitude	274.9	220.6	576	232.5	215.7	369
Coastal municipality	0.0712	0.257	576	0.0976	0.297	369
Urbanization	0.373	0.484	576	0.593	0.492	369

Note:

5 The effects of Law 81/1993

The first test of our theoretical prediction exploits the temporary introduction of gender quotas with Law 81/1993. In this analysis, we focus on the persistent effects of gender quotas after their removal.

5.1 Empirical Strategy

A key challenge in the empirical strategy is the time-invariant nature of the referendum data. Because these results do not vary over time, we cannot exploit the panel dimension of our dataset. We therefore rely on a cross-sectional analysis.

Our baseline specification is as follows:

$$y_{i,p} = \alpha + \beta x_{i,p} + \gamma w_{i,p} + \delta \left(x_{i,p} w_{i,p} \right) + \zeta z_{i,p} + \eta_p + \varepsilon_{i,p} \tag{1}$$

where the dependent variable $y_{i,p}$ is the share of women in municipal bodies in municipality i in province p; $x_{i,p}$ is a dummy equal to 1 if the municipality elected its council under gender quotas; $w_{i,p}$ is the measure of prejudice, proxied as described in Section 4.2 above; $x_{i,p}w_{i,p}$ is their interaction term; and $z_{i,p}$ is a vector of control variables (socio-demographic, economic, and geographic). Province- (or region-)level fixed effects η_p control for unobserved regional characteristics, and $\varepsilon_{i,p}$ is the idiosyncratic error term. Standard errors are clustered at the province level to account for spatial correlation in the error terms.

Identification of the causal effect relies on four main assumptions: a) quasi-random election timing: the timing of municipal elections is plausibly exogenous with respect to gender prejudice, as it depends on institutional rules and occasional early terminations due to political events (e.g., mayoral resignation); b) validity of the proxy: as discussed above, the proxies we use correlate well with broader societal attitudes toward women and gender roles; c) no reverse causality: the outcomes of municipal elections in the late 1990s cannot influence referendum results from the 1970s and 1980s, pre-existing electoral outcomes, or the pre-existing education gap ; d) no omitted variable bias: there are no omitted variables that are (i) correlated with gender prejudice, (ii) influence the share of women elected, (iii) are not included in the control vector, and (iv) are not absorbed by fixed effects.

These assumptions appear reasonable in our context. Election timing is largely driven by institutional processes. Our gender proxies reflect long-standing societal values. Lastly, any potential confounders that vary only at the provincial (or regional) level are accounted for by fixed effects.

5.2 Results

Overall, the regression results (Table 3) align with our theoretical expectations and support the testable implications derived from the model. Column (1) shows that the share of women elected to municipal councils increased significantly in municipalities subject to gender quotas, even after their abolition. This finding is consistent with the first implication: an exogenous increase in women holding office raises the likelihood of women being elected in subsequent elections. Among the control variables, the share of separated or divorced adults and average household size both exert significant effects on the share of elected women, with the expected signs, even after controlling for province fixed effects. These effects remain stable across specifications.

Column (2) augments the specification with a measure of prejudice, proxied by the share of votes against abortion rights in the 1981 referendum. Both the quota and prejudice indicators display the expected signs in relation to the share of women elected. However, the statistical significance of the prejudice measure varies across specifications (Table A2 in Appendix).

Column (3) introduces the interaction term, δ in equation (1), which is positive and statistically significant. This indicates that the impact of gender quotas on women's electoral success is stronger in municipalities with higher levels of prejudice — consistent with the theoretical prediction that quotas are most effective where barriers to female participation are greatest. Figure 1 illustrates this point: the effect of gender quotas is significantly positive only in high-prejudice municipalities, and not significant where prejudice is low. Moreover, the regression shows that prejudice reduces the share of women elected only when quotas are absent; in their presence, the effect of prejudice is offset, yielding a net impact close to zero (-0.0157, p-value 0.338). Taken together, the evidence from this model strongly supports the theoretical predictions.

Tables A1, A2, and A3 in the Appendix report the three model specifications with progressively added controls.

Table A4 in the Appendix reports the results using alternative proxies for prejudice. In columns (1) and (2), we replace the 1981 referendum with the share of votes against divorce rights in the 1974 referendum. Columns (3)–(6) report estimates using electoral

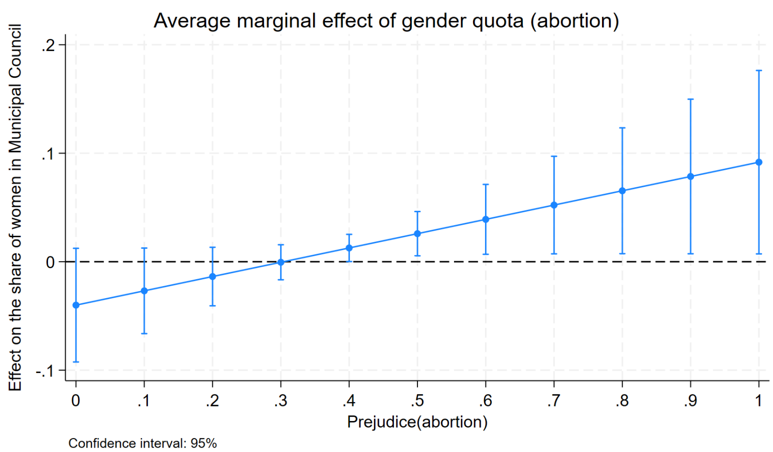
Table 3: Baseline model: interaction between gender quota and prejudice

	(1)	(2)	(3)
Dep. var: Share of women in Municipal Council			
Gender quota	0.012*	0.012*	-0.040
	(0.007)	(0.007)	(0.026)
Share of votes against abortion right (prejudice)		-0.021	-0.148**
		(0.016)	(0.070)
Gender quota × Share of votes against abortion right (prejudice)			0.132*
			(0.068)
Non-urban population (%)	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Male-to-female ratio	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Old-age dependency ratio (%)	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)
Legally separated or divorced (% of 18+)	0.004**	0.004*	0.004*
	(0.002)	(0.002)	(0.002)
Adults (25-64) with upper secondary or tertiary education (%)	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)
Mean household size	-0.028***	-0.028***	-0.027***
	(0.009)	(0.009)	(0.009)
Constant	0.287***	0.291***	0.341***
	(0.042)	(0.041)	(0.047)
Province FE	Yes	Yes	Yes
Observations	6048	6048	6048
Mean Outcome	0.180	0.180	0.180
N clusters	86	86	86

Note: Standard errors clustered at the province level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

support for the Christian Democracy party and the composite prejudice index. In these specifications, the estimated coefficients retain the same sign but are smaller in magnitude and not statistically significant. Finally, in columns (7) and (8), we show the estimates using the education gender gap. The interaction term is positive and statistically significant. Figure A.1 reports the average marginal effects for these alternative proxies and for the

Figure 1: Baseline model: interaction between gender quota and prejudice



composite index. Overall, the pattern is consistent with quotas having larger effects in municipalities with higher levels of prejudice.

5.3 Alternative specifications and robustness checks

The models discussed so far adopt the most demanding clustering strategy for standard errors: provinces are the smallest territorial unit above municipalities, and clustering at a different level than the fixed effects may be preferable. Following the 'rule of thumb' in Cameron and Miller (2015) and, more recently, MacKinnon et al. (2023), we replicate the regressions under alternative assumptions about standard errors. Table 4 presents results with unadjusted standard errors (column 1), heteroskedasticity-robust standard errors (column 2), clustering at the regional level (column 3), and clustering at the provincial level (column 4). The significance of the results remains broadly unaffected by the clustering level or by the choice of fixed effects, whether at the province (panel a) or region level (panel b). This provides reassurance that the main findings are not driven by a particular specification of clustering and fixed effects.

Table A5 reports two additional robustness checks. We shift attention to members of the executive committee, appointed by the mayor (column (1)), and to the gender of the mayor (column (2)), whose election was not subject to quotas. Results suggest some spillover from council composition to other municipal offices. However, these effects are

Table 4: Robustness check: alternative clustering

	(1)	(2)	(3)	(4)
Panel A: Dep. var: Share of women in Municipal Council				
Gender quota	-0.040*	-0.040	-0.040*	-0.040
	(0.024)	(0.025)	(0.022)	(0.026)
Prejudice (abortion)	-0.147**	-0.147**	-0.147**	-0.147**
	(0.057)	(0.062)	(0.053)	(0.069)
Gender quota × Prejudice (abortion)	0.131**	0.131**	0.131**	0.131*
	(0.057)	(0.062)	(0.052)	(0.068)
Other controls	Yes	Yes	Yes	Yes
Province FE	Yes	Yes	Yes	Yes
Observations	6025	6025	6025	6025
N clusters	.	.	15	86
Panel B: Dep.var: Share of women in Municipal Council				
Gender quota	-0.036	-0.036	-0.036	-0.036
	(0.024)	(0.025)	(0.022)	(0.025)
Prejudice (abortion)	-0.171***	-0.171***	-0.171***	-0.171**
	(0.057)	(0.061)	(0.051)	(0.069)
Gender quota × Prejudice (abortion)	0.126**	0.126**	0.126**	0.126*
	(0.058)	(0.061)	(0.052)	(0.065)
Other controls	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Observations	6025	6025	6025	6025
N clusters	.	.	15	86

Note: Other controls include the non-urban population share, male-to-female ratio, old age dependency ratio, legally separated or divorced adults (%), adults with upper secondary or tertiary education (%), and mean family size, all drawn from the 1991 census. Standard errors are OLS standard in column (1), heteroskedasticity-robust in column (2), clustered at the regional level in column (3), and clustered at the provincial level in column (4). Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

not statistically significant. Still, the direction and magnitude of the coefficients mirror those found at the council level, suggesting limited but possible indirect influence.

Finally, note that about 95% of municipalities adopted gender quotas. In principle, this does not undermine the validity of the regression analysis, as the descriptive statistics of the main variable are not markedly different (see Table A6). To further address this imbalance, we implement a Monte Carlo-based balanced resampling procedure: in each iteration, we retain all municipalities without quotas (Gender Quota = 0) and randomly undersample those with quotas (Gender Quota = 1) to construct more balanced samples

(20%, 30%, and 40% of untreated municipalities). We then re-estimate the regressions and summarize the coefficients of interest across iterations. Figure A.2 plots the distribution of estimated coefficients and their p-values across different levels of undersampling. The main results are confirmed even in smaller, more balanced samples.[11]

6 The effects of Law 215/2012

The second test of our theoretical prediction exploits the population threshold in the introduction of gender quotas with Law 215/2012. As we are interested in the long-lasting dynamics of the bias, we focus on both the first and second electoral cycles after the introduction of the quota.

6.1 Empirical strategy

To investigate the causal effect of the 2012 reform on female political representation, we implement a sharp regression discontinuity design exploiting the population threshold at 5,000 inhabitants that determines the applicability of the reform. To estimate heterogeneous treatment effects, we follow the framework for RD effects with pretreatment covariates proposed by Calonico et al. (2025b). Our baseline specification is given by:

$$\begin{aligned} y_m = {} & \alpha + \beta T_m + \theta\, Pop_m + \kappa\, T_m Pop_m + \zeta z_m \\ & + \gamma\, w_m + \delta\, T_m w_m + \lambda\, Pop_m w_m + \mu\, T_m Pop_m w_m + \varphi z_m w_m + \varepsilon_m, \end{aligned} \tag{2}$$

where y_m denotes the share of women in the municipal council of municipality m. The running variable, Pop_m, is municipal population, centered at the 5,000 cutoff. $T_m = \mathbf{1}\{Pop_m \geq 0\}$ denotes the treatment status, which equals one for municipalities above the cutoff. To study heterogeneity in treatment effects, we allow the impact of the reform to vary with a pre-treatment measure of prejudice, w_m, which is centered at its mean. To increase precision, we control for a vector of predetermined municipal characteristics, z_m, including altitude, an indicator for coastal municipalities, and the degree of urbanization. [12] ε_m denotes the error term.

[11] We report the results for 1,000 iterations, which is generally considered sufficient. Replicating the procedure with different numbers of iterations yields substantially unchanged results.

[12] With covariate adjustment, the regression includes both z_m and their interactions with the heterogeneity covariate, $z_m \times w_m$, while z_m are not interacted with treatment status (see Calonico et al. (2025a)).

This specification corresponds to a local linear RD with full interactions between treatment status, the running variable, and the heterogeneity covariate w_m. The coefficient β captures the local average treatment effect of the reform at the cutoff for municipalities with average levels of prejudice ($w_m = 0$). The coefficient δ measures how the treatment effect varies with the pre-treatment prejudice measure. The remaining interaction terms allow both the slope of the outcome with respect to population and its dependence on prejudice to differ on either side of the cutoff.

Estimation is conducted using standard RD procedures with observations restricted to a neighborhood around the cutoff and a uniform kernel. Inference for heterogeneous treatment effects follows the framework proposed by Calonico et al. (2025b), which delivers valid confidence intervals in RD settings with covariate interactions.

6.2 Results

We now turn to the results for the 2012 reform introducing gender quotas and gender-conditioned double preference voting in municipalities with more than 5,000 inhabitants. As discussed above, the strength of the analysis of this reform is that the population threshold allows for a regression discontinuity design. However, the fact that Law 215/2012 is still in place implies that we cannot exploit the removal of the quota to cleanly separate possible heterogeneous mechanical effects from changes driven by voters' belief updating. This distinction is crucial for the interpretation of heterogeneous effects.

Our theoretical framework predicts that exposure to female politicians reduces gender bias through learning. Hence, it predicts heterogeneous dynamics after voters have observed women in office. Therefore the heterogeneous effects that are (i) induced by gender quotas and (ii) due to a reduction of the prejudice should materialize primarily in subsequent electoral cycles. However, we may also observe heterogeneous effects in the first post-reform election, due to the direct constraint imposed by the quota and the double-preference rule, rather than to a change in voter attitudes. For this reason, we separately analyze the first and the second electoral cycles following the reform.

Table 5 reports the results for the first post-reform electoral cycle (2013–2017). Across all specifications, the reform has a large and precisely estimated positive effect on the share of women elected to municipal councils. The point estimates imply an increase of approximately 13–15 percentage points at the cutoff, confirming that the reform substantially altered electoral outcomes in treated municipalities. By contrast, the interaction between treatment status and the various proxies of pre-treatment prejudice is generally small and imprecisely estimated. This pattern is consistent with the interpretation that, in the first electoral cycle, the effects of the reform are predominantly mechanical and do not yet reflect learning-driven changes in voter behavior.

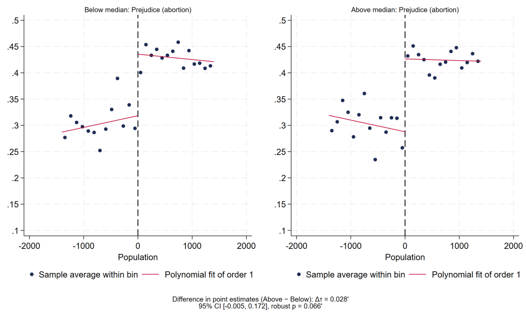
The picture changes markedly when we turn to the second electoral cycle. Table 6 shows that the positive effect of the reform on female representation persists and remains statistically significant. More importantly, the interaction terms between treatment and pre-treatment prejudice are now positive, sizable, and statistically significant across most proxies. Municipalities characterized by higher initial prejudice experience substantially larger gains in women's representation relative to less biased municipalities.

To further investigate the impact of pre-existing bias on the effects of quotas, we define *Low*/*High* prejudice groups (below vs above the median of each raw proxy). Then, using the RD heterogeneity methods proposed by Calonico et al. (2025b), we estimate β_{Low} and β_{High} (with group-specific optimal bandwidths) and test

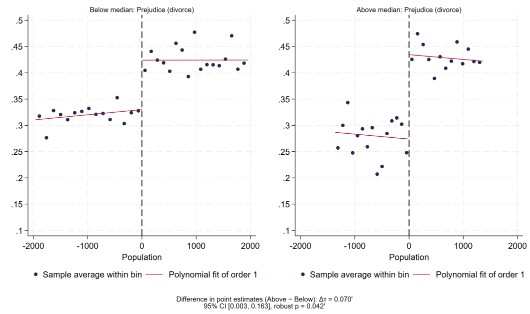
$$\Delta\tau = \beta_{\text{High}} - \beta_{\text{Low}} = 0.$$

Figure 2 illustrates these heterogeneous effects graphically. It also reports the difference in point estimates ($\Delta\tau$) between high- and low-prejudice municipalities together with robust bias-corrected confidence intervals and p-values, following Calonico et al. (2025a). In municipalities with high levels of pre-treatment prejudice, the discontinuity at the population threshold is clearly visible and substantially larger than in low-prejudice contexts. This pattern is consistent across alternative measures of bias, including referendum outcomes, political preferences, education gaps, and the composite index.

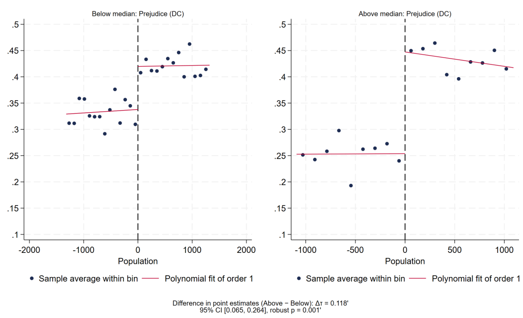
Taken together, these results closely match the predictions of the model. While the first post-reform election mainly captures the direct, mechanical impact of the quota, the emergence of stronger heterogeneous effects in the second electoral cycle points to a learning mechanism. Increased exposure to female politicians reduces gender bias more rapidly where initial prejudice is stronger, leading to persistent and increasingly unequal gains in women's political representation across municipalities.



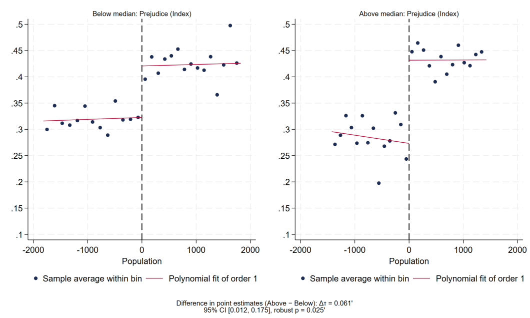
(a) Low vs. high prejudice (abortion)



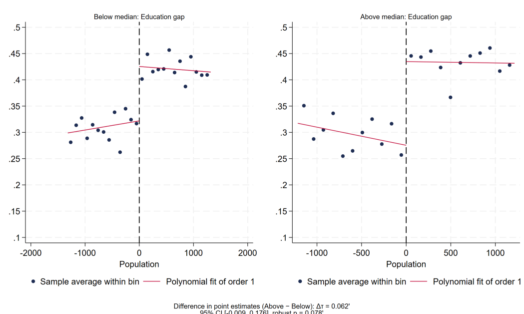
(b) Low vs. high prejudice (divorce)



(c) Low vs. high prejudice (DC)



(d) Low vs. high prejudice (index)



(e) Low vs. high prejudice (education gender gap)

Figure 2: Regression Discontinuity Estimates by Pre-Treatment Prejudice Measures

Table 5: Effects of the 2012 reform in the first post-reform electoral cycle (2013–2017)

	(1)	(2)	(3)	(4)	(5)
	Dep. var: Share of women in Municipal Council (1st electoral cycle (2013-2017))				
Treatment	0.137***	0.136***	0.146***	0.138***	0.140***
	(0.0245)	(0.0240)	(0.0238)	(0.0239)	(0.0244)
Treatment × Prejudice (abortion)	0.0390				
	(0.173)				
Treatment × Prejudice (divorce)		0.238			
		(0.150)			
Treatment × Prejudice (DC)			0.377*		
			(0.211)		
Treatment × Prejudice (index)				0.288	
				(0.194)	
Treatment × Education gap					0.0916
					(0.300)
Robust CI (Treatment)	[0.113 ; 0.209]	[0.112 ; 0.206]	[0.122 ; 0.216]	[0.114 ; 0.208]	[0.112 ; 0.208]
Robust CI (Interaction)	[-0.328 ; 0.350]	[-0.137 ; 0.450]	[-0.048 ; 0.781]	[-0.168 ; 0.592]	[-0.441 ; 0.736]
Bandwidth	1304.1	1304.1	1304.1	1304.1	1304.1
Obs (left)	507	507	507	507	507
Obs (right)	342	342	342	342	342
Mean Outcome	0.324	0.324	0.324	0.324	0.324
Geographical Controls	Yes	Yes	Yes	Yes	Yes

Note: The table reports regression discontinuity estimates of the effect of the 2012 reform on the share of women elected to municipal councils in the first post-reform electoral cycle. All regressions are estimated with a uniform kernel and optimal bandwidth. Robust bias-corrected standard errors in parentheses. Significance levels (based on robust p-values): * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6: Effects of the 2012 reform in the second post-reform electoral cycle

	(1)	(2)	(3)	(4)	(5)
	Dep. var: Share of women in Municipal Council (2nd electoral cycle)				
Treatment	0.128***	0.130***	0.143***	0.132***	0.131***
	(0.0226)	(0.0220)	(0.0216)	(0.0220)	(0.0228)
Treatment × Prejudice (abortion)	0.115*				
	(0.202)				
Treatment × Prejudice (divorce)		0.286**			
		(0.158)			
Treatment × Prejudice (DC)			0.503***		
			(0.175)		
Treatment × Prejudice (index)				0.380***	
				(0.202)	
Treatment × Education gap					0.193**
					(0.302)
Robust CI (Treatment)	[0.079 ; 0.167]	[0.085 ; 0.171]	[0.107 ; 0.191]	[0.089 ; 0.176]	[0.090 ; 0.179]
Robust CI (Interaction)	[-0.055 ; 0.736]	[0.063 ; 0.682]	[0.306 ; 0.992]	[0.142 ; 0.934]	[0.053 ; 1.235]
Bandwidth	1433.7	1433.7	1433.7	1433.7	1433.7
Obs (left)	566	566	566	566	566
Obs (right)	366	366	366	366	366
Mean Outcome	0.352	0.352	0.352	0.352	0.352
Geographical Controls	Yes	Yes	Yes	Yes	Yes

Note: The table reports regression discontinuity estimates of the effect of the 2012 reform on the share of women elected to municipal councils in the second post-reform electoral cycle. All regressions are estimated with a uniform kernel and optimal bandwidth. Robust bias-corrected standard errors in parentheses. Significance levels (based on robust p-values): * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

6.3 Falsification and robustness checks

We conduct a series of falsification tests and robustness checks to assess the validity of our regression discontinuity estimates.

A crucial assumption for the validity of the RD design is the absence of precise manipulation around the cutoff. We assess this assumption using the density continuity test for the running variable proposed by Cattaneo et al. (2020). The value of the statistic is -1.446 (p-value $= 0.148$), indicating no evidence of sorting around the cutoff. Figure 3 reports the estimated density.

Second, we assess the balancing of predetermined municipal characteristics and, crucially, the pre-treatment proxies of prejudice. As shown in Table A7, we find no discontinuities at the cutoff in our set of control covariates, including altitude, coastal municipality, and degree of urbanization. Table A8 similarly shows no discontinuities in the pre-treatment prejudice proxies used in the heterogeneity analysis.

Third, we examine whether the reform affects related political outcomes. We estimate RD specifications using alternative dependent variables, namely female representation in the municipal executive (Table A9 in the Appendix) and the probability of having a female mayor in the subsequent electoral cycle (Table A10 in the Appendix). We find no statistically significant effects on either the share of women in the executive or the probability of a female mayor.

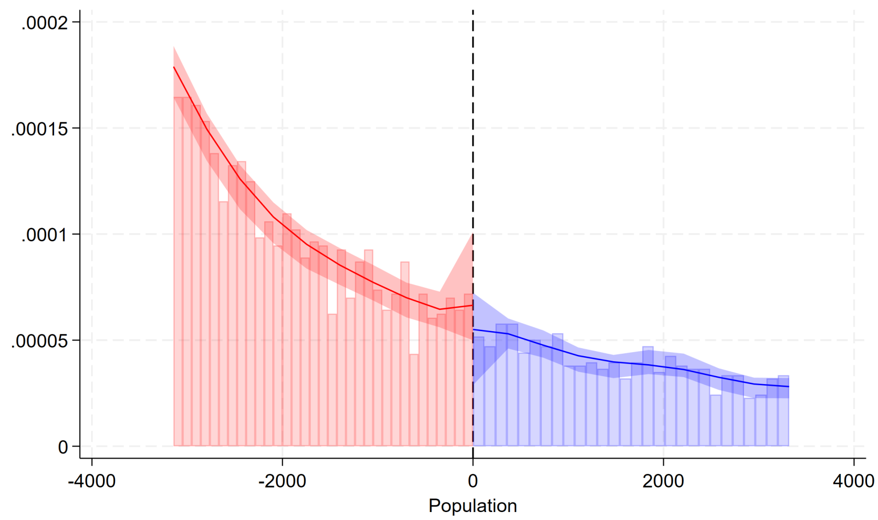
We further assess the sensitivity of our results to alternative functional form assumptions. In particular, we estimate specifications using a second-order polynomial in the running variable. Table A11 shows that our findings remain qualitatively similar to the baseline local linear results.

Finally, we verify the robustness of our findings to alternative RD estimation choices. We re-estimate Equation 2 using a triangular kernel, which gives more weight to observations near the cutoff. Table A12 in the Appendix shows that the estimates are very similar in magnitude to the baseline estimates using a uniform kernel.

7 Conclusions

This paper studies how voter prejudice against women in politics evolves over time and how gender quotas can accelerate this process through learning. Building on a Bayesian updating framework, we model prejudice as biased beliefs about female politicians' com-

Figure 3: Estimated density of the running variable



petence that are revised through exposure to women in office. The model delivers two key predictions: first, gender quotas may have effects that persist even after their removal; second, these effects should be stronger in contexts where initial prejudice is higher.

We test these predictions using two distinct institutional reforms in Italian municipal elections. The temporary gender quota introduced in 1993 provides a setting in which we can isolate learning effects from mechanical ones by examining electoral outcomes after the quota was abolished. We show that municipalities exposed to quotas experienced a persistent increase in women's representation, and that this effect was significantly stronger where pre-existing gender bias was more pronounced. These findings offer clear support for the learning mechanism implied by the model.

We then turn to the 2012 reform, which introduced permanent gender quotas and gender-conditioned double preference voting above a population threshold. In this context, separating mechanical effects from belief updating is more challenging. Using a regression discontinuity design with heterogeneous treatment effects, we show that the reform substantially increased women's representation in both the first and the second post-reform electoral cycles. Crucially, while heterogeneous effects by pre-treatment prejudice are weak or absent in the first cycle, they become large and statistically significant in the second cycle. This dynamic pattern is consistent with the idea that exposure to female politicians gradually reduces voter bias, particularly in more prejudiced environments.

Taken together, our results provide novel empirical evidence that gender quotas do more than mechanically alter electoral outcomes. By increasing the visibility of women in political office, quotas can change voter beliefs in a persistent way, generating larger and more durable gains where barriers to female political participation are initially stronger. This mechanism helps reconcile the short-run effectiveness of quotas with their long-run impacts on political representation.

From a policy perspective, our findings suggest that affirmative action policies may be most effective precisely where resistance to female leadership is greatest. Rather than reinforcing stereotypes, as some opposing affirmative action may argue, quotas can trigger a belief update that will reduce gender bias over time and build a more even playing field in political competition. Our analysis highlights the importance of looking at the dynamic and heterogeneity of responses when evaluating institutional reforms like those that promote gender equality.

References

Andreoli, F., Manzoni, E., and Margotti, M. (2022). Women at work: Gender quotas, municipal elections and local spending. *European Journal of Political Economy*, 75:102175.

Anzia, S. F. and Berry, C. R. (2011). The jackie (and jill) robinson effect: Why do congresswomen outperform congressmen? *American Journal of Political Science*, 55(3):478–493.

Bagde, S., Epple, D., and Taylor, L. (2016). Does affirmative action work? caste, gender, college quality, and academic success in india. *American Economic Review*, 106(6):1495–1521.

Baltrunaite, A., Bello, P., Casarico, A., and Profeta, P. (2014). Gender quotas and the quality of politicians. *Journal of Public Economics*, 118:62–74.

Baltrunaite, A., Casarico, A., Profeta, P., and Savio, G. (2019). Let the voters choose women. *Journal of Public Economics*, 180:104085.

Beaman, L., Chattopadhyay, R., Duflo, E., Pande, R., and Topalova, P. (2009). Powerful women: does exposure reduce bias? *The Quarterly journal of economics*, 124(4):1497–1540.

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

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.

Bordalo, P., Coffman, K., Gennaioli, N., and Shleifer, A. (2016). Stereotypes. *The Quarterly Journal of Economics*, 131(4):1753–1794.

Bozzano, M. (2017). On the historical roots of women's empowerment across italian provinces: religion or family culture? *European Journal of Political Economy*, 49:24–46.

Braga, M. and Scervini, F. (2017). The performance of politicians: The effect of gender quotas. *European Journal of Political Economy*, 46:1–14.

Broockman, D. E. (2014). Do female politicians empower women to vote or run for office? a regression discontinuity approach. *Electoral Studies*, 34:190–204.

Calonico, S., Cattaneo, M. D., Farrell, M. H., Palomba, F., and Titiunik, R. (2025a). rdhte: Conditional average treatment effects in rd designs. *arXiv preprint arXiv:2507.01128*.

Calonico, S., Cattaneo, M. D., Farrell, M. H., Palomba, F., and Titiunik, R. (2025b). Treatment effect heterogeneity in regression discontinuity designs. *arXiv preprint arXiv:2503.13696*.

Cameron, C. A. and Miller, D. L. (2015). A practitioner's guide to cluster-robust inference. *Journal of Human Resources*, 50(2):317–372.

Campbell, D. E. and Wolbrecht, C. (2006). See jane run: Women politicians as role models for adolescents. *The Journal of Politics*, 68(2):233–247.

Cattaneo, M. D., Jansson, M., and Ma, X. (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association*, 115(531):1449–1455.

Cella, M. and Manzoni, E. (2023). Gender bias and women's political performance. *European Journal of Political Economy*, 77:102314.

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

Clark, J. (1991). Getting there: Women in political office. *The Annals of the American Academy of Political and Social Science*, 515(1):63–76.

Craigmile, P. F. and Tirrerington, D. (1997). Parameter estimation for finite mixtures of uniform distributions. *Communications in Statistics-Theory and Methods*, 26(8):1981–1995.

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

Dolan, K. A. (2004). *Voting for women: How the public evaluates women candidates.* Westview Press.

Duflo, E. (2005). Why political reservations? *Journal of the European Economic Association*, 3(2-3):668–678.

Falk, E. and Kenski, K. (2006). Issue saliency and gender stereotypes: Support for women as presidents in times of war and terrorism. *Social Science Quarterly*, 87(1):1–18.

Fox, R. L. and Lawless, J. L. (2011). Gendered perceptions and political candidacies: A central barrier to women's equality in electoral politics. *American journal of political science*, 55(1):59–73.

Fox, R. L., Lawless, J. L., and Feeley, C. (2001). Gender and the decision to run for office. *Legis. Stud. Q.*, 26:411.

Franceschet, S., Krook, M. L., and Piscopo, J. M. (2012). *The impact of gender quotas.* OUP USA.

Gilardi, F. (2015). The temporary importance of role models for women's political representation. *American Journal of Political Science*, 59(4):957–970.

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

Gupta, A. and Miyawaki, T. (1978). On a uniform mixture model. *Biometrical Journal*, 20(7-8):631–637.

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

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

MacKinnon, J. G., Ørregaard Nielsen, M., and Webb, M. D. (2023). Cluster-robust inference: A guide to empirical practice. *Journal of Econometrics*, 232(2):272–299.

Pearson, K. and Dancey, L. (2011). Elevating women's voices in congress: Speech participation in the house of representatives. *Political Research Quarterly*, 64(4):910–923.

Pearson, K. and McGhee, E. (2013). What it takes to win: Questioning "gender neutral" outcomes in us house elections. *Politics & Gender*, 9(4):439–462.

Porter, C. and Serra, D. (2020). Gender differences in the choice of major: The importance of female role models. *American Economic Journal: Applied Economics*, 12(3):226–254.

Sanbonmatsu, K. (2002). Gender stereotypes and vote choice. *American Journal of Political Science*, 46(1):20–34.

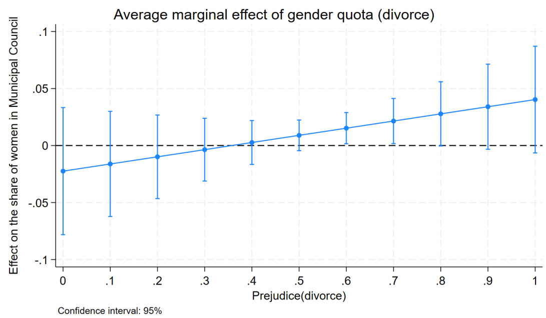
Tversky, A. and Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological review*, 90(4):293.

Volden, C., Wiseman, A. E., and Wittmer, D. E. (2013). When are women more effective lawmakers than men? *American Journal of Political Science*, 57(2):326–341.

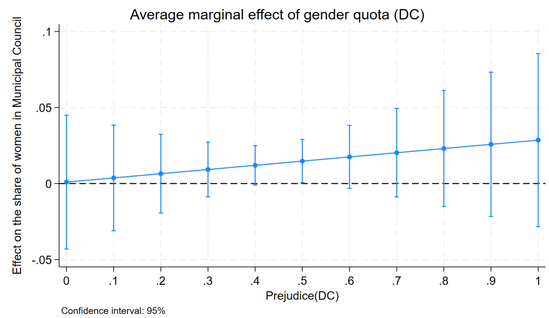
Welch, S. (1977). Women as political animals? a test of some explanations for male-female political participation differences. *American journal of political science*, pages 711–730.

A Appendix

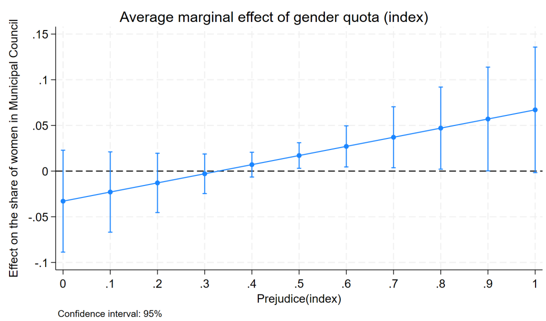
A.1 Additional figures



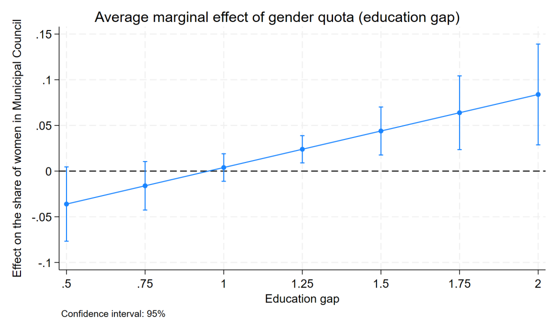
(A) Divorce



(B) Christian Democrats (DC)



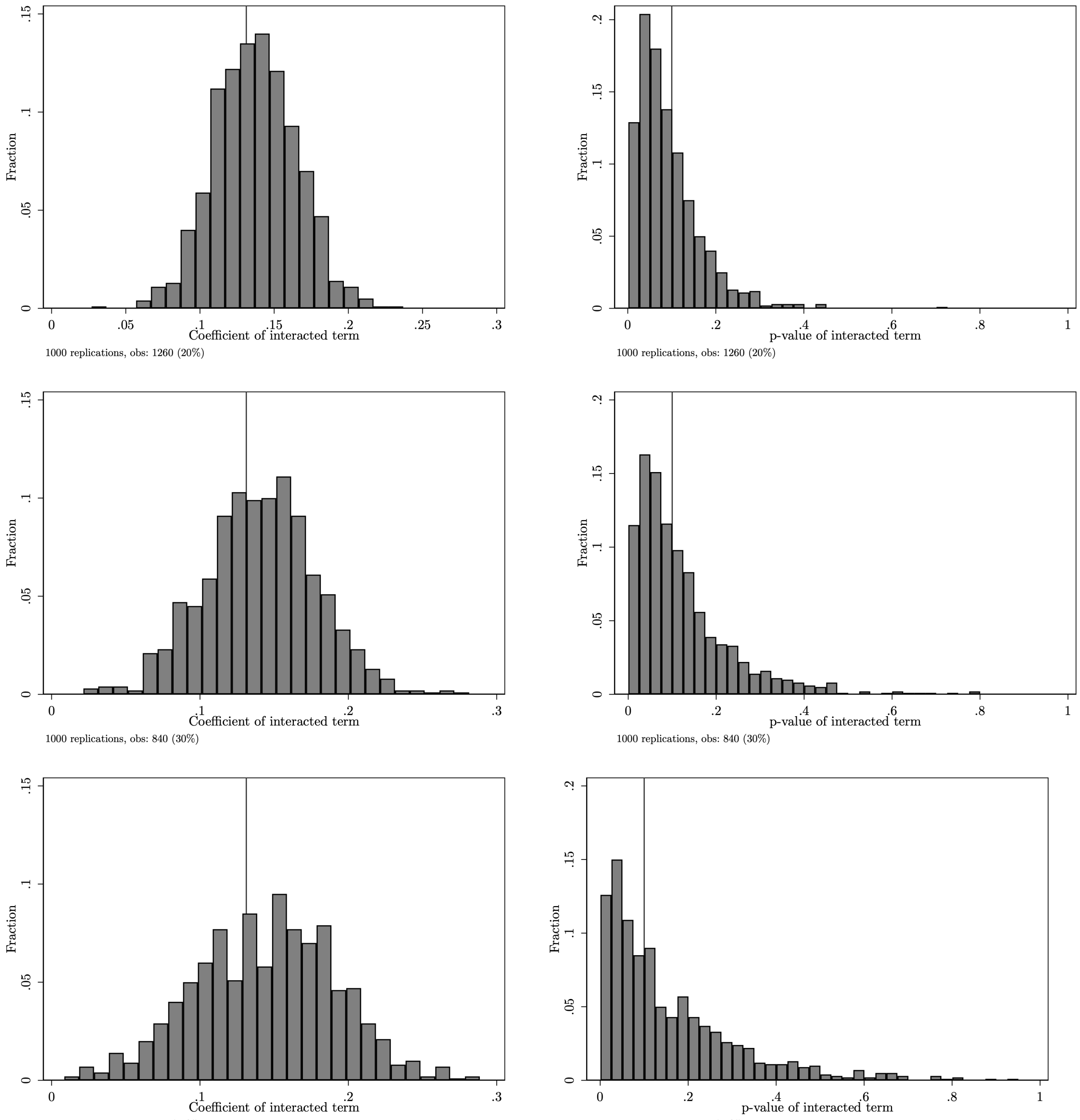
(C) Index



(D) Education gap

Figure A.1: Average marginal effects with different proxies

Figure A.2: Undersampling: distribution of the coefficient of the interaction term (left) and relative p-values (right) over 1,000 replications.



Note: Vertical lines are the estimated coefficient in the baseline model (left panels) and the conventional statistical significance (p-valu = 0.1).

Additional tables

Table A1: Effect of gender quota

	(1)	(2)	(3)	(4)
Dep.var: Share of women in Municipal Council				
Gender quota	0.038***	0.029***	0.014**	0.012*
	(0.007)	(0.007)	(0.006)	(0.007)
Non-urban population (%)		0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)
Male-to-female ratio		-0.000	-0.001**	-0.000
		(0.000)	(0.000)	(0.000)
Old-age dependency ratio (%)		-0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)
Legally separated or divorced (% of 18+)		0.015***	0.007***	0.005**
		(0.004)	(0.002)	(0.002)
Adults with upper sec. or tert. educ. (%)		-0.000	0.000	-0.000
		(0.000)	(0.000)	(0.000)
Mean household size		-0.054***	-0.023***	-0.027***
		(0.009)	(0.008)	(0.009)
Constant	0.143***	0.325***	0.278***	0.282***
	(0.009)	(0.044)	(0.041)	(0.043)
Region FE	No	No	Yes	No
Province FE	No	No	No	Yes
Observations	6025	6025	6025	6025
Mean Outcome	0.180	0.180	0.180	0.180
N clusters	86	86	86	86

Note: Standard errors reported in parentheses are clustered at the province level. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A2: Effect of gender quota and prejudice

	(1)	(2)	(3)	(4)
Dep.var: Share of women in Municipal Council				
Gender quota	0.037***	0.029***	0.013**	0.012*
	(0.007)	(0.007)	(0.006)	(0.007)
Prejudice (abortion)	-0.063***	-0.002	-0.050***	-0.021
	(0.023)	(0.019)	(0.015)	(0.017)
Non-urban population (%)		0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)
Male-to-female ratio		-0.000	-0.001**	-0.000
		(0.000)	(0.000)	(0.000)
Old-age dependency ratio (%)		-0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)
Legally separated or divorced (% of 18+)		0.015***	0.005**	0.004*
		(0.004)	(0.002)	(0.002)
Adults with upper sec. or tert. educ. (%)		-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)
Mean household size		-0.054***	-0.022***	-0.026***
		(0.009)	(0.008)	(0.009)
Constant	0.168***	0.327***	0.295***	0.287***
	(0.013)	(0.044)	(0.040)	(0.042)
Region FE	No	No	Yes	No
Province FE	No	No	No	Yes
Observations	6025	6025	6025	6025
Mean Outcome	0.180	0.180	0.180	0.180
N clusters	86	86	86	86

Note: Other controls include the non-urban population share, male-to-female ratio, and mean family size, all drawn from the 1991 census. Standard errors are reported in parentheses. In column (4), standard errors are clustered at the province level. In column (5), standard errors are clustered at the region level. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A3: Baseline model: interaction between gender quota and prejudice

	(1)	(2)	(3)	(4)
Dep. var: Share of women in Municipal Council				
Gender quota	-0.001	-0.012	-0.036	-0.040
	(0.028)	(0.026)	(0.025)	(0.026)
Prejudice (abortion)	-0.156**	-0.102	-0.171**	-0.147**
	(0.070)	(0.068)	(0.069)	(0.069)
Gender quota × Prejudice (abortion)	0.097	0.104	0.126*	0.131*
	(0.070)	(0.066)	(0.065)	(0.068)
Non-urban population (%)		0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)
Male-to-female ratio		-0.000	-0.001**	-0.000
		(0.000)	(0.000)	(0.000)
Old-age dependency ratio (%)		-0.000	0.000	0.000
		(0.000)	(0.000)	(0.000)
Legally separated or divorced (% of 18+)		0.015***	0.005**	0.004*
		(0.004)	(0.002)	(0.002)
Adults with upper sec. or tert. educ. (%)		-0.000	-0.000	-0.000
		(0.000)	(0.000)	(0.000)
Mean household size		-0.054***	-0.022***	-0.026***
		(0.009)	(0.008)	(0.009)
Constant	0.205***	0.365***	0.342***	0.336***
	(0.030)	(0.052)	(0.048)	(0.048)
Region FE	No	No	Yes	No
Province FE	No	No	No	Yes
Observations	6025	6025	6025	6025
Mean Outcome	0.180	0.180	0.180	0.180
N clusters	86	86	86	86

Note: Other controls include the non-urban population share, male-to-female ratio, and mean family size, all drawn from the 1991 census. Standard errors are reported in parentheses. In column (4), standard errors are clustered at the province level. In column (5), standard errors are clustered at the region level. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A4: Baseline model: interaction between gender quota and prejudice

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dep. var: Share of women in Municipal Council								
Gender quota	0.012*	-0.022	0.012*	0.001	0.012*	-0.033	0.012*	-0.076**
	(0.007)	(0.028)	(0.007)	(0.022)	(0.007)	(0.028)	(0.007)	(0.035)
Prejudice(divorce)	-0.005	-0.065						
	(0.016)	(0.054)						
Gender quota × Prejudice(divorce)		0.063						
		(0.050)						
Prejudice(DC)			-0.037***	-0.063				
			(0.014)	(0.050)				
Gender quota × Prejudice(DC)				0.028				
				(0.049)				
Prejudice(index)					-0.026	-0.122*		
					(0.018)	(0.065)		
Gender quota × Prejudice(index)						0.100		
						(0.061)		
Education gap							-0.010**	-0.087***
							(0.005)	(0.029)
Gender quota × Education gap								0.080**
								(0.031)
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	6036	6036	6045	6045	6048	6048	6040	6040
Mean Outcome	0.180	0.180	0.180	0.180	0.180	0.180	0.180	0.180
N clusters	86	86	86	86	86	86	86	86

Note: Standard errors clustered at the province level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A5: Robustness checks

	(1)	(2)
Dep. var: Share of women in:	**Municipal Executive**	**Mayor**
Gender quota	0.009	-0.048
	(0.048)	(0.050)
Prejudice(abortion)	-0.038	-0.169
	(0.125)	(0.118)
Gender quota × Prejudice(abortion)	0.001	0.120
	(0.120)	(0.116)
Constant	0.331***	0.050
	(0.093)	(0.089)
Other controls	Yes	Yes
Province FE	Yes	Yes
Observations	6019	6014
Mean Outcome	0.163	0.0700
N clusters	86	86

Note: Other controls include the non-urban population share, male-to-female ratio, old age dependency ratio, legally separated or divorced adults (%), adults with upper secondary or tertiary education (%), and mean family size, all drawn from the 1991 census. Standard errors clustered at the province level in parentheses. Significance levels: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A6: Descriptive statistics by quota

	Mean	Std. dev.	Obs.	Mean	Std. dev.	Obs.
		No quota			Quota	
Share of women in Municipal Council (1st election post gender quota)	0.143	0.119	252	0.182	0.109	5773
Share of women in Municipal Executive (1st election post gender quota)	0.126	0.206	252	0.164	0.214	5767
Female Mayor (1st election post gender quota)	0.052	0.222	252	0.071	0.257	5762
Municipality ever held election under gender quota	0.000	0.000	252	1.000	0.000	5773
Share of votes against abortion right (prejudice)	0.395	0.115	252	0.376	0.125	5773
Share of votes against divorce right (prejudice)	0.554	0.147	251	0.495	0.153	5762
Non-urban population (%)	18.241	19.432	252	22.045	20.071	5773
Male-to-female ratio	97.128	6.564	252	96.510	6.336	5773
Old-age dependency ratio (%)	28.782	12.231	252	29.584	13.053	5773
Legally separated or divorced (% of 18+)	1.009	0.843	252	1.330	0.871	5773
Adults (25-64) with upper secondary or tertiary education (%)	18.925	6.813	252	19.840	6.607	5773
Mean household size	2.787	0.378	252	2.692	0.358	5773

Table A7: Balancing of predetermined covariates

	(1) Altitude	(2) Coastal Municipality	(3) Urbanization
Treatment	-10.57	0.0331	0.0860
	(33.00)	(0.0381)	(0.0525)
Robust CI	[-85.559 ; 67.298]	[-0.054 ; 0.124]	[-0.027 ; 0.218]
Robust p-value	0.815	0.444	0.127
Bandwidth	987.5	1652.5	2197.1
Observations	642	1090	1519
Mean Outcome	257.0	0.0817	0.446

Note: Standard errors in parentheses. Significance levels * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A8: Balancing of prejudice measures

	(1) Abortion	(2) Divorce	(3) DC	(4) Index	(5) Education gap
Treatment	0.000791	0.0177	0.0106	0.00993	0.0127
	(0.0174)	(0.0207)	(0.0157)	(0.0164)	(0.0116)
Robust CI	[-0.041 ; 0.041]	[-0.031 ; 0.065]	[-0.022 ; 0.050]	[-0.024 ; 0.051]	[-0.014 ; 0.040]
Robust p-value	0.998	0.480	0.434	0.489	0.348
Bandwidth	1159.2	1306.4	1337.6	1217.7	966.3
Observations	761	857	880	808	634
Mean Outcome	0.383	0.501	0.332	0.404	1.020

Note: Standard errors in parentheses. Significance levels * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A9: Effects of the 2012 reform on female representation in the municipal executive

	(1)	(2)	(3)	(4)	(5)
	Dep. var: Share of women in Municipal Executive (2nd electoral cycle)				
Treatment	0.0316	0.0341	0.0229	0.0308	0.0287
	(0.0332)	(0.0336)	(0.0325)	(0.0329)	(0.0333)
Treatment × Prejudice (abortion)	-0.108				
	(0.293)				
Treatment × Prejudice (divorce)		-0.00527			
		(0.242)			
Treatment × Prejudice (DC)			-0.238		
			(0.304)		
Treatment × Prejudice (index)				-0.112	
				(0.313)	
Treatment × Education gap					-0.194
					(0.427)
Robust CI (Treatment)	[-0.056 ; 0.074]	[-0.056 ; 0.075]	[-0.069 ; 0.058]	[-0.057 ; 0.072]	[-0.056 ; 0.074]
Robust CI (Interaction)	[-0.706 ; 0.443]	[-0.645 ; 0.302]	[-0.983 ; 0.209]	[-0.864 ; 0.365]	[-1.006 ; 0.666]
Bandwidth	800.0	800.0	800.0	800.0	800.0
Obs (left)	289	289	289	289	289
Obs (right)	224	224	224	224	224
Mean Outcome	0.476	0.476	0.476	0.476	0.476
Geographical Controls	Yes	Yes	Yes	Yes	Yes

Note: The table reports regression discontinuity estimates of the effect of the 2012 reform on the share of women in the municipal executive in the second post-reform electoral cycle. All regressions are estimated with a uniform kernel and optimal bandwidth. Robust bias-corrected standard errors in parentheses. Significance levels (based on robust p-values): * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A10: Effects of the 2012 reform on the probability of electing a female mayor

	(1)	(2)	(3)	(4)	(5)
	Dep. var: Female mayor (2nd electoral cycle)				
Treatment	0.0198	0.0199	0.0314	0.0234	0.0236
	(0.0661)	(0.0672)	(0.0661)	(0.0672)	(0.0679)
Treatment × Prejudice (abortion)	-0.00899				
	(0.613)				
Treatment × Prejudice (divorce)		-0.231			
		(0.467)			
Treatment × Prejudice (DC)			0.187		
			(0.557)		
Treatment × Prejudice (index)				-0.103	
				(0.610)	
Treatment × Education gap					0.372
					(0.799)
Robust CI (Treatment)	[-0.157 ; 0.102]	[-0.154 ; 0.109]	[-0.150 ; 0.109]	[-0.152 ; 0.112]	[-0.145 ; 0.121]
Robust CI (Interaction)	[-0.931 ; 1.472]	[-1.084 ; 0.746]	[-1.249 ; 0.934]	[-1.281 ; 1.109]	[-1.303 ; 1.828]
Bandwidth	1856.3	1856.3	1856.3	1856.3	1856.3
Obs (left)	770	770	770	770	770
Obs (right)	459	459	459	459	459
Mean Outcome	0.164	0.164	0.164	0.164	0.164
Geographical Controls	Yes	Yes	Yes	Yes	Yes

Note: The table reports regression discontinuity estimates of the effect of the 2012 reform on the probability of having a female mayor in the second post-reform electoral cycle. All regressions are estimated with a uniform kernel and optimal bandwidth. Robust bias-corrected standard errors in parentheses. Significance levels (based on robust p-values): * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A11: Alternative – Second order polynomial in population

	(1)	(2)	(3)	(4)	(5)
	Dep. var: Share of women in Municipal Council (2nd electoral cycle)				
Treatment	0.133***	0.136***	0.152***	0.139***	0.141***
	(0.0232)	(0.0226)	(0.0222)	(0.0226)	(0.0233)
Treatment × Prejudice (abortion)	0.241				
	(0.208)				
Treatment × Prejudice (divorce)		0.377**			
		(0.162)			
Treatment × Prejudice (DC)			0.615***		
			(0.178)		
Treatment × Prejudice (index)				0.509**	
				(0.207)	
Treatment × Education gap					0.415
					(0.296)
Robust CI (Treatment)	[0.080 ; 0.171]	[0.085 ; 0.174]	[0.102 ; 0.189]	[0.088 ; 0.176]	[0.087 ; 0.179]
Robust CI (Interaction)	[-0.125 ; 0.690]	[0.008 ; 0.642]	[0.164 ; 0.862]	[0.055 ; 0.867]	[-0.234 ; 0.926]
Bandwidth	2359.4	2359.4	2359.4	2359.4	2359.4
Obs (left)	1053	1053	1053	1053	1053
Obs (right)	568	568	568	568	568
Mean Outcome	0.344	0.344	0.344	0.344	0.344
Geographical Controls	Yes	Yes	Yes	Yes	Yes

Note: The table reports regression discontinuity estimates of the effect of the 2012 reform on the share of women elected to municipal councils in the second post-reform electoral cycle using a second order polynomial in population. All regressions are estimated with a uniform kernel and optimal bandwidth. Robust bias-corrected standard errors in parentheses. Significance levels (based on robust p-values): * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A12: Alternative kernel function (triangular)

	(1)	(2)	(3)	(4)	(5)
	Dep. var: Share of women in Municipal Council (2nd electoral cycle)				
Treatment	0.127***	0.130***	0.144***	0.132***	0.133***
	(0.0232)	(0.0226)	(0.0224)	(0.0227)	(0.0235)
Treatment × Prejudice (abortion)	0.188				
	(0.210)				
Treatment × Prejudice (divorce)		0.316**			
		(0.163)			
Treatment × Prejudice (DC)			0.538***		
			(0.175)		
Treatment × Prejudice (index)				0.432**	
				(0.208)	
Treatment × Education gap					0.344*
					(0.304)
Robust CI (Treatment)	[0.082 ; 0.173]	[0.087 ; 0.175]	[0.107 ; 0.195]	[0.091 ; 0.180]	[0.093 ; 0.185]
Robust CI (Interaction)	[-0.107 ; 0.717]	[0.023 ; 0.660]	[0.264 ; 0.950]	[0.090 ; 0.905]	[-0.058 ; 1.134]
Bandwidth	1527.8	1527.8	1527.8	1527.8	1527.8
Obs (left)	594	594	594	594	594
Obs (right)	391	391	391	391	391
Mean Outcome	0.353	0.353	0.353	0.353	0.353
Geographical Controls	Yes	Yes	Yes	Yes	Yes

Note: The table reports regression discontinuity estimates of the effect of the 2012 reform on the share of women elected to municipal councils in the second post-reform electoral cycle. All regressions are estimated with a triangular kernel and optimal bandwidth. Robust bias-corrected standard errors in parentheses. Significance levels (based on robust p-values): * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.