

Gender Equality Through Turnover: Quasi-experimental Evidence from Term Limit Reforms in Italy

Carolina Kansikas and Manuel Bagues*

January 31, 2025

We study whether term limits can accelerate women’s access to top political positions by analyzing two reforms in Italian local elections that extended mayoral term limits from two to three five-year terms. The first reform affected municipalities with fewer than 3,000 inhabitants in 2014, and the second those below 5,000 in 2022 – a period marked by rapid growth in women’s political participation. Using a difference-in-discontinuities design, we find that longer term limits restrict opportunities for early-career politicians, with substantial effects for female representation: the share of female mayors would be 8 percentage points higher without the term limit extensions. The impact is larger in municipalities with more women in lower political positions and where gender quotas for council members are present, suggesting that entry-level quotas can be more effective when paired with policies promoting turnover in top positions.

Keywords: Term limits, female political representation, Italian local elections.

JEL codes: J16, J18, J48, D72

*University of Warwick. Kansikas: carolina.kansikas@warwick.ac.uk, Bagues: manuel.bagues@warwick.ac.uk. We thank for useful comments and suggestions Robert Akerlof, Eleonora Alabrese, James Fenske, Ludovica Gasse, Sophie Hatte, Julien Labonne, Andreas Stegmann, Natalia Zinovyeva and participants in presentations at the University of Warwick, the University of Barcelona, ENS Lyon, 2024 NICEP Conference in Nottingham, 2024 COSME Workshop on Gender Economics, the 2024 CAGE Workshop on Gender and Inequality, the EEA-ESEM 2024 Conference and the EALE 2024 Conference. Kansikas thanks the Yrjö Jahnsson Foundation for their generous support. A previous version of this paper was circulated under the title ‘Can term limits accelerate women’s access to top political positions? Quasi-experimental evidence from Italy’.

1 Introduction

Over recent decades Europe has witnessed a substantial increase in women’s participation in politics. Although gender parity remains elusive, women now constitute 33% of national parliamentarians and 32% of municipal councilors. Yet, despite this progress at entry-level positions - facilitated in part by the adoption of gender quotas - women remain severely underrepresented in top political roles, accounting for only 14% of European prime ministers and 17% of mayors.¹ The slow advancement of women to political leadership roles has been attributed to various factors, including societal gender norms (Folke and Rickne, 2020), bias within political parties (Esteve-Volart and Bagues, 2012), voter prejudice (Le Barbanchon and Sauvagnat, 2022), and gender differences in political persistence (Wasserman, 2023).

In this paper, we examine whether the gap in representation between entry-level and leadership positions might reflect electoral structures that favor extended incumbent tenures. Specifically, we investigate whether term limits can provide historically underrepresented groups, particularly women, greater access to political power and help align representation with evolving societal characteristics. We also analyze when term limits are most effective and how they can complement other equal representation policies, particularly entry-level quotas.

Term limits may promote female representation when a substantial portion of women have attained lower-level political positions but have been deterred from reaching top political positions due to incumbents’ entrenchment. Another potential channel is through the effect on the duration of political tenures. Female representation might increase or decrease depending on which gender places greater value on extended time in office. Furthermore, the short-term direct impact of term limits might be amplified if earlier exposure to female leaders has subsequent effects on voters, party members, and potential candidates.

Despite term limits’ theoretical potential to foster gender equality through increased turnover, empirical evidence of their impact on women’s political representation remains limited and inconclusive. There is cross-country evidence showing that the share of female MPs is larger in countries without term limits, although a causal interpretation is challenging due to potential underlying differences across countries (Schwindt-Bayer, 2005). Some studies have also found that women’s electoral success increases when incumbents do not seek reelection (Andersen and Thorson, 1984; Darcy, 1994) or fail to be elected (Lippmann, 2023), suggesting that term limits may have a positive impact on female representation.² Conversely, studies examining the introduction of term limits in various U.S. states during the early 1990s found no significant increase in the number of elected female politicians (Carroll and Jenkins, 2001; Carey et al., 2006; Pettey, 2018), which Carroll and Jenkins (2001) attribute primarily to a lack of female candidates participating in primary elections. The heterogeneity in these findings suggests that term

¹Data on female representation among local councilors, mayors, and parliamentarians are drawn from ? and cover 41 European countries. Data on the share of female heads of government correspond to the 28 European Union member countries and come from EIGE’s Gender Statistics Database, provided by the European Institute for Gender Equality, accessed on November 28, 2024, available at <https://eige.europa.eu/>.

²In a developing country context, Labonne, Parsa and Querubin (2021) find that term limits in the Philippines increased female representation, but this increase was driven by ‘dynastic’ women serving as placeholders for incumbents. This pattern likely reflects the capture of democratic institutions by elites (Martinez-Bravo, Mukherjee and Stegmann, 2017). However, dynastic succession is less common in Western democracies, given their distinct institutional frameworks and democratic traditions.

limits could promote female representation, but identifying the conditions for their effectiveness remains an open question.

To examine the conditions under which term limits can increase female representation, we exploit quasi-experimental variation in mayoral term limits across Italian municipalities. In this setting, mayors were initially restricted to two consecutive five-year terms. We analyze two reforms that extended term limits to three terms in smaller municipalities, a change reportedly motivated by difficulties in candidate recruitment. The first reform took place in 2014 and it increased the term limit in municipalities with fewer than 3,000 inhabitants, while larger municipalities remained unaffected. A second reform in 2022 extended this three-term cap to municipalities with populations below 5,000 inhabitants.

This setup presents several advantageous features. First, the population-based design of the reforms facilitates our empirical strategy. In a context where these population thresholds have been also used for other policies, we isolate the impact of term limits using a difference-in-discontinuities approach, where the evolution of slightly larger municipalities, unaffected by the increase in term limits, serve as a control group (Eggers et al., 2018). As we show below, treatment and control groups showed similar pre-reform patterns in the main outcome variables around the thresholds, supporting the validity of this empirical strategy. Second, its rich heterogeneity enables us to investigate the conditions under which term limits increase female representation. Specifically, we examine how the impact of term limits varies with the presence of women in local politics, exploiting the large variation in the share of female politicians across Italian municipalities. Furthermore, the available data enable us to identify the channels through which term limits affect female representation in top positions, distinguishing between the direct impact of incumbent replacement and the indirect effect of changes in the time horizon of mayoral positions. A further advantage of this setting is methodological.

Since results are qualitatively similar at both thresholds, we focus on the weighted average of our difference-in-discontinuities estimates across the two cutoffs. In the first electoral cycle after the reform, the extension of term limits decreased the probability that the incumbent was term limited by 45 percentage points (standard error=5 p.p.) and increased the probability of mayoral re-election by 19 p.p. (s.e.=7 p.p.). While the main objective of the policy was to increase the quantity and quality of mayoral candidates, instead we observe a significant decrease in the number of mayoral candidates ($\beta=-0.18$, s.e.=0.09), perhaps reflecting the decrease in political competition, and there is no significant impact on candidates' educational attainment, a variable that has been used in the literature to capture politicians' quality (e.g. Gagliarducci and Nannicini (2013); De Benedetto and De Paola (2019)).

Furthermore, extending term limits produced an unintended consequence: slower progress in women's access to mayoral positions. The share of female mayoral candidates decreased by 8 p.p. (s.e.=2 p.p.) in municipalities where term limits were extended, compared to slightly larger municipalities unaffected by the reform. This change in the composition of candidates translates into a proportional decrease in the share of female mayors, which falls by 8 p.p. (s.e.=3 p.p.). The size of this effect is approximately twice as large at the 5,000 threshold compared to the 3,000 one. This reduction in female representation is substantial, given that only about 18% of mayors in the control group are women. However, we find no evidence that

these changes stem from voter behavior, as neither electoral turnout nor the share of women voters is affected. Estimates using a difference-in-differences approach yield very similar results with greater precision.

Our heterogeneity analysis reveals that the effectiveness of term limits depends on the level of female representation in lower-level political positions, underscoring the importance of the availability of potential female mayoral candidates. This relationship helps explain the stronger impact of term limits on female representation observed in 2022 compared to 2014. Municipalities affected by the 2022 reform had been subject to gender quotas in their councilor candidate lists since 2013, which had created a broader pool of potential female mayoral candidates. In a context where entry-level quotas alone had failed to create short-term trickle-up effects to mayoral positions (Spaziani, 2022), this evidence suggests that term limits can complement gender quotas at lower levels by creating pathways for emerging female politicians to advance to higher positions.

Our analysis also indicates that term limit extensions slow women’s advancement to top positions primarily by delaying access for new cohorts of politicians, rather than through gender differences in preferences for longer tenure. The effect on female mayors appears only in municipalities where incumbents were in their second term, indicating that reduced female representation stems from incumbent retention rather than from gender-based differences in preferences for extended time in office.

An important challenge to our empirical strategy is the existence of several policies that vary at the same population thresholds as the term limit reforms. Except for the introduction of gender quotas in the executive, these policies were implemented years before the extension of term limits, and our difference-in-discontinuities design aims to disentangle the effect of term limits from these potential confounding factors. Most notably, historically mayors’ salaries increase discontinuously at the 3,000 and 5,000 population thresholds. Prior research by Gagliarducci and Nannicini (2013) shows that mayors in municipalities above the 5,000 threshold were more educated, suggesting that higher compensation attracts candidates with stronger outside options. However, they found no effect on gender composition—a finding our analysis confirms. The absence of gender differences around this threshold before 2022 indicates that salary differentials cannot explain the changes in female representation observed after the term limit reform.

The 5,000 population threshold also affects female representation in the council. In 2013, nine years before this same threshold was used to implement extended term limits, municipalities above 5,000 inhabitants became subject to gender quotas in council candidate lists and double preference voting. This reform increased the share of women among elected councilors, but it had no immediate effect on the proportion of female mayors. (Baltrunaite et al., 2019; Spaziani, 2022). Although a trickle-up effect to top positions might emerge over time, the timing of changes around the 5,000 threshold suggests that they are primarily driven by the extension of term limits. Specifically, consistent with previous studies, we observe no significant gender differences in mayors around the cutoff either in the first election cycle after quotas (2013-2017) or during the first four years of the second cycle (2018-2021). The discontinuity in female mayors appears in 2022, within the second election cycle after gender quotas had been

introduced, coinciding with the term limits extension.

Another challenge to our empirical strategy comes from the 2014 introduction of gender quotas for executive committees in municipalities with more than 3,000 inhabitants, which coincided with the term limits extension and relied on the same population threshold. Several findings suggest that executive gender quotas are unlikely to have affected the share of female mayors. First, the gap in female mayors around the 3,000 threshold emerges strongly in the first post-reform election cycle (2014-2018) but it becomes smaller and not significant in the following cycle (2019-2023). This pattern aligns with term limits' effects, as all incumbents could run again in the first election while some reached their three-term limit in subsequent elections. A trickle-up effect from executive quotas should strengthen over time, yet we observe the opposite pattern. Second, we find stronger effects in municipalities where female politicians were already well-represented before the quota, making the gender quota less binding. Third, the increase in the share of female mayors appears only in municipalities where incumbents were serving their second term and thus affected by the term limit extension.

This paper contributes to several strands of the literature. First, it extends research on term limits, which has traditionally focused on balancing two competing effects: preventing incumbent entrenchment versus the costs of restricting voter choice and removing re-election incentives (e.g. Besley and Case (1995); Alt, Bueno de Mesquita and Rose (2011)). Our work highlights an underexplored benefit of term limits: facilitating representational congruence between elected officials and the population over time. While previous studies on this question have yielded mixed results, our analysis shows that term limits effectively increase women's representation when there is already substantial female representation at lower levels of the political hierarchy. Without this condition, as seen in some U.S. states in the 1990s, term limits fail to increase women's representation in top positions (Carroll and Jenkins, 2001; Caress, 1999; Pettey, 2018)). Our research also provides novel evidence on the mechanisms through which term limits affect female representation. We show that the effect operates mainly through incumbent replacement rather than gender differences in preferences for political career length. We also find little evidence of an exposure effect, where women in top positions influence party members and voters. Our findings show that countries seeking to increase female representation should carefully consider the significant negative consequences of relaxing term limits, as revealed by the Italian experience.

Second, we contribute to the literature on gender representation policies and the persistence of inequality in politics. Prior work has largely focused on gender quotas, which have been policymakers' primary tool for fostering gender equality. In Europe, where quotas typically target lower levels of the political hierarchy, they have substantially increased women's representation at targeted levels, but their impact on higher-ranked positions remains limited.³ Our results indicate that term limits can complement quotas by creating pathways for women to advance from lower-level positions to mayoral roles.

³Mandatory gender quotas in candidate lists have not impacted significantly the share of female mayors within two electoral cycles in Spain (Bagues and Campa, 2021), Italy (Baltrunaite et al., 2019; Spaziani, 2022) or France (Lassébie, 2020). In contrast, O'Brien and Rickne (2016) find that the voluntary introduction of gender quotas in candidate lists by the Swedish Social Democratic Party did increase the share of female mayors. In the context of India, O'Connell (2020) show that mandated female representation at the local level increases the number of female candidates who later contest seats in state and national legislatures.

Finally, our work contributes to the evaluation of the extension of term limits in Italy. Our work confirm previous findings by De Benedetto and De Paola (2019), who showed that the 2014 reform did not improve the quality of mayoral candidates, as measured by their educational attainment. However, in contrast with De Benedetto and De Paola (2019), our results using a larger sample suggest that the extension of term limits did not affect electoral participation and it increased the share of female candidates and mayors.⁴

The paper proceeds as follows. Section 2 discusses the different channels through which term limits can affect female representation in top positions. Section 3 describes the institutional framework, including the competencies of Italian local governments and the electoral system, as well as other policies that rely on the 3,000 and 5,000 population thresholds. Section 4 presents our data, and Section 5 provides the empirical analysis, using a difference-in-discontinuities approach. Section 6 concludes by summarizing our main findings, discussing their policy implications, and outlining directions for future research.

2 Conceptual framework

Term limits can create opportunities for new candidates to replace the existing political class. There are at least three mechanisms through which term limits may affect women’s representation: (i) the likelihood that a woman replaces a term-limited incumbent as the party’s new candidate (‘replacement’ effect), (ii) how shorter time horizons differently influence the attractiveness of political careers for men and women (‘selection’ effect), and (iii) how exposure to female politicians shapes the attitudes of voters, party members, and potential candidates over time (‘exposure’ effect).

The strength of the replacement effect is likely larger in contexts with a strong incumbency advantage and where term limits are more binding. Furthermore, if the incumbent is term-limited, the probability that the new party candidate is a woman will be higher if there is a large pool of female politicians at lower levels, which may be fostered by the existence of gender quotas at lower levels of the hierarchy. It is also crucial that party leaders and members are not biased against female candidates. However, women may still fail to reach top positions if voters have prejudices against female leaders.

In addition to this direct effect through the replacement of incumbents, a second channel - selection - may affect female candidates if the possibility of a longer tenure makes politics relatively more attractive to women or men, for example through greater stability in the position. If term limits affect female representation through this channel, then we would expect a change in the female representation also in municipalities where the incumbent is not yet term-limited.

⁴De Benedetto and De Paola (2019) analyze how the extension of term limits in small Italian municipalities in 2014 affected electoral participation using a difference-in-discontinuities approach and data for years 2011-2017, three and four years before the reform, with a sample of around 600 municipalities within the optimal bandwidth. They find that the extension of term limits decreased turnout by 4.6 p.p. (s.e.=2.1 p.p.) and they point out that, in unreported results, “extended term limits reduce the probability of having female candidates running for election (significant at 10 percent level),” though they caution that this result is not robust to the choice of bandwidth. Instead, in our analysis of the impact of the 2014 reform we use information for all elections in the previous and following electoral cycles (years 2009-2018), with a sample of around 3,000 municipalities within the optimal bandwidth. We find a relatively precise zero estimate for electoral turnout (-0.1 p.p, s.e.=0.4 p.p.) and a significant impact on the share of female candidates (-7 p.p, s.e.=3 p.p.) that is robust to different bandwidths. Our analysis of the 2022 reform yields similar results.

Finally, the long-term effects of increased political turnover depend on whether it shifts the equilibrium level of representation through an exposure effect, or it just accelerates convergence to a pre-existing equilibrium level that reflects underlying voter preferences and the supply of minority candidates. Term limits may directly shift this equilibrium upwards, through a feedback cycle between elected representatives and the supply of candidates Baskaran and Hessami (2018), or via role-model effects Beaman et al. (2009).

Analyzing the impact of term limits through the lenses of replacement, selection, and exposure can help to elucidate how their effects may vary across different contexts and time periods. A substantial replacement effect can be expected during periods of rapid growth in women's political presence and in contexts with strong incumbency advantages. This replacement effect would be observable in races where the incumbent is term-limited. Selection effects are more likely to emerge when men and women have different preferences or face differing labor market opportunities. In this case term limits might affect the gender mix of mayoral candidates in all races, irrespective of the incumbent's term-limited status. Finally, if an exposure effect is present, we would expect to observe an additional increase in the share of female politicians in subsequent elections in municipalities that experienced greater mayoral turnover and, by proxy, more exposure to female mayors in the intervening period. This exposure effect should be more pronounced in contexts where lack of prior exposure to female mayors has been a barrier to representation. We will revisit these concepts in our analysis.

3 Institutional framework

3.1 Local government and electoral system

Italian municipalities are responsible for a wide range of public services, including water supply, waste management, municipal policing, infrastructure maintenance, housing provision, and welfare policy implementation. The local administration is formed by a mayor, an executive committee, and a council. The mayor serves as the executive head, guiding allocation of public funds and appointing the executive committee. Members of the executive are assigned to areas of competence, and each of them supersedes a specific policy domain (e.g. education), assisting decision-making in their area. Council members can propose policies and local legislation and vote on them. While mayors and council members are elected directly by voters, executive committees are chosen among elected council members by the mayor.

The electoral system varies depending on municipalities' population. In municipalities with less than 15,000 inhabitants, which are the focus of this paper, the mayor is elected through a majoritarian system and the list associated with the mayor receives 2/3 of the seats in the council through a proportional representation system with open lists. During the period of our analysis, local elections took place every five years but, for historical reasons, such as the dissolution of the local government, not all municipalities hold elections on the same year.

3.2 Term limit reforms: 2014 and 2022

While term limits are relatively uncommon in Europe, Italian mayors have been subject to a two-term limit since 1993.⁵ In April 2014 term limits were extended from a maximum of two to a maximum of three consecutive five-year terms in municipalities with less than 3,000 inhabitants, as measured in the latest census.⁶ The main political motivation for the term limit extension was the difficulty in finding mayoral candidates who could replace incumbents in small municipalities.⁷ This reform was not implemented in five regions with “special status”, which have autonomy to decide on electoral policies.⁸ In 2022, term limits were extended to three terms in all municipalities with less than 5,000 inhabitants.

3.3 Other policies based on population cutoffs

The population thresholds used for the extension of term limits are also relevant for several institutional features that might potentially affect mayor characteristics.

3.3.1 Salaries and size of municipal council

Mayoral salaries are determined based on the number of inhabitants of the municipality, following a step function that varies at different population cutoffs, including the 3,000 and the 5,000 cutoffs. Over time, differences in compensation according to municipality size have been stable, except for a 15% salary increase in 2020 in municipalities below 3,000 inhabitants.⁹ The existing empirical evidence indicates that, before term limits were extended, attaining a mayoral position had only a modest positive impact on earnings (Bertoni et al., 2023).¹⁰

The size of the municipal council also varies at some population thresholds. While it does not change at the 5,000 threshold, the municipal council has been always slightly larger in municipalities with more than 3,000 inhabitants. The specific difference has changed several times during the period of our study.¹¹

⁵Term limits in Europe mainly apply to heads of State under executive presidencies (Baturu and Elgie, 2019). At the sub-national level, only Italy and some Spanish regions apply term limits (European Commission for Democracy Through Law - Report on Term Limits Part 2, 2019, available at https://www.te.gob.mx/vota_elections/media/files/4a7c8c5e9bc8cc8.pdf). Term limits are more prevalent in other regions of the world. In the United States, the president and governors of several states are subject to term limits, but such restrictions do not apply to members of the Senate or House of Representatives. In Africa, term limits were widely introduced during the 1990s at the presidential level (Zamfir, 2016).

⁶Law 7 April 2014, n. 56, available at <https://www.gazzettaufficiale.it/eli/id/2014/4/7/14G00069/sg>.

⁷See discussion in law proposal n. 3539, available at https://www.camera.it/leg17/995?sezione=documenti&tipoDoc=lavori_testo_pdl&idLegislatura=17&codice=17PDL0038480

⁸Three of these regions remained with a two-term restriction for all municipalities (Sardinia, Sicily, Friuli Venezia-Giulia), while two regions (Aosta Valley and Trentino-South Tyrol) already had a three-term limit.

⁹The law decree 2019-10-26, article 57, increased mayors' remuneration in municipalities below 3,000 to 85% of the salary in municipalities just above this threshold. The salary increase was effective on 25-12-2019. See <https://www.normattiva.it/uri-res/N2Ls?urn:nir:stato:decreto.legge:2019-10-26;124>

¹⁰Employing a close elections research design, Bertoni et al. (2023) find that successful mayoral candidates earn approximately 6% more annually over a 15-year horizon compared to unsuccessful candidates. The availability of one additional term is unlikely to significantly increase this premium.

¹¹The municipal council included 12 members in municipalities below 3,000 before 2011, 9 members in 2011, 6 in 2012 and 2013, and 10 after 2014, while in municipalities with a population size between 3,000 and 10,000, councils were composed by 16 members before 2011 municipal, 12 members in 2011, 7 in 2012 and 2013, and 12 since 2014.

3.3.2 Equal representation policies

Two different gender quotas were implemented in Italy in recent years to increase female representation at different levels of the political hierarchy. In 2013, municipalities with more than 5,000 inhabitants adopted a 40% gender quota on candidate lists along with double preference voting, allowing voters to cast up to two preference votes - one for a female and one for a male candidate for the council. While this reform substantially increased the share of female council members, it did not significantly affect the share of female mayors or executive members (Baltrunaite et al., 2019; Spaziani, 2022).

One year later, in 2014, municipalities with more than 3,000 inhabitants implemented a 50% gender quota on executive committees, affecting the mayor's choice of committee members. Notably, this population threshold of 3,000 inhabitants matches the cutoff used for term limit extensions, though term limits were extended in municipalities below this threshold while gender quotas were implemented in those above it.

3.3.3 Other policies

There are two other policies relying on population thresholds during this period. The Domestic Stability Pact (DSP), a regulation aimed at improving national financial stability, was introduced in 1999 in all municipalities, constraining their capability to run deficits. In 2001, municipalities below 5,000 were excluded and, in 2013, the rule was extended to all municipalities between 1,000 and 5,000 inhabitants. Finally, in 2015 the DSP was replaced with a budget balance rule for all local authorities. Given that the term limits reform relying on the 5,000 threshold was introduced much later, in 2022, and it was implemented taking into account the population according to the 2021 census, the Domestic Stability Pact is unlikely to be a relevant confounding effect.

Similarly, in 2010 a law was introduced requiring small municipalities to manage jointly a number of basic public services (Law no. 78/2010).¹² This law affected municipalities with fewer than 5,000 inhabitants, except in the case of municipalities legally considered as 'mountain municipalities', where the threshold was lowered to 3,000 residents. Cremaschi et al. (2023) have shown that this reform reduced access to local public services and increased the electoral support for far-right parties in national elections by around 1.5 percentage points. Given its timing, we do not expect this policy to affect our analysis of the term limits at the 5,000 threshold. However, this policy might have potentially affected some electoral outcomes in 2014 in 'mountain municipalities' below the 3,000 threshold, coinciding with the extension of term limits.

¹²Municipalities were required to manage jointly manage at least three public services by January 1, 2013, at least three other services by September 30, 2014, and all remaining services by December 31, 2014.

4 Data

We use data from local elections held between 1992 and 2023.¹³ We exclude from the analysis Italian regions with a special status, as they were not subject to the policy reform. Given that our difference-in-discontinuities analysis typically selects an optimal bandwidth of around 2,000 inhabitants, next we describe the main characteristics for municipalities with populations between 1,000 and 7,000 inhabitants, comprising 4,082 municipalities. Around 24% of the Italian population lives in these municipalities. Our focus on small municipalities implies that we are estimating the impact of the policy in areas that are relatively more rural, with slightly fewer women, and older population. Close to 40% of these municipalities have the status of mountain municipality.

4.1 Electoral information

As shown in Table 1, turnout tends to be high, around 75% of eligible residents vote, and the gender composition of voters is balanced, with 50% of voters of each gender. On average, there are around 2.4 candidates to mayor per municipality.

4.2 Mayor characteristics

On average, mayors in our sample are 49 years old with 15 years of education. Dynastic candidates, which we proxy using surname information, seem to be rare. Around 1% of mayors have the same surname as the previous incumbent.¹⁴ Mayors tend to have substantial political experience. They average 10 years of prior service as council members and 0.62 terms as mayor. Although 31% of incumbents are term-limited, among those eligible to run again, more than half win re-election, with re-elected mayors accounting for 39% of all mayors. The high share of re-elected mayors stems from two factors. First, most eligible incumbents, regardless of gender, pursue re-election. Second, incumbents tend to enjoy an electoral advantage, winning approximately two-thirds of their re-election bids. Thus, given this incumbency advantage, term limits create crucial opportunities for new candidates to access mayoral positions.

4.3 Female representation

While female representation in mayoral positions has risen from just over 4% in 1993 to 16% in 2023, women remain substantially underrepresented (see Figure 1a). Representation is higher at lower levels of government, in municipal councils and executive committees. This representation has grown substantially, particularly following two key reforms: the 2013 introduction of gender quotas in candidate lists and double preference voting for municipalities above 5,000 inhabitants, and the 2014 implementation of gender quotas in executive committees for municipalities above 3,000 inhabitants.

¹³Electoral data was retrieved from the Ministry of Internal Affairs *Eligendo* <https://elezioni.interno.gov.it>, see Data Appendix C for details

¹⁴The scarce presence of dynastic candidates is in contrast with findings for the Philippines by Labonne, Parsa and Querubin (2021), reflecting the difference in institutional context between the two countries, and suggests that the role of kinship in shaping the impact of term limits is less important in this context.

Women are more prevalent among younger cohorts of politicians, with higher representation among newly elected mayors and mayoral candidates compared to incumbents (see Figure 1b). This generational pattern suggests that female representation in mayoral positions is likely to increase through natural turnover over time.

5 Empirical analysis

The 3,000 and 5,000 population thresholds have historically been used for multiple policies, potentially invalidating a standard regression discontinuity design in this setting. To address this challenge, in our main analysis we implement a difference-in-discontinuities approach (RD-DID). By examining outcome variables in differences using a regression discontinuity design, this method aims to account for time-invariant factors that vary at the threshold (e.g. Eggers et al. (2018), Grembi, Nannicini and Troiano (2016)). The RD-DID relies on stronger assumptions than the standard regression discontinuity design. A potential threat to the validity of RD-DID is that, if municipalities above and below the threshold are somehow different, they might be affected differently by time-variant shocks. For instance, the economic cycle might potentially affect mayors' characteristics differently depending on the level of salaries. Although we cannot fully eliminate these concerns, several placebo exercises support the validity of our RD-DID strategy, as we show below.

We report estimates using a difference-in-differences (DID) approach in Appendix D. These estimates are similar in magnitude but more precise than those from the RD-DID. Event study analyses and placebo tests support the validity of the DID approach. However, we present RD-DID estimates in the main text and DID in an appendix because the former method provides more credible control over the running variable through its regression discontinuity framework, albeit with reduced statistical power.

We focus our analysis of the effects of the term limits extension on three classes of outcomes: (i) impact on incumbents, (ii) electoral competition and voter behavior, and (iii) characteristics of elected politician. To understand the extent to which the policy affected incumbents, we assess the impact of the extension on the probability an incumbent was term-limited and re-elected as mayor. Second, we turn to the policy's effects on electoral competition and voting behavior, by estimating effects on number of candidates and turnout, focusing in particular on the share of female candidates and female share of voters. Finally, we assess to what extent these potential changes generated by the policy have consequences for the characteristics of elected politicians by analyzing its impact on the share of female mayors, the age and experience profile of politicians (with experience proxied by previous tenure in council) and years of education - a commonly used proxy for the quality of elected officers.

5.1 Validity of the difference-in-discontinuities approach

To assess the credibility of the RD-DID empirical strategy, we (i) examine the distribution of municipalities around the thresholds to test for manipulation in population counts, and (ii) conduct placebo estimations for the outcome variables in previous electoral cycles.

5.1.1 Density function

Term limit reforms were implemented based on population counts from the census valid at the time of each election. Elections between 2013 and 2022 used the 2011 census, and elections from 2023 onwards were determined by the 2021 census. Given that census counts preceded the term limit reform and this was difficult to anticipate, the primary threat of manipulation stems not from this reform itself, but from prior policies that used these thresholds, such as higher mayoral salaries.

Despite several policies varying at the 3,000 and 5,000 inhabitants thresholds, we find no evidence of manipulation in the 2011 census at either threshold, which determined term limits for the 2014 and 2022 reforms (see Figure 2). Similarly, we find no manipulation at the 5,000 threshold in the 2021 census, which governed term limits from 2023 onwards. The absence of bunching in population counts supports our identification strategy.¹⁵

5.1.2 Pre-existing discontinuities at the threshold

The presence of other policies at the 3,000 and 5,000 thresholds may have influenced the characteristics of municipalities around these cutoffs. We first examine whether our outcome variables exhibit discontinuities prior to the reform. Given that we observe some pre-existing differences, we then investigate whether examining changes in outcomes rather than levels can account for these differences. As we show below, placebo tests using outcome variables in differences support the validity of our RD-DID approach.

Pre-existing differences in outcome variables in levels We estimate the following equation:

$$(1) \quad Y_{it} = f(\text{population}_{it}) + \beta \text{AboveThreshold}_{it} + \varepsilon_{it}$$

where Y_{it} denotes the outcome variable in municipality i and election t , $f(\text{population}_{it})$ is a flexible function of the running variable population , as measured in the census, and AboveThreshold is an indicator that takes value one for municipalities above the corresponding population threshold. We use local polynomial estimation and robust bias-corrected inference procedures, employing the mean square error optimal bandwidth proposed by Calonico, Cattaneo and Titiunik (2014). Our main specification uses a linear polynomial and a uniform kernel, which assigns equal weights to all observations within the bandwidth.¹⁶

We examine twelve outcome variables: (i) whether the incumbent is term limited, (ii) number of candidates, (iii) share of female candidates, (iv) turnout, (v) share of female voters, (vi) whether the incumbent was re-elected, (vii) mayor's age, (viii) mayor's years of education, (ix) mayor's previous experience in the council, (x) mayor's gender, (xi) share of female council members and (xii) share of female executive members. Due to data availability, for the 3,000

¹⁵In unreported results, we detect a discontinuity at the 3,000 threshold in the 2001 census, but this discontinuity had disappeared by the time term limits were extended below this threshold. A discontinuity emerges again at the 3,000 threshold in the 2021 census, after this threshold had become irrelevant for term limits.

¹⁶The uniform kernel specification allows us to display in RD plots the same specification reported in RD tables.

threshold, we test for pre-existing discontinuities using the three electoral cycles prior to the 2014 term limit extension: 1999-2003, 2004-2008, 2009-2013. For the 5,000 threshold, we examine the five previous electoral cycles: 1997-2001, 2002-2006, 2007-2011, 2012-2016, 2017-2021.

At the 3,000 threshold, we conduct 36 placebo tests, examining twelve variables over three electoral cycles. As shown in Figure A.1, five of these tests yield significant discontinuities: the share of term-limited incumbents and the number of candidates in 2004-2008, the share of female candidates and female mayors in 1999-2003, and mayors' age in 2009-2013. At the 5,000 threshold, our examination of twelve variables over five cycles yields 60 different placebo tests. We observe a large discontinuity in the share of women in the council during 2012-2016 and 2017-2021 (see Figure A.2). This discontinuity reflects the 2013 introduction of gender quotas in municipalities above 5,000 inhabitants, which first affected elections during 2013-2017. For the remaining eleven variables, we find no systematic pattern of discontinuities prior to the reform. Only three out of 55 placebo tests show significant differences at the 5% level: mayoral age in 2003-2007, experience in 2017-2021, and women in the executive in 2012-2016.

Pre-existing differences in outcome variables in differences The presence of discontinuities in lagged outcome variables, particularly at the 3,000 threshold, invalidates a standard regression discontinuity design. To examine whether using a RD-DID can help to address this issue, we estimate:

$$(2) \quad \Delta Y_{it} = f(\text{population}_{it}) + \beta \text{AboveThreshold}_{it} + \varepsilon_{it}$$

This specification mirrors equation (1) but uses the dependent variable in differences, where ΔY_{it} measures changes in Y_{it} from the previous election.

As shown in Figures 3 and 4, when we examine changes in outcomes rather than levels, most discontinuities disappear, supporting the validity of the RD-DID strategy. The main exception is a significant increase at the 5,000 threshold in the share of women in the council during 2012-2016 and 2017-2021, reflecting the 2013 introduction of gender quotas in council candidate lists. At the 3,000 threshold, only two out of 36 placebo tests show significant estimates (turnout and mayoral years of education in 1999-2003), consistent with what would be expected by chance. Similarly, at the 5,000 threshold, excluding the share of women in the council, only one out of 52 placebo tests shows a significant difference, the share of women in the executive in 2012-2016.

5.2 Short-term impact of term limit extensions

We analyze the impact of term limit extension in the first election following the reform. We report our estimates in Table 2, and the corresponding RD plots in Figure 5 (3,000 threshold) and Figure 6 (5,000 threshold), where we also include information on the optimal bandwidth. Since estimates are qualitatively similar at both thresholds, we focus our discussion on their weighted average.

We begin by examining incumbents' eligibility to run again to understand how many incumbents were retained in the pool of potential candidates to mayoral positions. As shown in the upper panel of Table 2, column 1, in municipalities not affected by the extension of term limits,

the probability that the incumbent is term-limited and unable to run for re-election is 45 p.p. higher (s.e.=5 p.p.). This estimate mechanically reflects the share of second-term mayors who can run again in smaller municipalities due to the term limit extension but must step down in larger ones. Since incumbents typically enjoy an electoral advantage, their inability to run may create opportunities for new candidates, altering the candidate pool. Indeed, municipalities with extended term limits experience both a decrease in the number of candidates ($\beta=0.18$, s.e.=0.09) and a lower share of female candidates ($\beta=8$ p.p., s.e.=2 p.p.), relative to larger municipalities not affected by the reform.

While the presence of more incumbents in the candidate pool could affect voting behavior by reducing uncertainty in electoral outcomes, and a larger presence of female candidates could stimulate female voter participation, we find no evidence of such changes in voter behavior. Neither turnout (95% confidence interval=[-0.01, +0.01]) nor the share of female voters (95% confidence interval=[-0.00, +0.00]) show significant changes.

The reform has a significant impact on the selection of mayors. Term extensions increase the probability of incumbent re-election by 19 percentage points (s.e.=7 p.p.). These extensions result in the retention of more experienced mayors: affected incumbents have on average 1.5 more years of council experience (s.e.=0.7) and are half a year older, though this age difference is not statistically significant at conventional levels. These results indicate that term limits extensions delayed the access to mayoral positions to new cohorts of politicians.

Additionally, municipalities affected by the reform have significantly lower shares of female mayors ($\beta=8$ p.p., s.e.=3 p.p.). This effect is notably larger at the 5,000 threshold than at the 3,000 threshold (16 p.p. vs 6 p.p.), a difference we discuss in Section 5.3 when examining how women’s presence in lower political positions affects term limits’ impact.

Turning to other characteristics of elected mayors, we find no impact on mayors’ education levels, consistent with previous findings by De Benedetto and De Paola (2019). In terms of female representation at lower levels, we do not observe any changes in terms of the share of female council members but, consistently with the introduction of gender quotas in the executive in 2013, there is an increase in the share of female executive members at the 3,000 threshold.

5.2.1 Robustness

Results on the impact of the reform on our main outcomes of interest - female mayor and female candidates - are robust to alternative specifications, including different choices of bandwidth, polynomial order, and kernel. As shown in Figure A.3, our estimates of term limits’ impact on the share of female candidates and mayors remain stable when using bandwidths that are half or twice the size of the optimal bandwidth, suggesting results are not driven by a particular sample of municipalities included in the optimal bandwidth. Results are also unchanged when we consider a second order polynomial or a triangular kernel (see Table B.1).

Finally, our results are also qualitatively similar but more precise when using a difference-in-differences strategy, where we consider municipalities within 2,000 inhabitants of the corresponding threshold as treatment and control groups (see Appendix D). As shown in Table D.2), the difference-in-differences estimates indicate that the 2014 term limit extension decreased the share of female mayoral candidates by 2 p.p. (s.e.=1 p.p.) and female mayors by 4 p.p. (s.e.=1

p.p.), while the 2022 reform reduced the share of female candidates by 7 p.p. (s.e.=3 p.p.) and female mayors by 11 p.p. (s.e.=4 p.p.).

5.3 Mechanisms

As discussed in section 2, term limits can affect female representation through several channels. Beyond the direct effect of increased turnover, the prospect of longer tenure could have altered the selection of politicians into the candidate pool. To disentangle the roles of replacement and selection effects, we examine whether the impact varies with incumbents' tenure, focusing on whether they had completed their first or second term and were thus directly affected by the term limit extension. We also exploit heterogeneity in representation at lower political levels to understand under which conditions term limits can foster female representation.

Our results reveal two key patterns. First, the reform's impact on female representation is significant only when the incumbent was completing their second term, consistent with a replacement effect. Second, the effect is substantially larger in municipalities that previously had more women in their councils, suggesting that a pool of potential female mayoral candidates is crucial for term limits to effectively foster female representation.

5.3.1 Selection versus replacement effects? Impact by number of terms served by the previous incumbent

This heterogeneity analysis helps distinguish between two potential mechanisms through which term limit extensions might decrease female representation. The first is a direct 'replacement effect,' whereby term-limited incumbents are allowed to run again. The second is a 'selection effect,' which could arise if men and women differ in their preferences for longer political careers.

We focus on the first cycle of elections after the reform. The comparison between municipalities with first- and second-term mayors provides a way to separate these effects. In municipalities with second-term mayors, the reform has two impacts: it allows incumbents to run again and it extends the time horizon for future mayors. In contrast, in municipalities with first-term mayors, incumbents can seek re-election regardless of the reform, so the main effect is likely to operate through changes in the expected length of political careers.

As shown in Table 3, in municipalities where incumbents were in their first term, term limit extensions increase the share of female candidates but they have no effect on mayoral characteristics. However, in municipalities where incumbents were in their second term, the reform leads to a 33 p.p. increase (s.e.=5 p.p.) in incumbent re-election probability, a 4.4-year increase in mayors' council experience (s.e.=1.5), and a 13 p.p. decrease in the share of female mayors (s.e.=6 p.p.).

This evidence suggests that the 'replacement effect' is the primary mechanism driving the decrease in female representation. Term limit extensions allow more experienced incumbents to remain in office, while municipalities where incumbents must step down see their replacement by less experienced candidates, who are more likely to be women.

5.3.2 The role of supply: Impact by previous level of female representation at lower hierarchy levels

Term limits should theoretically increase female representation in top positions when there is a sufficient pool of potential female candidates. We test this prediction by examining the 2014 reform’s impact across municipalities with different levels of female representation, using the share of women in the two previous councils as a proxy for women’s presence in local politics.¹⁷

We estimate equation (2) separately for municipalities above and below the mean share of female council members in the previous two terms, which was 20%. As shown in Table 4, in both samples the impact of the reform on the probability that the incumbent was term limit and the probability that was re-elected is very similar. However, term limits have a substantially stronger effect on women’s representation in municipalities where women had greater presence in prior councils. When the representation of women in prior councils was above 20%, term limit extensions decrease the share of female candidates by 8 p.p. (s.e.=3 p.p.), compared to only 1 p.p. (s.e.=2 p.p.) in municipalities below the median. Similarly, in municipalities with more women in councils the share of female mayor decreases by 9 p.p. (s.e.=4 p.p.), relative to only 4 p.p. (s.e.=3 p.p.) in municipalities below the median.

While these patterns suggest the importance of having a pipeline of potential female candidates for mayoral positions, the gender composition of councils is not exogenous. The existing composition of councils may reflect municipal characteristics that correlate with female representation. For example, municipalities with more female councilors may have voters or politicians who are generally more supportive of female candidates.

To further isolate the impact of women’s council presence from other factors, we compare two sets of estimates: first, the effect of term limits at the 3,000 threshold in 2014-2018, when female councilors were relatively rare (around 20%); and second, at the 5,000 threshold in 2022-2023, where gender quotas increased women’s council representation to about 40% through an exogenous reform at this threshold. As we saw in Table 2, the negative impact of term limit extensions on female representation is twice as large at the 5,000 threshold, suggesting that council composition has a causal effect, and that there may be a complementarity between quota policies at lower levels of the hierarchy, and faster turnover in higher political positions.

5.4 Is there an exposure effect? Short- vs. medium-term effects

The impact of changes in term limits may be amplified or reduced over time. On the one hand, as mayors in small municipalities hit the new three-term limit, the share of term-limited mayors is likely to increase. On the other hand, the increased exposure to female politicians in the first cycle of elections could have a knock-on effect on voters and parties, leading to further improvements in the next elections. This exposure effect would be particularly important in cases where voters or party members have concerns about the quality of female politicians, and learn about their quality over time through exposure to newly elected women.

¹⁷We cannot conduct this analysis for the 2022 reform because the 2013 introduction of gender quotas at the 5,000 threshold created a discontinuity in women’s council representation. This prevents us from constructing a valid counterfactual for municipalities with low female representation above the 5,000 threshold.

To examine potential exposure effects, we estimate the mid-term impact of the 2014 reform using equation (2), treating 2009-2013 as the pre-reform period and 2019-2023 as the post-reform period. The differences between municipalities above and below 3,000 inhabitants decreased over time. In the first post-reform election cycle, incumbents in larger municipalities had a 43 percentage point higher probability of being term-limited. This gap narrows to 11 percentage points (s.e. = 3 p.p.) in the second cycle (see Table 5, top panel), reflecting both that many mayors in smaller municipalities have reached their third term and the 2022 change in threshold from 3,000 to 5,000. The probability of mayoral re-election, initially higher in larger municipalities, becomes 13 percentage points lower (s.e. = 6 p.p.). Similar patterns emerge in female representation. The gap in female candidates' share decreases from 7 percentage points (s.e. = 3 p.p.) in the first cycle to 4 percentage points (s.e. = 3 p.p.) in the second cycle, with a similar pattern in the probability of having a female mayor ($\beta = 3$ p.p., s.e. = 4 p.p.). Overall, we find no evidence of additional exposure effects within the second electoral cycle.

As expected, in the second electoral cycle the share of women in the executive remains higher above the 3,000 threshold ($\beta = 11$ p.p., s.e. = 2 p.p.), reflecting the continued presence of executive gender quotas. The decrease in the gender gap during this period suggests that these quotas had limited influence on female representation in mayoral positions, as we discuss in detail below.

5.5 Confounding effects of other policies

As discussed in section 3.3, there are other policies that rely on the 3,000 and the 5,000 inhabitants cutoffs. Below we provide several pieces of evidence suggesting that the change in female representation at the mayoral level that we observe in 2014 at the 3,000 threshold and in 2022 at the 5,000 threshold was induced by the extension of term limits, and not by these other policies.

5.5.1 Gender quotas in executive committees

In 2014, at the same time as term limits were extended in municipalities below 3,000 inhabitants, gender quotas for executive committees were introduced in municipalities above this threshold. Given that the two policies were implemented at the same time and relied on the same threshold, in principle our RD-DID estimates for the 2014 reform capture the combined effect of these two policies. However, several pieces of evidence suggest that the term limits reform, rather than the gender executive quota, drove the increase in female mayors' share in larger municipalities.

First, the effect is stronger in the initial election cycle after the reform (2014-2018) than in the second cycle (2019-2023). This pattern aligns with term limits, which have their largest impact on incumbent term-limitation in the first election. Executive committee quotas, by contrast, would likely show increasing effects over time. These quotas have limited potential to affect mayoral selection in the first post-reform election, since mayors select executive committee members from council members after elections. However, their impact could grow over time as women who enter politics through quota positions advance up the political ladder.

Second, the effect is concentrated in locations where women's representation in previous terms was higher than the mean (Table 4). This pattern aligns with term limits being the

main driver, as their impact should be stronger when a larger pool of potential female mayoral candidates exists. Executive quotas, by contrast, should have larger effects in municipalities with historically low female representation, where the quota would be more binding. As shown in Table 4, column 7, in municipalities with historically fewer women in council, the executive quota increased women’s share by 19 percentage points (s.e. = 2 p.p.) but produced no significant changes in female representation in higher offices. Conversely, in more gender-balanced municipalities, the quota led to a smaller increase in women’s executive share (12 percentage points, s.e. = 2 p.p.) but generated significant increases in both female candidates and mayors.

Third, we only observe an impact on mayoral characteristics in municipalities where the incumbent mayors were in their second term, which is consistent with the replacement effect induced by term limits. If the executive quota was increasing the probability of a female mayor, we would have expected an effect also when the previous mayor was in the first term.

5.5.2 Gender quotas in candidate lists and double preference voting

Since 2013, municipalities with more than 5,000 inhabitants have been subject to gender quotas in candidate lists and double preference voting. Previous studies show that while this reform substantially increased female presence in councils, it had no immediate impact on the share of female mayors within the first election cycle (Baltrunaite et al., 2019; Spaziani, 2022). However, a potential trickle-up effect over time could influence our 2022-2023 estimates at the 5,000 threshold.

To verify that the higher share of female mayors we observe in municipalities above 5,000 inhabitants after the 2022 term-limits reform is not driven by gender quotas’ long-term effects, we estimate the impact of two-term exposure to gender quotas using an RD-DID strategy and data for years 2018-2021. As shown in Table 6, during years 2018-2021 we observe no variation at the 5,000 threshold in either the share of female candidates or the probability of having a female mayor. However, when analyzing elections taking place in 2022 —when municipalities below 5,000 were also affected by term limit extensions— we find a large and significant increase in the share of female mayors. The fact that this gap appears only in 2022, coinciding with the term-limit reform, suggests that gender quotas in candidate lists to the council cannot explain the increase in female representation at the 5,000 threshold.

5.5.3 Joint provision of basic public services

We also verify that our results are not driven by the 2010 reform requiring ‘mountain municipalities’ with less than 3,000 inhabitants to provide jointly some public services. We re-estimate the impact of the 2014 extension of term limits separately for mountain and non-mountain municipalities (see Table B.2). While estimates are less precise due to the decrease in sample size, they are very similar across the two groups. For instance, the share of female mayors decreased by 5 p.p. (s.e.=4 p.p.) in non-mountain municipalities with less than 3,000 inhabitants, compared to a 6 p.p. (s.e.=5 p.p.) decrease in mountain municipalities.

6 Conclusion

Exploiting evidence from term limit extensions in Italy in 2014 and 2022, we find that without these extensions, female representation among mayors would have been 8 percentage points higher. Longer term limits delayed access to mayoral roles for new cohorts, substantially slowing down growth in female representation. The impact on gender representation was particularly pronounced in municipalities with higher female representation at lower levels of the hierarchy and where gender quotas existed in candidate lists. Our analysis suggests this effect was driven primarily by the persistence of incumbents in mayoral positions, rather than gender differences in politicians' preferences for longer tenure.

Our evidence reveals an overlooked benefit of term limits: they facilitate access to political positions for underrepresented groups and narrow the gap in representation between entry-level and top political positions, particularly during periods of rapid societal change. Moreover, term limits can complement and accelerate the effects of entry-level gender quotas. At a time when policymakers worldwide seek equal representation in higher-level positions, term limits emerge as an effective policy tool.

Our work leaves several open questions. First, the effect of term limits may be asymmetric. While we study the impact of lengthening term limits, reductions in term limits (as implemented in the US in the 1990s) may not have an immediate effect if there is no readily available pool of female candidates. Further research is needed on the impact of introducing or shortening term limits. Second, we only study the short-term effect of the reform within two elections. In the longer term, exposure effects may potentially amplify (or reduce) its impact. Third, as discussed in section 2, the effects of term limits likely depend on the characteristics and preferences of voters, party members, and politicians. Term limits may be less effective in contexts where voters and parties show stronger bias against women. These same factors may also influence the complementarity between term limits and quotas. Fourth, our work focuses on term limits' ability to promote female representation in politics, rather than racial, ethnic, or other forms of diversity. While we expect our results to be relevant for these other dimensions, future work could investigate the effects of term limits on representation of these groups. Finally, our work has implications for labor markets more broadly. Long pipelines or structures that favor the persistence of existing cohorts can impede female representation (Arellano-Bover et al., 2024). Policies that accelerate the careers of younger cohorts in the corporate world, such as term limits for executive boards, may similarly affect female representation and selection into these positions.¹⁸

¹⁸Some authors have advocated for term limits on corporate boards in the US to increase female representation, although to the best of our knowledge, no causal evidence is available (e.g. Rosenblum and Nili (2019)).

References

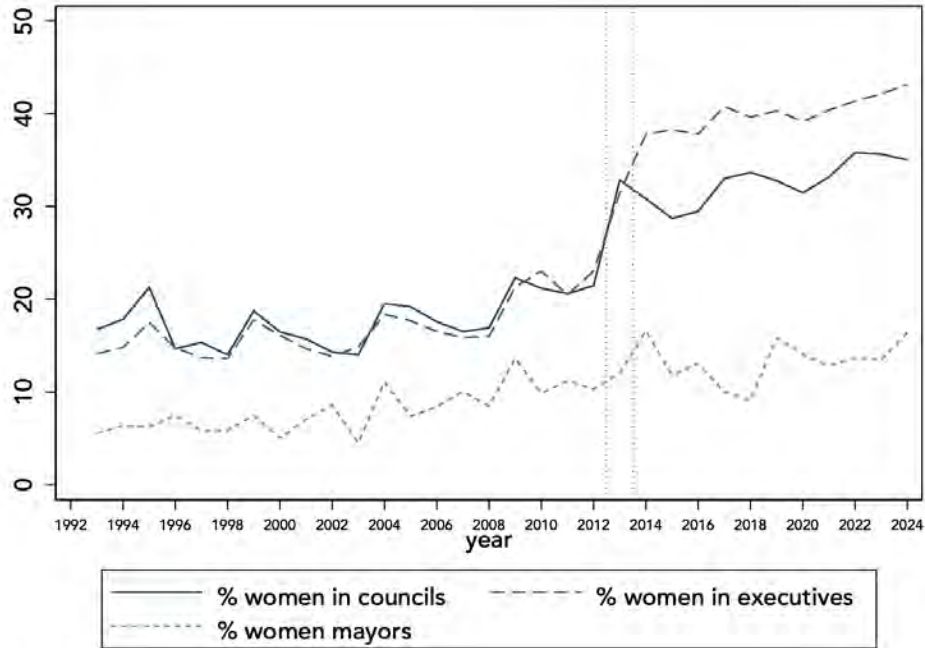
- Alt, James, Ethan Bueno de Mesquita, and Shanna Rose.** 2011. “Disentangling accountability and competence in elections: evidence from US term limits.” *The Journal of Politics*, 73(1): 171–186.
- Andersen, Kristi, and Stuart J Thorson.** 1984. “Congressional turnover and the election of women.” *Western Political Quarterly*, 37(1): 143–156.
- Arellano-Bover, Jaime, Nicola Bianchi, Salvatore Lattanzio, and Matteo Paradisi.** 2024. “One Cohort at a Time: A New Perspective on the Declining Gender Pay Gap.” Available at SSRN.
- Bagues, Manuel, and Pamela Campa.** 2020. “Women and power: Unpopular, unwilling, or held back? A Comment.” *Journal of Political Economy*, 128(5): 2010–2016.
- Bagues, Manuel, and Pamela Campa.** 2021. “Can gender quotas in candidate lists empower women? Evidence from a regression discontinuity design.” *Journal of Public Economics*, 194: 104315.
- Baltrunaite, Audinga, Alessandra Casarico, Paola Profeta, and Giulia Savio.** 2019. “Let the voters choose women.” *Journal of Public Economics*, 180: 104085.
- Baskaran, Thushyanthan, and Zohal Hessami.** 2018. “Does the election of a female leader clear the way for more women in politics?” *American Economic Journal: Economic Policy*, 10(3): 95–121.
- Baturo, Alexander, and Robert Elgie.** 2019. *The Politics of Presidential Term Limits*. Oxford University Press.
- Beaman, Lori, Raghavendra Chattopadhyay, Esther Duflo, Rohini Pande, and Petia Topalova.** 2009. “Powerful women: does exposure reduce bias?” *The Quarterly journal of economics*, 124(4): 1497–1540.
- Bertoni, Marco, Giorgio Brunello, Lorenzo Cappellari, and Maria De Paola.** 2023. “The long-run earnings effects of winning a mayoral election.” *Luxembourg Institute of Socio-Economic Research (LISER) Working Paper Series*, , (2023-02).
- Besley, Timothy, and Anne Case.** 1995. “Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits.” *The Quarterly Journal of Economics*, 110(3): 769–798.
- Calonico, Sebastian, Matias D Cattaneo, and Rocio Titiunik.** 2014. “Robust data-driven inference in the regression-discontinuity design.” *The Stata Journal*, 14(4): 909–946.
- Caress, Stanley M.** 1999. “The influence of term limits on the electoral success of women.” *Women & Politics*, 20(3): 45–63.

- Carey, John M, Richard G Niemi, Lynda W Powell, and Gary F Moncrief.** 2006. “The effects of term limits on state legislatures: a new survey of the 50 states.” *Legislative Studies Quarterly*, 31(1): 105–134.
- Carroll, Susan J, and Krista Jenkins.** 2001. “Do term limits help women get elected?” *Social Science Quarterly*, 82(1): 197–201.
- Cattaneo, Matias D, Michael Jansson, and Xinwei Ma.** 2020. “Simple local polynomial density estimators.” *Journal of the American Statistical Association*, 115(531): 1449–1455.
- Crevaschi, Simone, Paula Rett, Marco Cappelluti, and Catherine E De Vries.** 2023. *Geographies of Discontent: Public Service Deprivation and the Rise of the Far Right in Italy*. Harvard Business School.
- Darcy, Robert.** 1994. *Women, elections, and representation*. Vol. 1, U of Nebraska Press.
- De Benedetto, Marco Alberto, and Maria De Paola.** 2019. “Term limit extension and electoral participation. Evidence from a diff-in-discontinuities design at the local level in Italy.” *European Journal of Political Economy*, 59: 196–211.
- Eggers, Andrew C, Ronny Freier, Veronica Grembi, and Tommaso Nannicini.** 2018. “Regression discontinuity designs based on population thresholds: Pitfalls and solutions.” *American Journal of Political Science*, 62(1): 210–229.
- Esteve-Volart, Berta, and Manuel Bagues.** 2012. “Are women pawns in the political game? Evidence from elections to the Spanish Senate.” *Journal of Public Economics*, 96(3-4): 387–399.
- Folke, Olle, and Johanna Rickne.** 2020. “All the single ladies: Job promotions and the durability of marriage.” *American Economic Journal: Applied Economics*, 12(1): 260–287.
- Gagliarducci, Stefano, and Tommaso Nannicini.** 2013. “Do better paid politicians perform better? Disentangling incentives from selection.” *Journal of the European Economic Association*, 11(2): 369–398.
- Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano.** 2016. “Do fiscal rules matter?” *American Economic Journal: Applied Economics*, 1–30.
- Labonne, Julien, Sahar Parsa, and Pablo Querubin.** 2021. “Political dynasties, term limits and female political representation: Evidence from the Philippines.” *Journal of Economic Behavior & Organization*, 182: 212–228.
- Lassébie, Julie.** 2020. “Gender quotas and the selection of local politicians: Evidence from French municipal elections.” *European Journal of Political Economy*, 62: 101842.
- Le Barbanchon, Thomas, and Julien Sauvagnat.** 2022. “Electoral competition, voter bias, and women in politics.” *Journal of the European Economic Association*, 20(1): 352–394.
- Lippmann, Quentin.** 2023. “Persistence of incumbents and female access to political positions.” *Journal of Economic Behavior & Organization*, 206: 327–349.

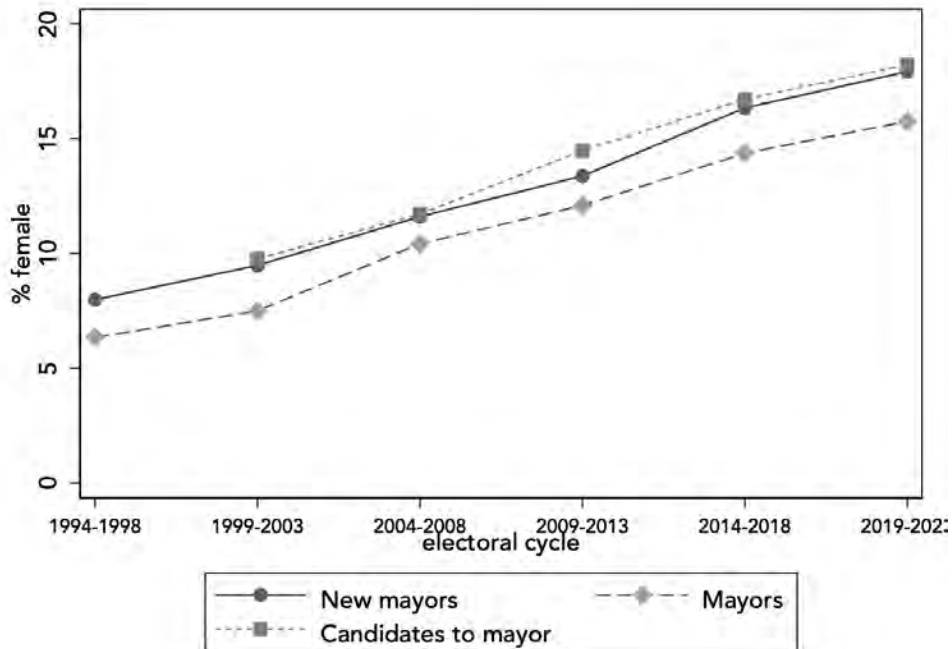
- Martinez-Bravo, Monica, Priya Mukherjee, and Andreas Stegmann.** 2017. “The non-democratic roots of elite capture: Evidence from Soeharto mayors in Indonesia.” *Econometrica*, 85(6): 1991–2010.
- O’Connell, Stephen D.** 2020. “Can quotas increase the supply of candidates for higher-level positions? Evidence from local government in India.” *Review of Economics and Statistics*, 102(1): 65–78.
- O’Brien, Diana Z, and Johanna Rickne.** 2016. “Gender quotas and women’s political leadership.” *American Political Science Review*, 110(1): 112–126.
- Pettey, Samantha.** 2018. “Female Candidate Emergence and Term Limits: A State-Level Analysis.” *Political Research Quarterly*, 71(2): 318–329.
- Rosenblum, Darren, and Yaron Nili.** 2019. “Board diversity by term limits.” *Alabama Law Review*, 71: 211.
- Roth, Jonathan, and Pedro HC Sant’Anna.** 2023. “When is parallel trends sensitive to functional form?” *Econometrica*, 91(2): 737–747.
- Schwindt-Bayer, Leslie A.** 2005. “The incumbency disadvantage and women’s election to legislative office.” *Electoral Studies*, 24(2): 227–244.
- Spaziani, Sara.** 2022. “Can gender quotas break the glass ceiling? Evidence from Italian municipal elections.” *European Journal of Political Economy*, 102171.
- Wasserman, Melanie.** 2023. “Gender differences in politician persistence.” *Review of Economics and Statistics*, 105(2): 275–291.
- Zamfir, Ionel.** 2016. *Democracy in Africa: Power alternation and presidential term limits*. EPRS, European Parliamentary Research Service, Members’ Research Service.

Figure 1: Female Political Representation

(a) Share of Female Politicians at Different Levels of the Political Hierarchy

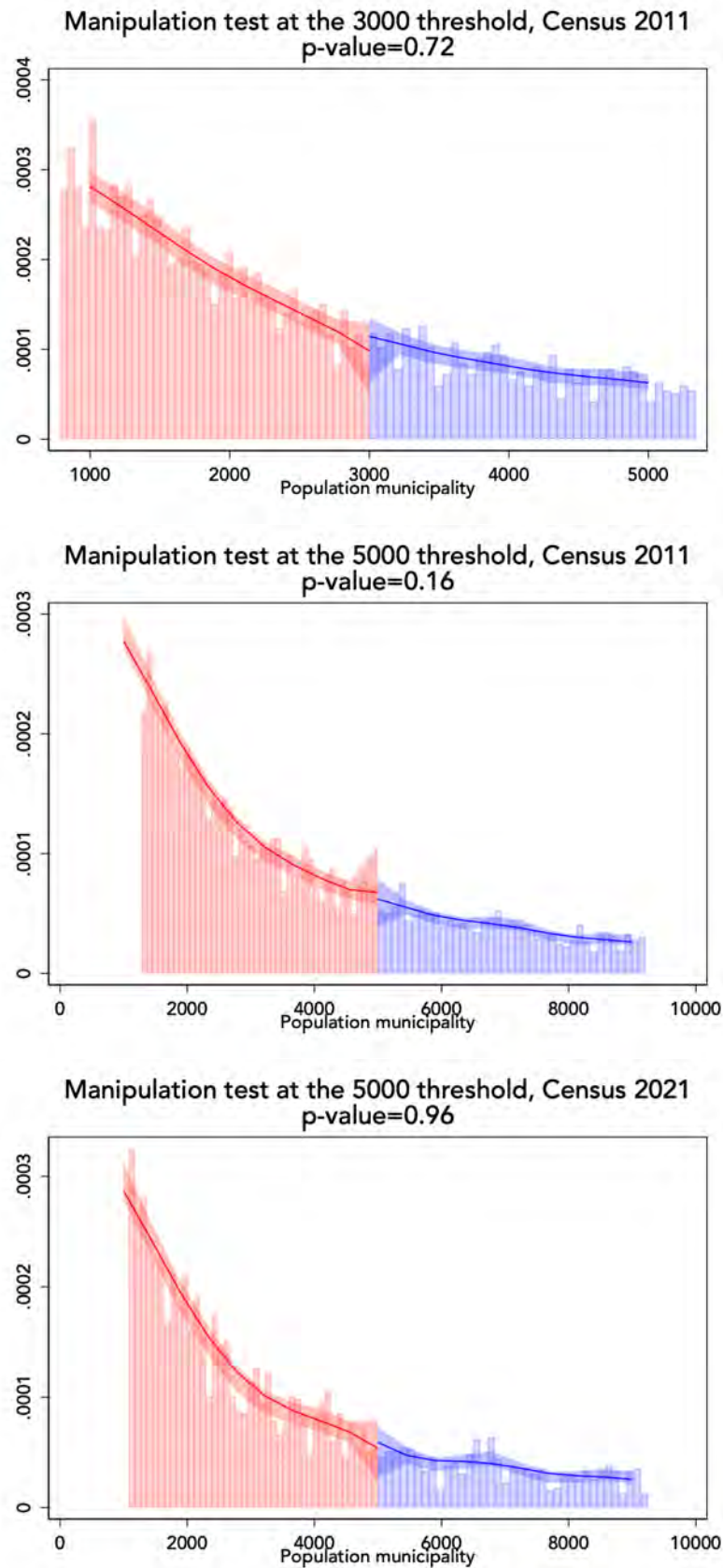


(b) Share of Women among Mayoral Candidates, New Mayors, and Incumbents



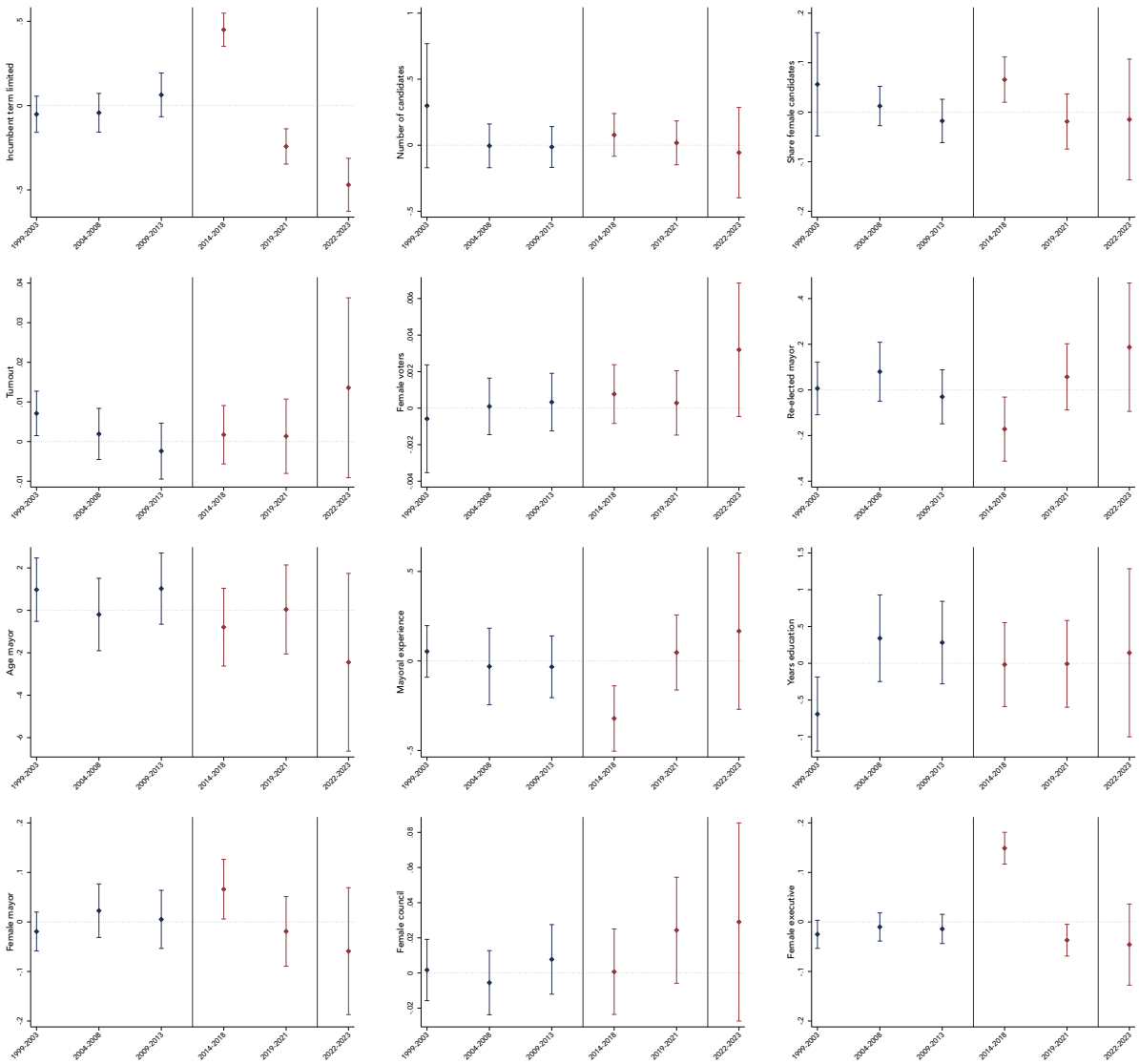
Note: Panel (a) shows the evolution of the share of women at different levels of local government: among mayors (short-dash line), executives (dash line), and councillors (solid line). Data is presented as yearly averages of municipalities having an election within the year on the x-axis. The two vertical lines indicate the introduction in 2013 of a gender quota for councillor candidates in municipalities with more than 5,000 inhabitants and the introduction in 2014 of a gender quota for executive committees in municipalities with more than 3,000 inhabitants. Panel (b) presents the share of women at different levels, grouped by election cycle. The solid line plots the average share of women among newly elected mayors, the dash line among all mayors, and the short-dash line among candidates for mayor.

Figure 2: Density Functions; 3,000 and 5,000 Inhabitants; Census 2011 and 2021



Note: The above figures report the results from manipulation testing procedures using the local polynomial density estimators proposed in Cattaneo, Jansson and Ma (2020) and the robust bias-corrected statistic. The top two figures report the test at the 3,000 and the 5,000 thresholds using the 2011 census. The bottom figure reports the test at the 5,000 threshold using the 2021 census. The sample includes all municipalities in the census.

Figure 3: Difference-in-Discontinuities Estimates, 3,000 Threshold, Years 1999-2023



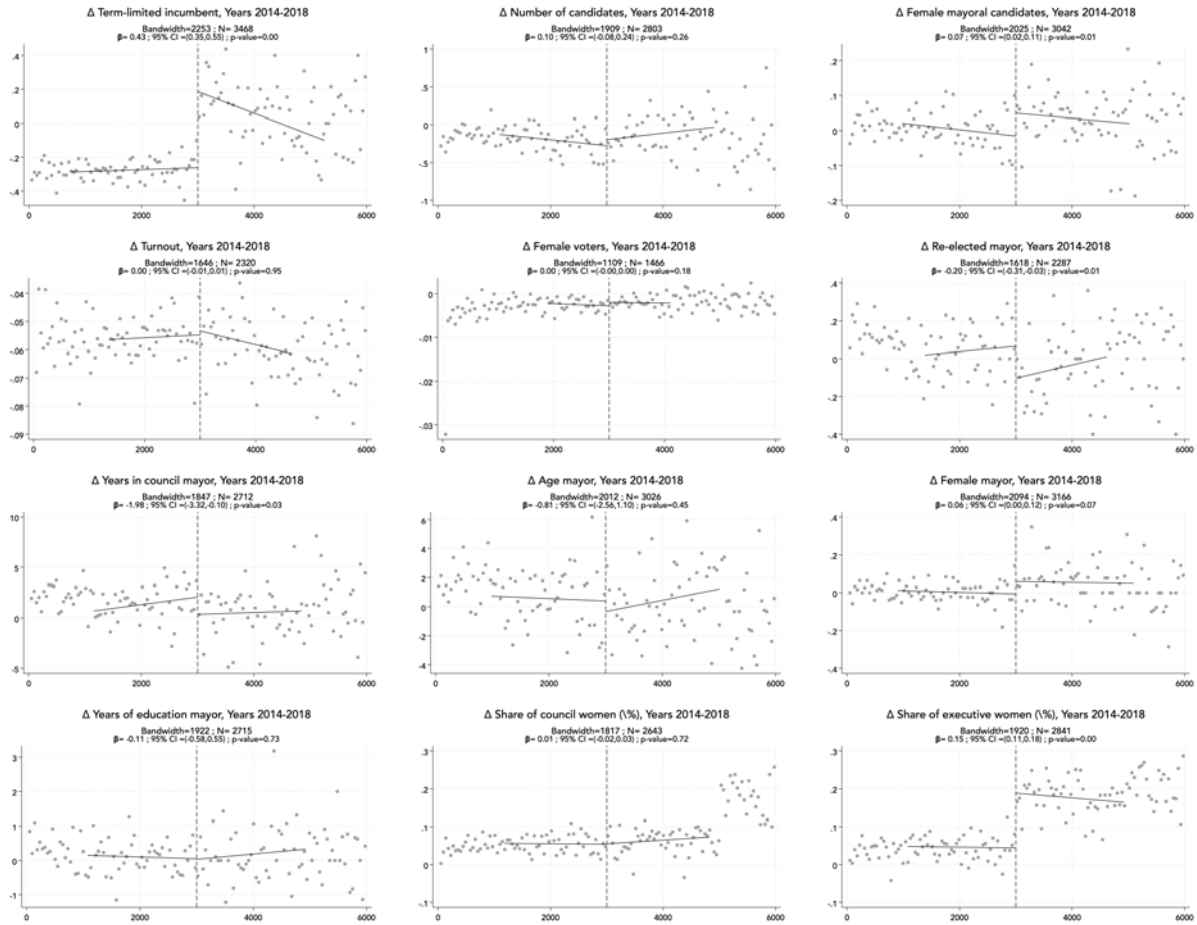
Note: The figure shows difference-in-discontinuities estimates and 95% confidence intervals around the 3,000 inhabitants threshold, based on equation (2). Estimates are shown separately for six periods: 1999-2003, 2004-2008, 2009-2013, 2014-2018, 2019-2021, and 2022-2023. The first vertical line marks the 2014 term limit extension reform, while the second vertical line indicates the 2022 reform that raised the threshold to 5,000 inhabitants.

Figure 4: Difference-in-Discontinuities Estimates, 5,000 Threshold, Years 1997-2023



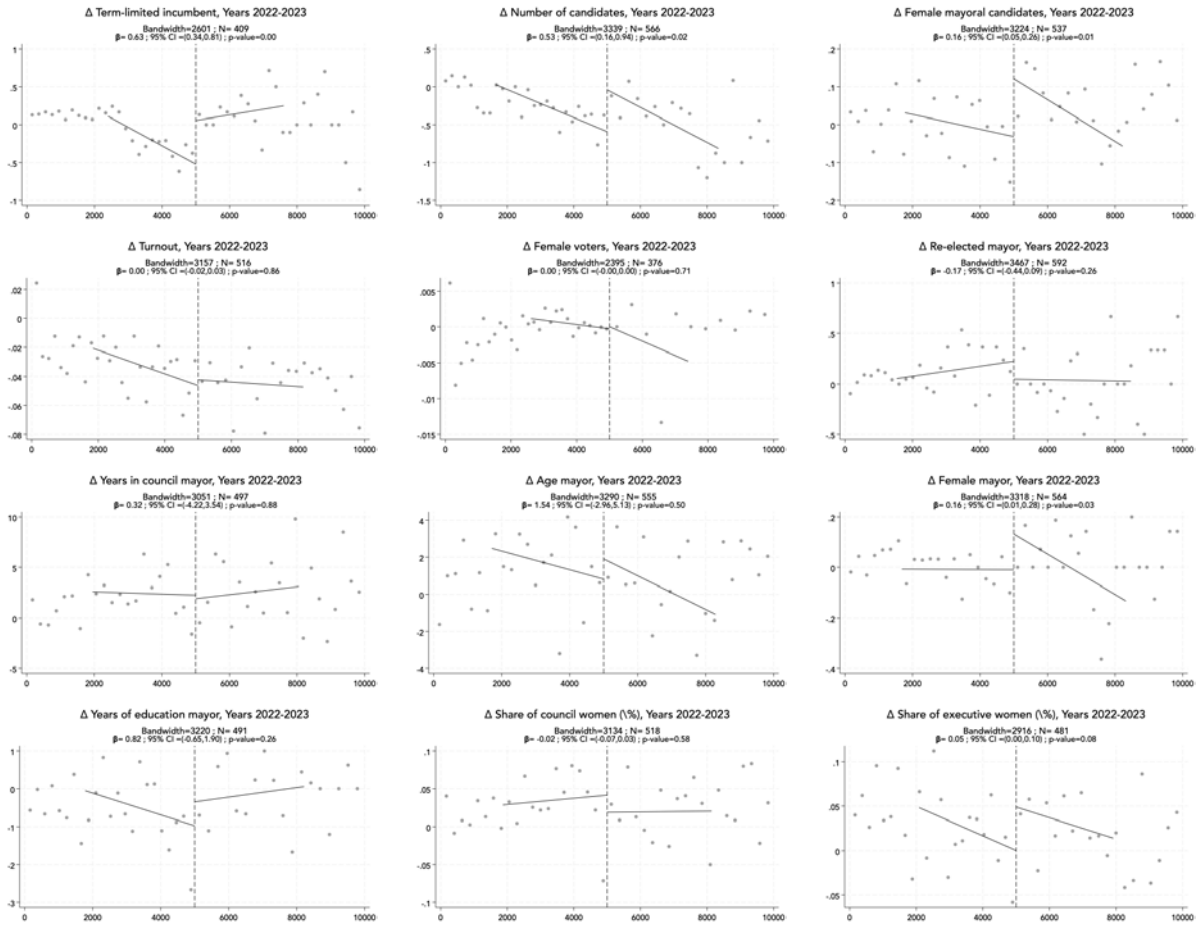
Note: The figure presents regression discontinuity estimates and their confidence intervals for municipalities around the 5000 inhabitants population threshold, that 2022 term extension reforms, using estimates from equation (2). Bars represent 95% confidence intervals. The first dotted line at the 2014 cohort represents the timing of the first term limit extension reform.

Figure 5: Difference-in-Discontinuities Plots, 3000 Cutoff, Years 2014-2018



Note: The above Regression Discontinuity (RD) plots report binned sample means for several outcome variables in differences, e.g. Δ *Term-limited incumbent* represents the difference between the value of this variable in the election that took place in year t and its value in the previous election (typically in $t - 5$). The RD plots report information above and below the 3,000 threshold (i.e. the population threshold for the 2014 reform, below which term limits were extended from two to three terms), in period 2014-2018. The plots also report the results from a difference-in-discontinuities estimation, based on the MSE optimal bandwidth proposed by Calonico, Cattaneo and Titiunik (2014) and using a first-degree polynomial and a uniform kernel.

Figure 6: Difference-in-Discontinuities Plots, 5000 cutoff, Years 2022-2023



Note: The above Regression Discontinuity (RD) plots report binned sample means for several outcome variables in differences, e.g. Δ *Term-limited incumbent* represents the difference between the value of this variable in the election that took place in year t and its value in the previous election (typically in $t - 5$). The RD plots include information above and below the 5,000 threshold (i.e. the population threshold for the 2022 reform, below which term limits were extended from two to three consecutive terms), in period 2022-2023. The plots also report the results from a regression discontinuity estimation, using the MSE optimal bandwidth proposed by Calonico, Cattaneo and Titiunik (2014), using a first-degree polynomial and a uniform kernel.

Table 1: Descriptive Statistics

	N	Mean	St. dev.	Min.	Max.
Turnout	22376	0.75	0.11	0.23	1
Female voters	18811	0.50	0.013	0.38	1
Number of candidates	18768	2.42	0.82	1	7
Age mayor	23368	49.0	10.2	18.6	91.3
Years of education mayor	22615	14.7	3.41	0	22
Same surname	23368	0.014	0.12	0	1
Terms as mayor	23368	0.62	0.83	0	6
Years in council mayor	23368	9.60	7.95	0	38
Term-limited incumbent	22686	0.31	0.46	0	1
Re-elected mayor	23275	0.39	0.49	0	1
Female mayoral candidates	18767	0.14	0.24	0	1
Female mayor	23368	0.11	0.31	0	1
Share of executive women (%)	23368	0.20	0.19	0	1
Share of council women (%)	23257	0.23	0.13	0	1

Notes: Summary statistics at the ‘municipality*election year’ level for 4,082 Italian municipalities with more than 1,000 inhabitants and less than 7,000 between 1992 and 2023. Data on candidates and voters is only available from 1994.

Table 2: Impact of the Extension of Term Limits

	(1)	(2)	(3)	(4)	(5)	(6)
	Incumbent term-limited	Number candidates	Female candidates	Voter turnout	Female voters	Incumbent re-elected
3000	0.43*** (0.05)	0.10 (0.09)	0.07** (0.03)	0.00 (0.00)	0.00 (0.00)	-0.20*** (0.08)
5000	0.63*** (0.13)	0.53** (0.23)	0.16*** (0.06)	0.00 (0.01)	0.00 (0.00)	-0.17 (0.15)
Weighted	0.45*** (0.05)	0.18** (0.09)	0.08*** (0.02)	0.00 (0.00)	0.00 (0.00)	-0.19*** (0.07)
N	3877	3369	3579	2836	2836	2879
	(7)	(8)	(9)	(10)	(11)	(12)
	Terms in council	Age mayor	Female mayor	Education mayor	Female council	Female executive
3000	-1.98** (0.91)	-0.81 (1.07)	0.06* (0.04)	-0.11 (0.32)	0.01 (0.01)	0.15*** (0.02)
5000	0.32 (2.16)	1.54 (2.30)	0.16** (0.07)	0.82 (0.73)	-0.02 (0.03)	0.05* (0.03)
Weighted	-1.62* (0.84)	-0.45 (0.97)	0.08** (0.03)	0.03 (0.29)	0.00 (0.01)	0.13*** (0.02)
N	2879	3581	3730	3206	3161	3322

Notes: The table includes estimates from different regressions using a difference-in-discontinuities design for twelve different outcome variables, six in each panel. In both panels, the first row (i.e. ‘3000’) provides estimates from the analysis of the impact of extending term limits at the 3,000 threshold on elections held in years 2014-2018, the second row (i.e. ‘5000’) the impact of extending term limits at the 5,000 threshold on elections held in years 2022-2023, and the third row (i.e. ‘Weighted’) reports the weighted average of both estimates. ‘N’ denotes the number of observations used in computing the weighted average. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 3: Heterogeneity Analysis, by Previous Incumbent's Term

Panel A. Previous Incumbent in 1st Term						
	(1)	(2)	(3)	(4)	(5)	(6)
	Incumbent re-elected	Female candidates	Experience in council	Female mayor	Female council	Female executive
Weighted	0.02 (0.05)	0.08*** (0.03)	1.54 (0.98)	0.05 (0.04)	0.01 (0.02)	0.12*** (0.02)
Panel B. Previous Incumbent in 2nd Term						
	(1)	(2)	(3)	(4)	(5)	(6)
	Incumbent re-elected	Female candidates	Experience in council	Female mayor	Female council	Female executive
Weighted	-0.33*** (0.05)	0.09** (0.04)	-4.42*** (1.46)	0.13** (0.06)	-0.02 (0.02)	0.14*** (0.03)

Notes: The table includes estimates from different regressions using a difference-in-discontinuities design. In the top (bottom) panel the sample includes elections in municipalities where the previous mayor was in their 1st (2nd) term. In both panels, we report the weighted estimate from the analysis of the impact of extending term limits at the 3,000 threshold on elections held in years 2014-2018 and the impact of extending term limits at the 5,000 threshold on elections held in years 2022-2023. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table 4: Heterogeneity Analysis, by Share of Women in Previous Councils

Panel A. Share of women in previous councils above 20%							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Incumbent term-limited	Incumbent re-elected	Female candidates	Experience in council	Female mayor	Female council	Female executive
Above 3,000	0.40*** (0.03)	-0.12** (0.06)	0.08*** (0.03)	-3.06*** (0.98)	0.09** (0.04)	-0.00 (0.01)	0.12*** (0.02)
Panel B. Share of women in previous councils below 20%							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Incumbent term-limited	Incumbent re-elected	Female candidates	Experience in council	Female mayor	Female council	Female executive
Above 3,000	0.36*** (0.03)	-0.12** (0.05)	0.01 (0.02)	-0.92 (1.03)	0.04 (0.03)	0.00 (0.01)	0.19*** (0.02)

Notes: The table includes difference-in-discontinuities design estimates for the effect of the term extension. In the top panel the sample includes municipalities where the share of women in the previous two councils was above 20%. In the bottom panel the sample includes municipalities where the share of female councilors was below 20%. Heteroskedasticity-robust standard errors are reported in parentheses, and statistical significance is denoted: *** p<0.01, ** p<0.05, * p<0.1.

Table 5: Medium-Term Effects

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Incumbent term-limited	Incumbent re-elected	Female candidates	Experience in council	Female mayor	Female council	Female executive
Above 3,000	0.11** (0.05)	-0.12** (0.06)	0.04 (0.03)	-1.16 (0.95)	0.03 (0.04)	0.03* (0.02)	0.11*** (0.02)

Notes: The table reports difference-in-discontinuities estimates (equation 2) of the medium-term effects of the 2014 term limit extension, comparing outcomes between 2009-2013 and 2019-2023. Heteroskedasticity-robust standard errors are reported in parentheses, and statistical significance is denoted: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

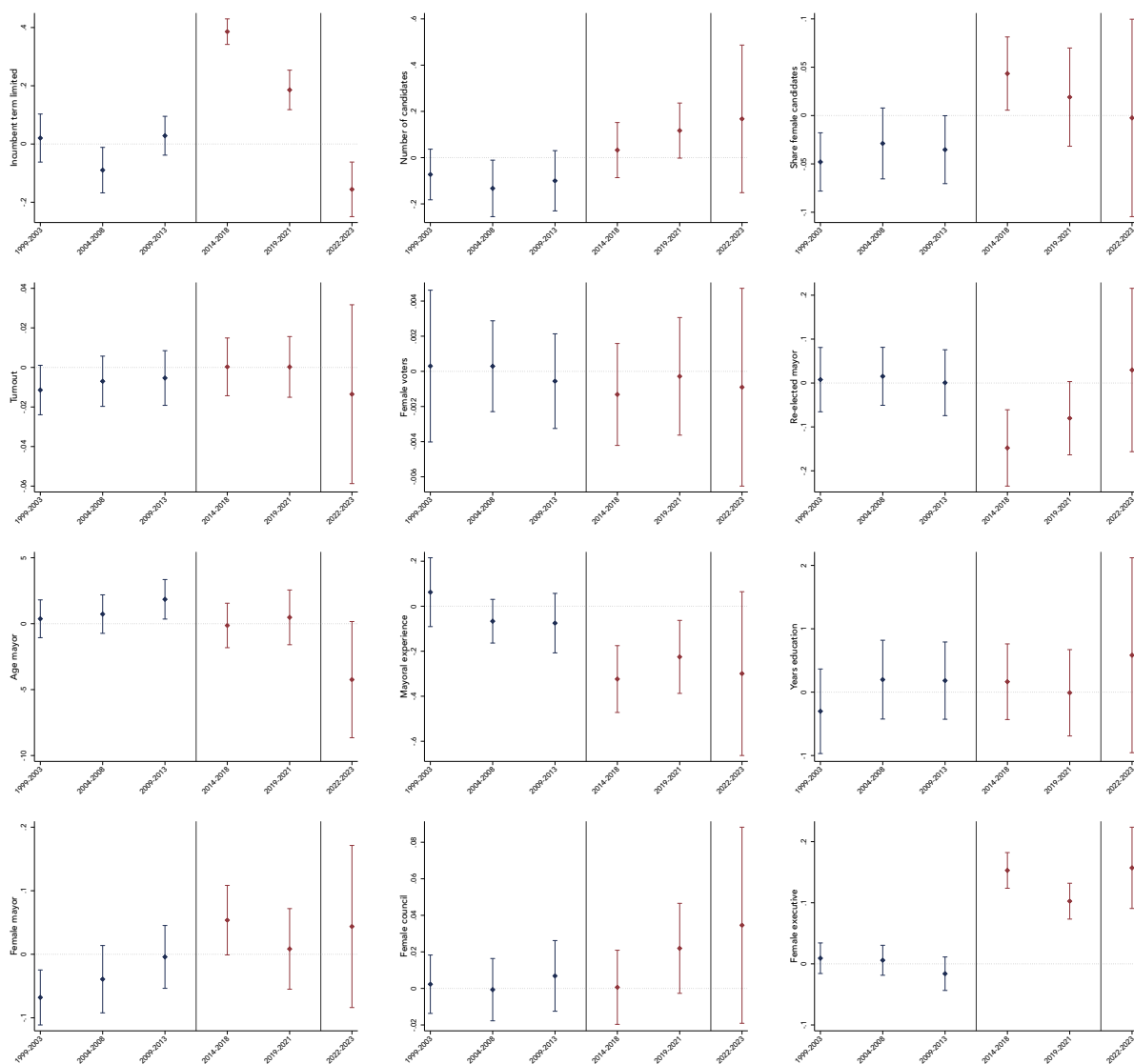
Table 6: Disentangling the Effects of Term Limits and Gender Quotas in Candidate Lists

	(1)	(2)	(3)	(4)	(5)	(6)
	Incumbent term-limited	Female candidates	Female mayor	Incumbent term-limited	Female candidates	Female mayor
Above 5,000	-0.09 (0.06)	0.01 (0.03)	0.04 (0.05)	0.48*** (0.15)	0.06 (0.07)	0.27*** (0.09)
Years	2018-2021	2018-2021	2018-2021	2022	2022	2022

Notes: The table presents difference-in-discontinuities estimates at the 5,000 inhabitants threshold for two periods: 2018-2021 (columns 1-3) and 2022 (columns 4-6). In 2018-2021, municipalities above the 5,000 threshold were subject to gender quotas in candidate lists for the second time. In 2022, municipalities above the 5,000 threshold continued to face these gender quotas for the second time, while municipalities below this threshold were affected by the term limit extension. Heteroskedasticity-robust standard errors are reported in parentheses, and statistical significance is denoted: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

A Appendix Figures

Figure A.1: Regression Discontinuity Design, 3,000 Threshold



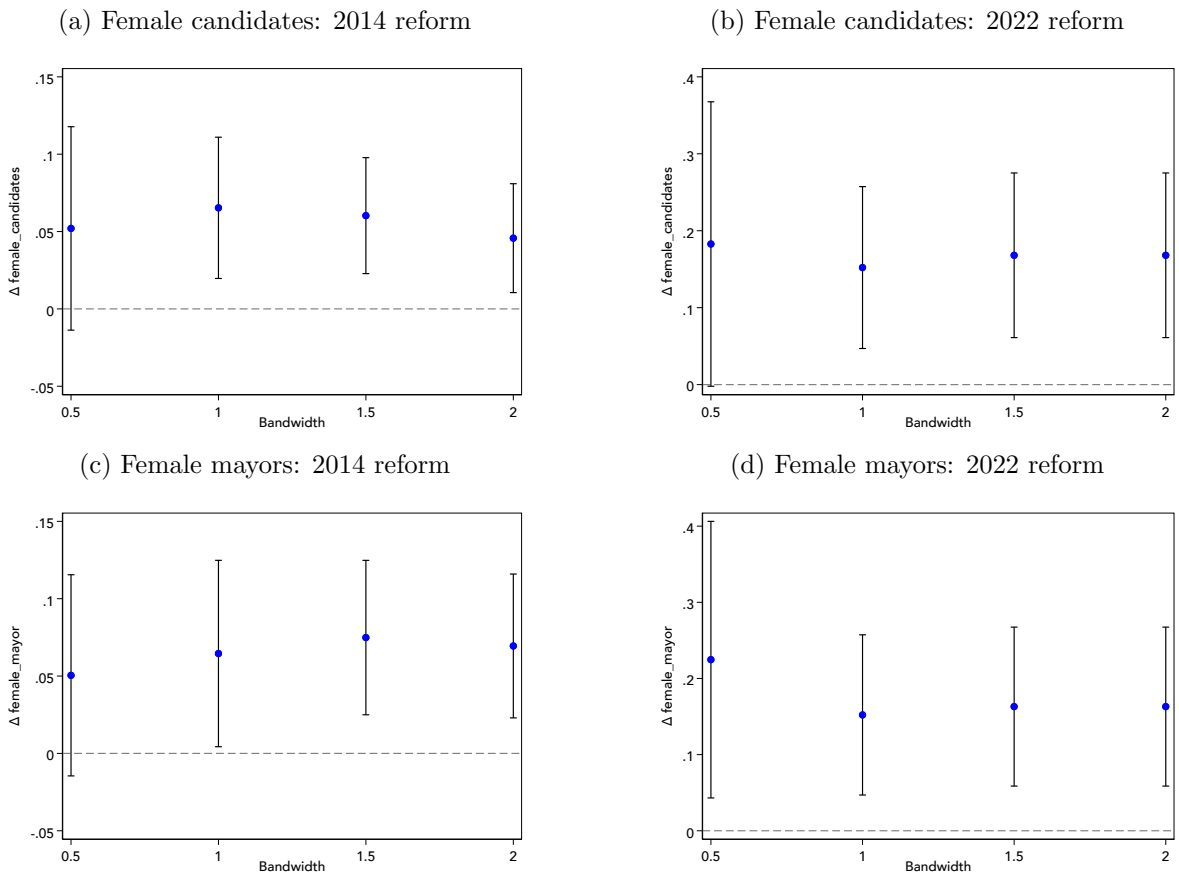
Note: These figures present regression discontinuity estimates and confidence intervals for municipalities around the 3,000 inhabitants population threshold using estimates from equation (1) across election cycles. Bars represent 95% confidence intervals. The first vertical line marks the 2014 term limit extension reform, and the second vertical line indicates when the threshold for a three-term limit was raised to 5,000 inhabitants.

Figure A.2: Regression Discontinuity Design, 5,000 Threshold



Note: These figures present regression discontinuity estimates and confidence intervals for municipalities around the 5,000 inhabitants population threshold using estimates from equation (1) across election cycles. Bars represent 95% confidence intervals. The second vertical line indicates when the threshold for a three-term limit was fixed at 5,000 inhabitants.

Figure A.3: Sensitivity of Estimates to Choice of Bandwidth



Note: These figures plot coefficient estimates and confidence intervals under different choices of bandwidth. The outcome variables are the first difference in female mayors for both reforms, with the 2014 reform exploiting the 3000 population cutoff and the 2022 reform exploiting the 5000 population cutoff. The x -axis shows the ratio of the chosen bandwidth to the optimal bandwidth used for sensitivity analysis. 1 represents the optimal bandwidth. Dots represent point estimates and whiskers represent 95% confidence intervals.

B Appendix Tables

Table B.1: Robustness: Kernel and Polynomial Order

	(1)	(2)	(3)	(4)
	Female	Female	Female	Female
	candidates	candidates	mayor	mayor
3000	0.07**	0.07**	0.06*	0.06*
	(0.03)	(0.03)	(0.04)	(0.03)
5000	0.16***	0.16***	0.16**	0.16**
	(0.06)	(0.06)	(0.07)	(0.07)
Weighted	0.08***	0.08***	0.08**	0.08**
	(0.02)	(0.02)	(0.03)	(0.03)
Kernel	Uniform	Triangular	Uniform	Triangular
Polynomial order	2	1	2	1

Note: The table reports estimates for the share of female candidates (columns 1 and 2) and female mayors (columns 3 and 4), for different specifications of kernel (uniform or triangular) and polynomial (first or second order). Standard errors clustered at the municipality level in parentheses. Statistical significance is denoted: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

Table B.2: Impact of the Extension of Term Limits in 2014, ‘Mountain Municipalities’

Panel A. Non mountain municipalities						
	(1)	(2)	(3)	(4)	(5)	(6)
	Incumbent term-limited	Incumbent re-elected	Female candidates	Female mayor	Female council	Female executive
Above 3000	0.40*** (0.03)	-0.09* (0.06)	0.04 (0.03)	0.05 (0.04)	0.00 (0.01)	0.15*** (0.02)
Panel B. Mountain municipalities						
	(1)	(2)	(3)	(4)	(5)	(6)
	Incumbent term-limited	Incumbent re-elected	Female candidates	Female mayor	Female council	Female executive
Above 3000	0.41*** (0.05)	-0.20*** (0.07)	0.07* (0.04)	0.06 (0.05)	0.01 (0.02)	0.15*** (0.03)

Note: The top panel shows regression estimates from a differences-in-discontinuities design at the 3,000 threshold, years 2014-2018, estimated on the sample of municipalities that have not been legally defined as ‘mountain municipalities’. The bottom panel reports a similar analysis for ‘mountain municipalities’. Standard errors clustered at the municipality level in parentheses. Statistical significance is denoted: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

C Data

C.1 Electoral data: administrators

The data used for the analysis was collected from the Open Data section of the Eligendo portal of Italian elections provided by the Italian Ministry of Interior. The dataset includes the universe of Italian administrators on the 31st December each year, and contains information on their name, age, gender, professional title, educational achievement, party affiliation and assigned position. The dataset was cleaned for consistency and appended for years 1992-2023. We also collected population information directly from the Census.

C.2 Electoral data: candidates

The data used for the analysis was collected from the Eligendo portal of Italian elections, and complemented with data received directly from the Ministry of the Interior. The dataset contains information on candidates for mayoral positions, including their age and gender, party affiliation, turnout, and votes received.

D Difference-in-differences

In this appendix we report results using a difference-in-differences (DID) strategy. Our DID and RD-DID estimates are very similar, but more precise in the former case.

To ensure that we compare municipalities with similar institutional characteristics and to minimize the possibility that they are exposed to differential time-variant shocks, we restrict our sample to municipalities within 2,000 inhabitants of the population thresholds used for extending term limits. More precisely, to estimate the effect of the 2014 reform, which extended term limits in municipalities below 3,000 inhabitants, we compare the evolution of (i) municipalities with more than 1,000 and less than 3,000 inhabitants (in what follows, ‘small’ municipalities) versus municipalities with more than 3,000 and less than 5,000 inhabitants (‘medium’ municipalities). This sample restriction excludes municipalities above 5,000 inhabitants, which were affected by the introduction of quotas in candidate lists and double preference voting in 2013, as well as municipalities with less than 1,000 inhabitants, which was the threshold used for the application of the *Domestic Stability Pact* in 2013. This sample includes around 3,500 municipalities. Similarly, in our analysis of the 2022 reform, which extended term limits in municipalities with more than 3,000 and less than 5,000 inhabitants, we compare (i) municipalities with more than 3,000 and less than 5,000 inhabitants (‘medium’ municipalities) versus municipalities with more than 5,000 and less than 7,000 inhabitants (‘large’ municipalities). In this case we are only able to include in our study municipalities that had elections in 2022 and 2023. There are around one 320 municipalities in this sample.

As we show in more detail below, our analysis suggests that DID is an appropriate empirical strategy in this context. The treatment and control groups were very similar before term limits were extended, both in terms of the average values for the main outcome variables as well as their evolution over time, providing support for the parallel trends assumption.¹⁹ The absence of significant pre-existing differences in levels strengthens the credibility of our difference-in-differences strategy, as it implies that parallel trends tests are not sensitive to functional form assumptions which, as Roth and Sant’Anna (2023) point out, are often non-trivial. For instance, in our context, it is not immediately clear whether the outcome variable should be measured in percentage point changes or percent changes. While we use the former, the lack of pre-existing differences between treatment and control groups ensures that our results are robust to this choice.

D.1 Evolution over time of main characteristics, by size of the municipality

We start by showing graphically the evolution over time of the main variables in the treatment and control groups. These graphs provide suggestive visual evidence indicating that (i) while there are no significant differences between the control and treatment groups prior to the policy change, (ii) both groups diverge after the policy change. Figures D.1 and D.2 provide descriptive

¹⁹In some settings, larger municipalities may not provide a valid counterfactual for the evolution of smaller ones. For instance, in the context of Spanish local elections, Bagues and Campa (2020) show that voting behavior evolves quite differently in large and small municipalities, invalidating the consistency of standard DID estimates that rely on population thresholds to assess the impact of policy changes on voting behavior, e.g. the introduction of gender quotas in Spain in municipalities with less than 5,000 inhabitants in 2007.

information on the evolution over time of the main variables by municipality size, for the 2014 and the 2022 reforms respectively. Given that there are some years with very few elections, for clarity in our graphical analysis we group elections by electoral cycles. When we analyze the 2014 reform, we consider as the baseline electoral cycle years 2014-2018 (i.e. the first cycle of elections after the reform). In our analysis of the 2022 reform, we restrict our sample to municipalities that had one election after the reform, i.e. years 2022-24, and track them in previous electoral cycles.

As illustrated in both figures, municipalities above and below the threshold appear highly comparable prior to the policy change across all dimensions except one. The main difference lies in the number of candidates, which seems to evolve in parallel in the treatment and control groups, but it tends to be greater in larger municipalities. This observation aligns with the policy’s motivation, which was driven by the scarcity of mayoral candidates in smaller municipalities. More precisely, before 2014 there are around 2.3 candidates in small municipalities, compared to 2.6 in medium-sized ones. Similarly, before 2022 the number of candidates was around 2.9 in large municipalities, compared to 2.6 in medium size ones.

The other dimensions appear to be balanced across the treatment and control groups, both at the 3,000 threshold and the 5,000 threshold. Around 35% of mayors were term-limited before terms limits were extended. As expected, the share of term-limited mayors decreases to zero in treated municipalities in the first election after the reform, while in the control group it remains roughly constant. The gap between both groups is slightly smaller in the second cycle of elections, which we can observe in the case of the 2014 reform (years 2019-2021), reflecting that in treated municipalities some mayors reached their third term and became term-limited. The probability that the mayor is re-elected follows a very similar pattern. Before the reforms there is no apparent difference between the treatment and control groups, but a gap arises when term limits are extended in smaller municipalities.

Mayors in treatment and control groups were also very similar before the reform in terms of their age and their political experience, as measured by the number of years since they entered the council. However, a gap seems to appear when term limits are extended, with a relative increase in the political experience and age of mayors in ‘treated’ municipalities.

Finally, the share of female candidates and female mayors, which was growing in all municipalities prior to the reforms, stagnates when term limits are extended, while in municipalities in the control group it keeps growing at the same speed as before.

D.2 Balance test

The above descriptive graphs visually suggest that there were no differences between the treatment and control groups prior to the policy change, except for the number of candidates. Next we formally test whether the treatment and control groups were similar in levels before the intervention, controlling for electoral year fixed effects and clustering standard errors at the municipal level. For the 2014 reform, we compare ‘small’ and ‘medium’ municipalities in years 1993-2013. Similarly, for the 2022 reform, we compare ‘medium’ and ‘large’ municipalities between 1993 and 2021. As shown in Table D.1, the treatment and control groups are statistically similar before the reform in all dimensions except for the number of candidates and for a small

difference in the probability of being term-limited in 2022. The lack of pre-existent differences lends support to our empirical strategy. As the main outcome variables were similar before the introduction of the policy, it is more reasonable to expect that the two groups of municipalities would have evolved similarly, absent the introduction of the policy.

D.3 Event study analysis

We conduct an event study analysis to examine the validity of the parallel trends assumption. In particular, we estimate the following equation:

$$(3) \quad Y_{me} = \alpha_m + \beta_e + \sum_{k=T_0}^{-2} \gamma_k \times TreatmentGroup_m + \sum_{k=0}^{T_F} \gamma_k \times TreatmentGroup_m + \varepsilon_{me},$$

where Y_{ie} is the outcome of interest (e.g. whether a female mayor was elected in municipality m in the election that took place in electoral cycle e), α_e and β_m represent a set of dummies for electoral cycles and municipalities respectively, and T_0 and T_F represent respectively the first and the last electoral cycle available in the dataset. *TreatmentGroup* is a binary variable taking value 1 for municipalities what were affected by the extension of term limits (e.g. municipalities with more than 1,000 and less than 3,000 inhabitants when we analyze the 2014 reform). Standard errors ε_{me} are clustered at the municipality level.

Overall, the event study estimates support the validity of a difference-in-differences strategy in this context (see Figures D.3 and D.4). Out of 42 tests, we only observe two significant coefficients at the 5% level, which is consistent with the underlying probability of finding false positives. One of the violations corresponds to the evolution in the share of female candidates in small and medium municipalities between 2004-2008 and 2009-2013, and the other to the evolution in the number of candidates in large and medium municipalities between 2012-2016 and 2017-2021.

D.4 Difference-in-differences estimates

We estimate the following difference-in-differences equation separately for each reform:

$$(4) \quad Y_{mt} = \alpha_m + \beta_t + \gamma Three\ Terms_{mt} + \varepsilon_{mt},$$

where Y_{mt} is the outcome variable of interest, α_m and β_t represent two sets of municipality and year fixed effects, and *Three Terms* is a binary variable taking value 1 when the municipality is subject to a three-term limit. Standard errors are clustered at the municipality level.

In our analysis of the 2014 reform we restrict the sample to elections between 2004 and 2019 in municipalities with more than 1,000 and less than 5,000 inhabitants. We exclude elections in years 2020 onwards, as municipalities with less than 3,000 inhabitants experienced this year a small increase in mayoral salaries and in 2022 they were affected by the extension of term limits. In the post-treatment period, years 2014-2018 provide information on the impact within the first cycle of elections that took place after the reform, and year 2019 informs about the

effect of the reform in the second cycle of elections.²⁰ In our analysis of the 2022 reform we consider municipalities between 3,000 and 7,000 inhabitants in years 2013-2023. We exclude observations before 2013 given that in this year gender quotas in candidate lists were introduced in municipalities above 5,000 inhabitants. We report results in Table D.2 (i.e. 2004-2019 for the 3,000 threshold and 2013-2023 for the 5,000 threshold).

We start by quantifying the impact of the term limit extension on the probability that the incumbent mayor can run again and is re-elected. As shown in the upper panel of Table D.2, column 1, the extension of term limits in small municipalities in 2014 decreases the probability that the incumbent is term-limited by 32 p.p (s.e.=1 p.p.). The impact of the 2022 reform is slightly larger, a 46 p.p. decrease (s.e.=3 p.p.), as shown in the bottom panel of Table D.2, column 1. We observe a decrease in the number of candidates at the 3,000 threshold, -0.08 p.p. (s.e.=0.03) and no significant change at the 5,000 threshold (0.04, s.e.=0.06).

There is a decrease in the share of female candidates at both thresholds, 2 p.p. (s.e.=1 p.p.) at the 3,000 threshold and 6 p.p. (s.e.=2 p.p.) at the 5,000 one, compared to larger municipalities unaffected by the reform.

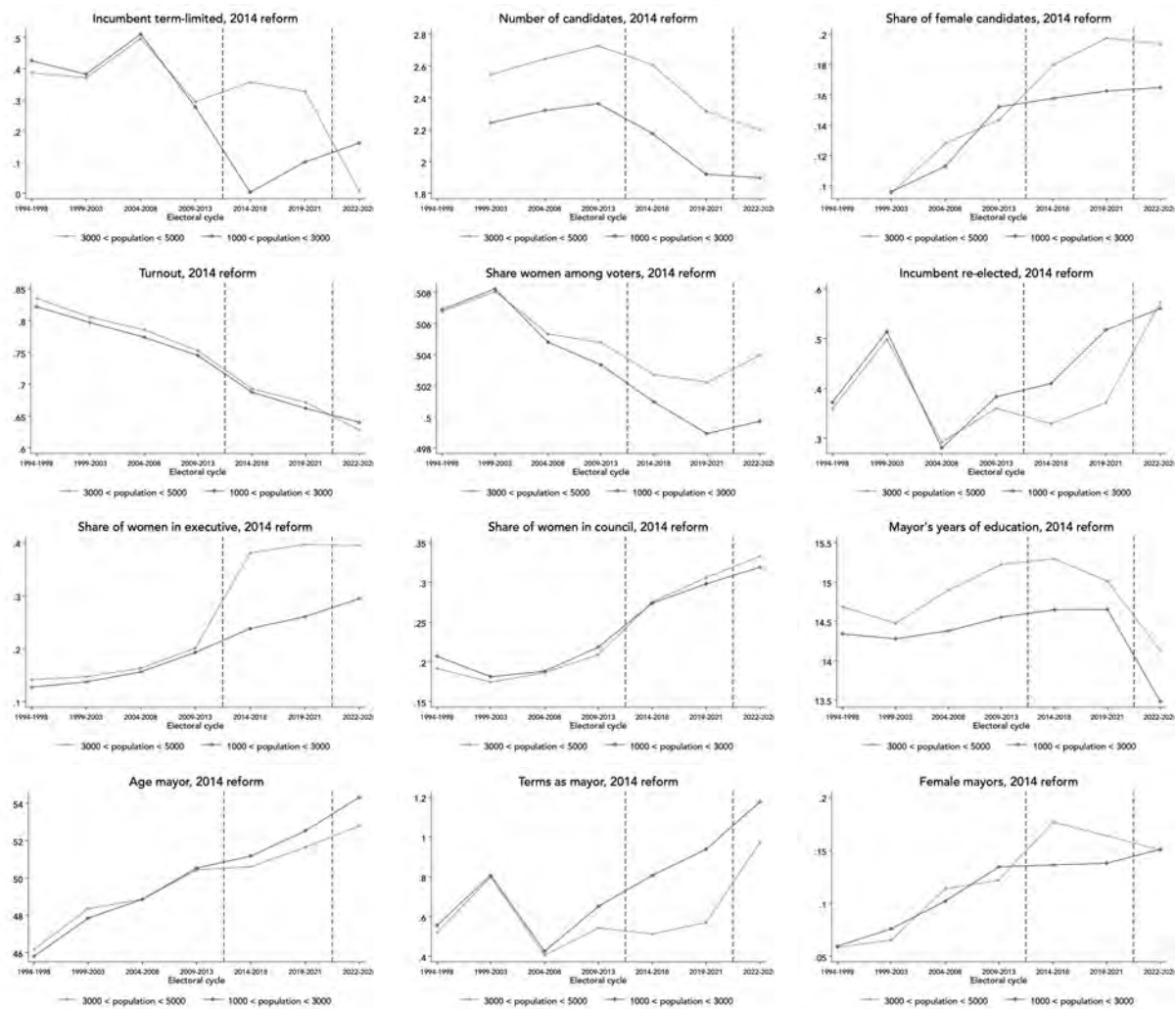
To induce a replacement effect, the reform should also affect the probability that incumbents are re-elected. Indeed, we find that the reform increased the probability that the incumbent mayor was re-elected by 11 p.p. in 2014 (s.e.: 2 p.p.) and by 5 p.p. in 2022 (s.e.: 4 p.p.), although the latter effect is not statistically significant.

Next, we analyze the impact on the characteristics of mayors. The extension of term limits helped to retain mayors that are older and also more experienced. In 2014 the reform increase the average age of mayors by 0.86 years (s.e.=0.45) and in 2022 by 0.01 (s.e.=0.70 p.p.), although only the former coefficient is statistically significant. The number of years since the mayor had entered the municipal council increased by 1.12 (s.e.=0.35) when term limits were extended in 2014, and by 0.66 (s.e.=1.07) with the 2022 extension.

Finally, we also observe a clear and significant decrease in the share of female mayors in smaller municipalities right after the reform, compared to larger ones that were not affected by the extension of term limits. The share of female mayors decreased by 4 p.p. in municipalities affected by the 2014 reforms (s.e.: 1 p.p.), compared to an average of 17% in the control group. In 2022 the effect is around 10 p.p. (s.e.: 4 p.p.), relative to a 16% baseline. The effect of the 2022 reform is substantially larger in magnitude than the 2014 reform. We return to this observation in the heterogeneity section, where we focus on how supply considerations - the presence of women at lower positions of the political hierarchy - on the impact of term limits.

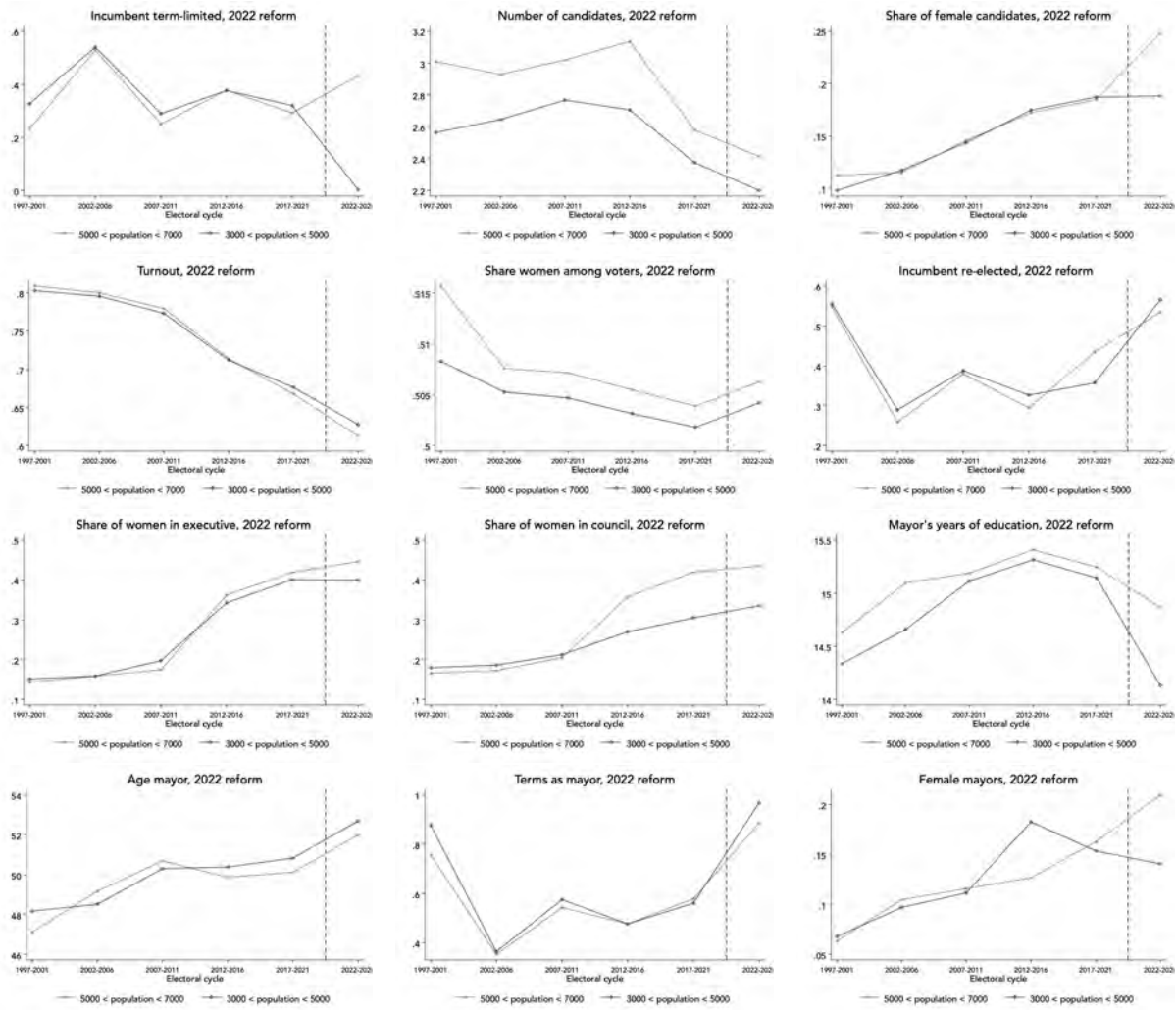
²⁰In a small number of municipalities there were two elections between 2014-2018 due to the resignation of the mayor or some other exceptional event requiring a new election before the end of the standard five year term.

Figure D.1: Main Characteristics over Time, Small and Medium Municipalities



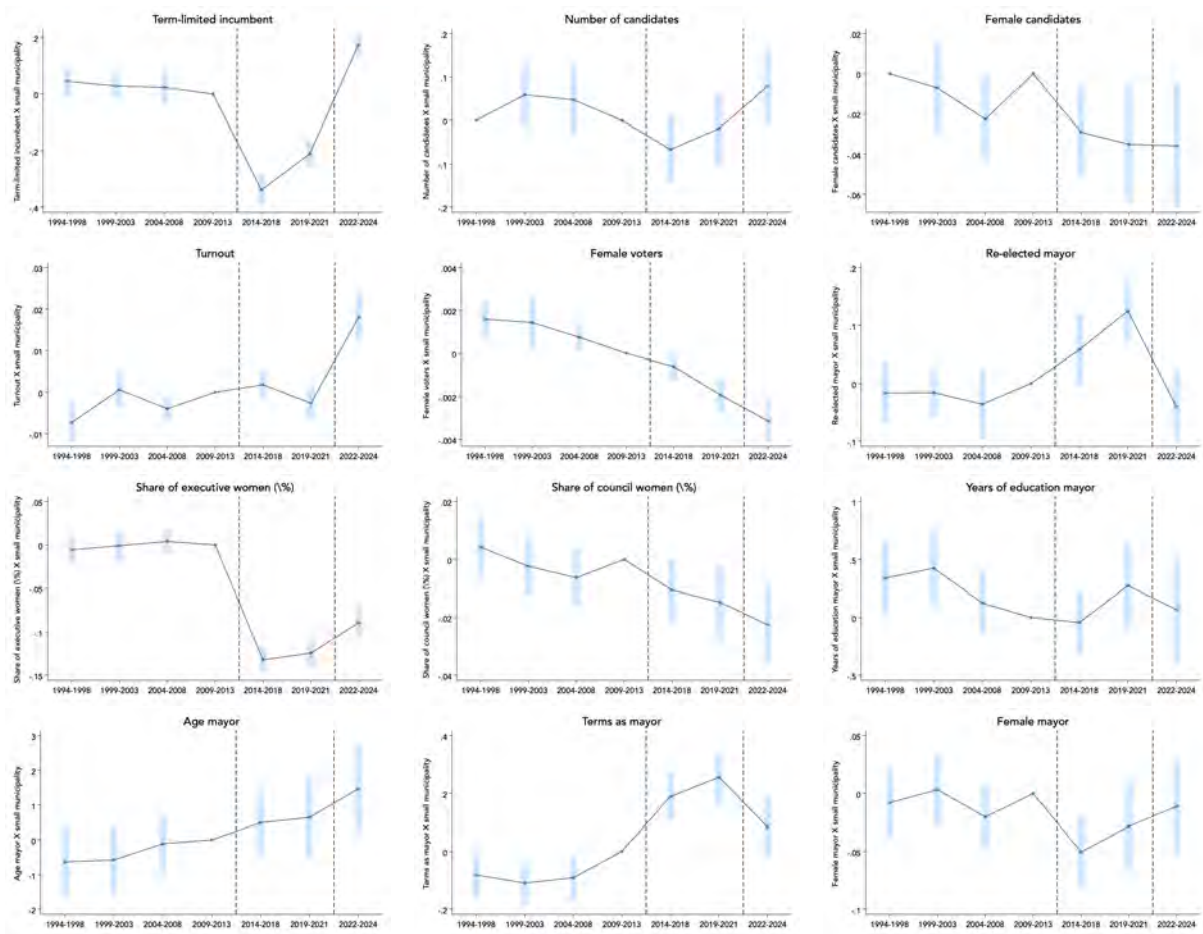
Note: The above figures provide the average value for the main outcome variables in different electoral cycles by size of the municipality. Small municipalities (more than 1,000 and less than 3,000 inhabitants) are represented by circle gray bins and medium ones (more than 3,000 and less than 5,000 inhabitants) by diamond navy bins. The vertical line at year 2014 depicts the time at which term limits were extended in municipalities below 3,000 inhabitants.

Figure D.2: Main Characteristics over Time, Medium and Large Municipalities



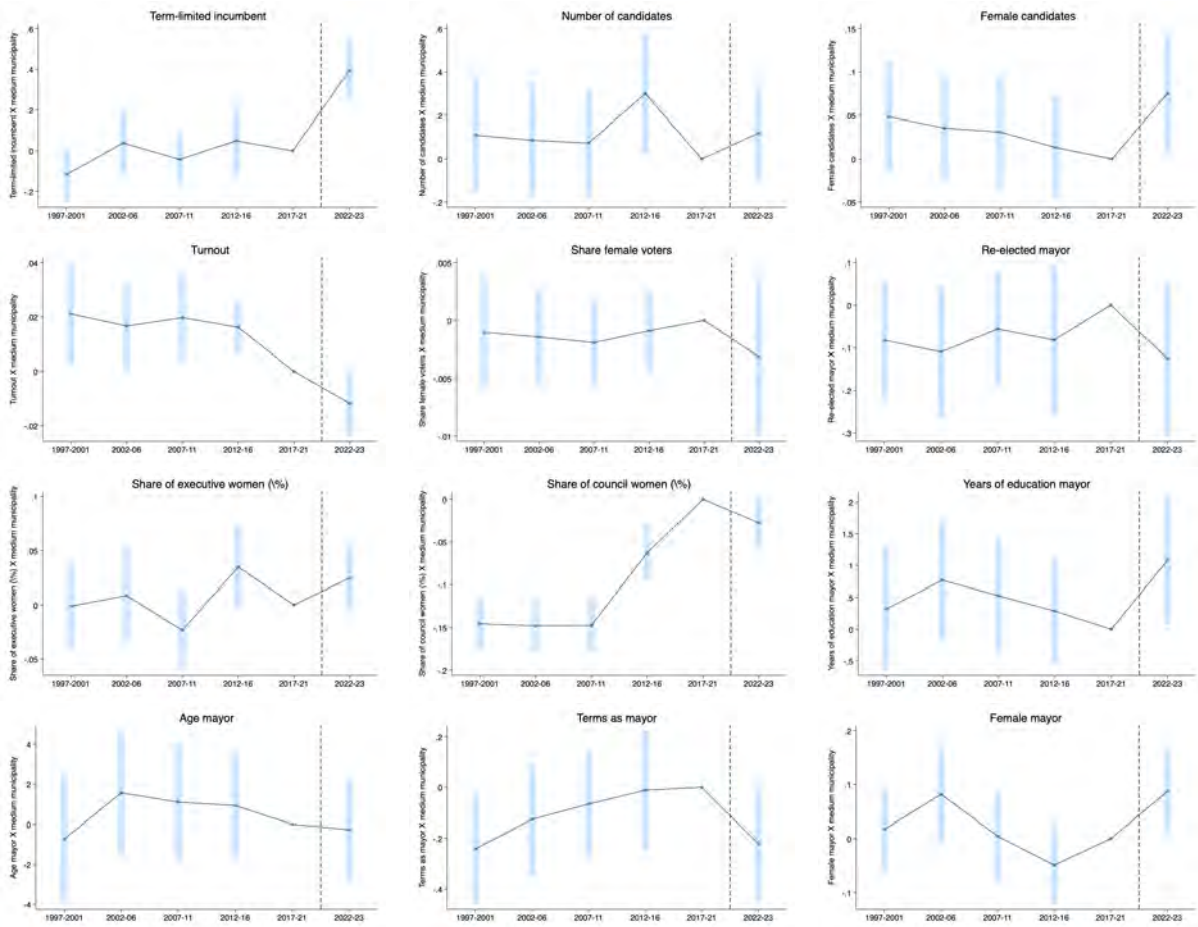
Note: The above figures display the main characteristics over time for two groups of municipalities: with more than 3,000 and less than 5,000 inhabitants ('medium') and with more than 5,000 and less than 7,000 ('large'). Each bin represents the average within an election cycle (5 years). The vertical line at year 2022 depicts the time when term limits were extended in municipalities with less than 5,000 and more than 3,000 inhabitants.

Figure D.3: Event Study Graphs, 2014 Reform



Note: The figure presents event study graphs for the 2014 term extension reforms, using estimates from equation 3. Coefficients are normalized to zero in t-1. Bars represent 95% confidence intervals. The red vertical line represents the time of reform, 2014.

Figure D.4: Event Study Graphs, 2022 Reform



Note: The figure presents event study graphs for the 2022 term extension reforms, using estimates from equation 3. Coefficients are normalized to zero in t-1. Bars represent 95% confidence intervals. The red vertical line represents the time of reform, 2022.

Table D.1: Balance Tests: Treatment and Control Groups Before the Reforms

Panel A. Before the 2014 reform					
	3,000 < population < 5,000		1,000 < population < 3,000		Difference
	N	Mean	N	Mean	
Term-limited incumbent	3917	0.39	8715	0.40	0.010
Number of candidates	2931	2.64	6561	2.31	-0.328***
Female candidates	2935	0.12	6562	0.12	-0.002
Re-elected mayor	3917	0.38	8715	0.39	0.008
Age mayor	3919	48.34	8721	48.14	-0.212
Years in council mayor	3919	8.38	8721	8.43	0.016
Female mayor	3919	0.09	8721	0.09	0.003

Panel B. Before the 2022 reform					
	5,000 < population < 7,000		3,000 < population < 5,000		Difference
	N	Mean	N	Mean	
Term-limited incumbent	2755	0.34	4873	0.36	0.019**
Number of candidates	2635	2.90	4683	2.58	-0.315***
Female candidates	2633	0.15	4687	0.15	-0.007
Re-elected mayor	2755	0.38	4873	0.37	-0.010
Age mayor	2773	49.20	4890	49.70	0.409
Years in council mayor	2773	9.84	4890	10.23	0.285
Female mayor	2773	0.11	4890	0.12	0.006

Note: The two panels report a balance test for municipalities in the control group (columns 2 and 3) and in the treatment group (columns 4 and 5) for the 2014 reform (Panel A) and the 2022 reform (Panel B). Panel A provides information for years 1993-2013 for municipalities with more than 3,000 and less than 5,000 inhabitants according to the 2011 census (columns 2-3) and for municipalities with more than 1,000 and less than 3,000 (columns 4-5). Panel B provides information for years 1993-2021 for municipalities with more than 5,000 and less than 7,000 inhabitants according to the 2011 census (columns 2-3) and municipalities with more than 3,000 and less than 5,000 inhabitants (columns 4-5). In both panels, column 6 indicates the difference between the corresponding treatment and control groups, conditional on years fixed effects, with standard errors clustered at the municipality level. Stars indicate whether this difference is significant: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D.2: Impact of the Extension of Term Limits, Difference-in-Differences Estimates

Panel A. Extension of term limits in 2014							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Incumbent term-limited	Number candidates	Female candidates	Incumbent re-elected	Years in council	Age mayor	Female mayor
Three terms	-0.32*** (0.01)	-0.08*** (0.03)	-0.02** (0.01)	0.11*** (0.02)	1.12*** (0.35)	0.87* (0.45)	-0.04** (0.01)
Observations	11,118	11,102	11,102	11,118	11,136	11,136	11,136
Mean	0.353	2.505	0.189	0.345	12.01	50.81	0.175
Panel B. Extension of term limits in 2022							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Incumbent term-limited	Number candidates	Female candidates	Incumbent re-elected	Years in council	Age mayor	Female mayor
Three terms	-0.39*** (0.05)	0.03 (0.10)	-0.07** (0.03)	0.10 (0.07)	-0.65 (1.08)	0.12 (1.23)	-0.11*** (0.04)
Observations	4,597	4,623	4,621	4,597	4,647	4,647	4,647
Mean	0.345	2.399	0.185	0.446	13.97	51.84	0.163

Note: *Three terms* is a dummy that takes value one when the term limit has been extended to three terms. Panel A includes results from regressions in the sample of elections between 2004 and 2019 in municipalities with more than 1,000 and less than 5,000 inhabitants. Panel B includes elections between 2013 and 2023 in municipalities with more than 3,000 and less than 7,000 inhabitants. Standard errors clustered at the municipality level in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$