ISSN 2282-6483

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Political power, conflict and backlash: theory and evidence from Italy

Giovanni Righetto
Paolo Vanin

Quaderni - Working Paper DSE N° 1213

Political power, conflict and backlash: theory and evidence from Italy

Giovanni Righetto[1] Paolo Vanin [2]

September 22, 2025

Abstract

While democratization and enfranchisement are known to benefit minority groups in the long run, sudden increases in political representation can disrupt existing power balances, provoke resistance, and lead to worse policy outcomes in the short run. We document and explain this pattern. In our theoretical model, conflict and backlash are triggered by a sufficient increase in political power if preferences are sufficiently different. We exploit the introduction of an affirmative action measure in Italian local elections, which led to an exogenous increase in female political representation in small municipalities. Using a Difference in Discontinuity design, we document that, in line with the theory, moderate increases in female representation led to higher day care spending, while large increases resulted in lower spending on this gender-sensitive issue. Higher council dissolution rates and null effects on non-gender-related policy areas support the interpretation of the evidence suggested by the theory. Several robustness checks and evidence from Spanish data also support the internal and external validity of our findings.

JEL classification numbers: D71, H53, I38

Keywords: Conflict, committee, gender quotas, day care investments

[1] University of Bologna. Email: giovanni.righetto3@unibo.it

[2] University of Bologna. Email: paolo.vanin@unibo.it

We would like to thank Nicola Mastrorocco, Andrea Mattozzi, Manuel Bagues, Giorgio Gulino, Giacomo Ponzetto, Pamela Campa, Mara Squicciarini, and all the participants to the following conferences and seminars: BSE PhD Jamboree organized by the Universitat Pompeu Fabra, NSEF PhD and Postdoc Workshop organized by the University Federico II, Internal seminar at the University of Bologna, Brown Bag seminar at the University of Milan, Conference on "Institutions, knowledge diffusion and economic development" organized by CRENOS, Bomopav 2024 organized by the Universities of Bologna, Modena, Venice and Padua, Sites-GLO conference 2024.

Non-Technical Summary

Expanding political rights and representation for marginalized groups has historically brought long-term social and economic benefits. Examples include the extension of suffrage to the poor, to women, and to African Americans, as well as the more recent adoption of gender quotas in politics. Yet, these reforms often produce short-term tensions: when a previously underrepresented group suddenly gains influence, existing power balances may be disrupted, sometimes leading to conflict and even worse policy outcomes in the immediate aftermath.

This paper asks whether, and under what conditions, an increase in minority political power can backfire in the short run. To answer this question, we develop a simple theoretical model of conflict and backlash. In the model, when group preferences are sufficiently divergent, a non-monotonic pattern arises: as the minority strengthens, it first secures policy concessions, then encounters backlash and conflict, and only later returns to a trajectory of policy gains. When preferences are closer, however, no conflict emerges.

We test these predictions in the context of Italian local politics. A 2012 reform of municipal elections introduced gender quotas that substantially raised the share of female councillors in municipalities above 5,000 inhabitants. Using a *Difference-in-Discontinuity* design, we find that moderate increases in female representation led to higher spending on day care, a highly gender-sensitive policy area. By contrast, large increases in female representation reduced such spending, consistent with the emergence of conflict and backlash. Supporting evidence comes from higher rates of council dissolution in these municipalities and the absence of effects on non-gender-sensitive policies. We further show that these patterns replicate in Spanish local elections, confirming the external validity of our findings.

The contribution of the paper is twofold. First, it offers a parsimonious theoretical framework to understand when minority empowerment generates cooperation versus conflict. Second, it provides novel causal evidence on the short-run political and fiscal consequences of gender quotas, showing that greater representation does not always translate into immediate policy gains.

1 Introduction

During the last two centuries, democratization in many countries has empowered social groups that were previously excluded from political decision-making. Examples include the extension of voting rights to the poor, to women, and to African Americans, as well as the more recent introduction of gender quotas in politics across both developed and developing countries. In the long run, these enfranchisement processes have often benefited underrepresented groups, but along the way they have also generated episodes of conflict and temporary backlash.[3]

While these phenomena may have complex historical, political, economic and identity-related roots, in this paper we first propose a simple stilized model of conflict and backlash, in which a group has the power to set a policy and another group may peacefully accept the decision or engage in conflict. If preferences are sufficiently divergent, a distinctive non-monotonic pattern arises: as the second group grows stronger, it initially secures greater policy concessions, then encounters a backlash and conflict, and eventually returns to a trajectory of policy gains. By contrast, when preferences are closer, conflict never emerges: as the second group gains strength, it may achieve further concessions or no gains at all, but without any escalation.

We next present empirical evidence consistent with the model. We exploit a 2012 reform of local elections in Italy that, as documented by Baltrunaite et al. 2019, substantially increased the share of female councillors in a subsample of small municipalities. We take such share as a measure of women's political strength. The reform was an affirmative action measure that was only applied to municipalities with more than 5,000 inhabitants, allowing a difference in discontinuity strategy to assess its impact on policy outcomes. We focus mainly on a highly gender-sensitive policy, local expenditures for day care, for which there

[3] For example, consider democratization and coups and their effects on redistribution to the poor (Acemoglu and Robinson 2001); discrimination and employment opportunities for Black Americans (Aneja and Avenancio-León 2019); women's suffrage and public health spending in the United States (Miller 2008); the enfranchisement of the less educated and political responsiveness in Brazil (Fujiwara 2015); or, more recently, the adoption and repeal of diversity and inclusion policies in the United States.

is evidence of significant differences in preferences between men and women.

In line with the model, we document a non-monotonic policy impact. Relative to the control group, the reform led to a 22% increase in day care expenditure in municipalities with low shares of female councillors and to a nearly 24% reduction in those with high shares. We also show that political conflict increased in the latter municipalities, leading to more frequent municipal council dissolution. By contrast, we find no effect on expenditure categories that are not gender-sensitive, again consistent with the model's predictions when group preferences are similar or only moderately different.

We run a number of checks to ensure the internal validity of our main finding. We also support its external validity by studying the effects of a gender quota introduced in 2007 in Spanish local elections. The external validity is further supported by Baskaran and Hessami 2025, who provide similar evidence for a different context.[4]

Why do a few more women in a municipality council lead to higher expenditure in day care, whereas many more lead to a reduction or at best no increase? This may appear surprising and counterintuitive.[5] Our model provides a natural and parsimonious explanation: while an increase in representation initially leads to policy gains, above a certain level the majoritarian group prefers to push back and trigger conflict. The reason is that the requests of the minority become more and more costly to accomodate, and at some point the cost of policy concessions becomes higher for the majority than that of conflict. In the model, as the minority's chances of victory in case of conflict approach one, there is a further reversal and the majority eventually turns to (even higher) concessions. In our empirical investigation, we do not observe this last phase, coherently with the fact that observed increases in female

[4]They exploit a regression discontinuity design on mixed-gender races for last party-specific council seats at local council elections in Bavaria and document that an additional female councillor accelerates the expansion of public child care by 40%, but, crucially, only in councils with few women. They also document that when there are more female councillors, they speak more frequently in council meetings and child care is discussed more often. In terms of standard economic analysis, this 'change of the discourse' is effective if it is more than cheap talk, for instance if it is paired by the threat of conflict, as in our model.

[5]Personal characteristics of elected politicians do not matter for policy making in the median voter theory (Downs 1957), whereas they do in the citizen-candidate model (Besley and Coate 1997), but then one might expect that as female representation increases, policies move in the direction preferred by women. See also Besley and Case 1995.

representation do not make women's chances of winning close to one.

Our model is a variation of the ultimatum bargaining game, in which rejection of the proposed policy is not followed by a null payoff for both groups, but rather by conflict, that is, by a lottery that determines who ultimately sets the policy. This simple structure makes the proposer's expected cost of conflict linear in the responder's strength, and the costs of the concessions required to ensure peace concave, thus generating a possible backlash at intermediate strength levels. To the best of our knowledge, this insight is new.[6]

Our theory belongs to the long-standing approach that, since Nash 1953 and Schelling 1960, studies conflict as an inefficient outcome due to commitment problems.[7] Much attention has been devoted to situations in which the dominant group would like to avoid conflict but cannot, because its current resources are not enough to buy the other group out of conflict and it cannot commit credibly to transfer enough future resources to the other group.[8] We change the perspective: in our simple two-period model, if there is conflict it is because the dominant group actually prefers it to peace. The commitment problem arises from the fact that, when conflict emerges in equilibrium, the underrepresented group would like to avoid it, and would be happy to promise peace in exchange for moderate concessions that are at most as costly to the majority as conflict, but this promise would not be credible: if those concessions were proposed, the minority would have an incentive to try to achieve more favorable outcomes precisely through conflict.

In the context of municipality councils, our simple explanation is appealing because in

[6]See Güth and Kocher 2014 for a discussion of the many variations of the ultimatum game studied in the literature after Güth, Schmittberger, and Schwarze 1982.

[7]Two-sided commitment problems are studied in Crawford 1982, Ellingsen and Miettinen 2008 and Baliga and Sjöström 2020, among others. We focus on a one-sided commitment problem. The literature on power shifts (e.g., Fearon 1995, Fearon 1998, Powell 1999 and Powell 2006, among others) emphasize that the expectation of an unfavorable future changes may trigger a preventive conflict. Our investigation of the role of strength is related but different because it relies on comparative statics and not on future changes in the strategic environment. See Baliga and Sjöström 2024 for a discussion of conflict due to commitment problems, and Acemoglu and Wolitzky 2024 for conflict due to asymmetric information.

[8]See in the seminal work by Rubinstein 1982, in which the baseline bargaining game is indefinitely repeated, and many applied models based on that framework. Powell 2004 argues that the same logic essentially applies to revolutions in Acemoglu and Robinson 2000, coups in Acemoglu and Robinson 2001, and secessions in Fearon 2004, among others.

many political fights, the dominant group has both agenda power and the means to avoid conflict but is just unwilling to use them. Moreover, while in some cases the minority may be able to acquire commitment devices that temporarily limit its requests and grant a peaceful agreement, this may not happen because self-restrain is politically costly and some parts of the minority may not support it or adhere to it. At the same time, our stilized model can be applied more broadly to think about conflict situations beyond the specific application pursued in this paper.[9]

An important premise of our interpretation of the empirical evidence is that male and female councillors have sufficiently different preferences over day care expenditure. While we are not aware of direct measures of such preferences among municipality councillors in Italy, Marchese, Profeta, and Savio 2025 present clean and closely related evidence for mayors, revealing an almost 40% increase in day care expenditure due to having a female mayor rather than a male one.[10] Unless councillors are substantially and implausibly different from mayors, the large gender gap in day care preferences is likely to extend to them.[11] The reform itself may have further contributed to polarize along gender lines councillors' views on day care, opening the doors to gender-based confrontation on this but not on other issues.[12] It

[9]If strength reflects group shares, our model features conflict for high levels of preference polarization, which is a hump-shaped function of group shares (Esteban and Ray 1994). Relative to Esteban and Ray 2011, we emphasize the sequential nature of policy setting, but we simplify investment in conflict resources and restrict confrontation to being over public goods and between two homogeneous and cohesive groups. In the sequential theory of bargaining in legislatures developed by Baron and Ferejohn 1989, low impatience and the possibility of amendments make initial proposals not necessarily accepted. The cost of conflict and concessions play a similar role in our model.

[10]This finding is based on a regression discontinuity design on close mixed-gender elections in Italian municipalities below 5,000 inhabitants for the pre-Covid-19 pandemic period. By contrast, exploiting close mixed-gender races in the US, Ferreira and Gyourko 2014 find that the election of a female mayor did not affect the size and composition of public expenditures, but this is clearly a different context.

[11]One reason for it may be that Italian women, in particular mothers, tend to provide more informal care and household work than their partners, as documented by Barigozzi, Cremer, Monfardini, et al. 2019, among others. The pandemic affected this balance and reduced the gender gap in day care preferences (Marchese, Profeta, and Savio 2025), but our empirical exercise stops before it. Funk and Gathmann 2015 and Slegten and Heyndels 2020 present related evidence of gender differences in preferences, the former based on citizens' votes in Swiss referenda and the latter on stated preferences among Flemish local politicians, but neither of them addresses day care in a specific way.

[12]Preference polarization may be seen as acting in the background of our model. As we take identity as given, our approach is complementary to those that endogenize group identification. See Bonomi, Gennaioli, and Tabellini 2021 and Grossman and Helpman 2021 for theories of identity politics and conflict, and Genicot and Ray 2024 for a discussion of this literature. Increased confrontation on gender-sensitive issues is in line

is also the case that the majority has the means to avoid conflict through concessions: it could always choose to devote a larger budget share to day care, but it may be unwilling to pay the associated political cost.

We contribute to the empirical literature that exploits the exogenous imposition of gender quotas to understand how the sex ratio of politicians in charge affects policy choices. See Hessami and Fonseca 2020 for a review. Here the evidence is not univocal and ranges from substantial effects in India to no effects in Spain or Norway.[13] For Italy, Baltrunaite et al. 2019 find no effects on local expenditures and Andreoli, Manzoni, and Margotti 2022 find small changes in public security spending and administration costs.[14] For day care expenditure, we confirm that the reform had no effect on average but we show that this is due to a composition of positive and negative effects that depend on the share of female councillors in the way predicted by our model. For expenditure categories that are not gender-sensitive, we find no effect on average and no heterogeneity based on the sex ratio of councillors, again as predicted by the model.

The remainder of the paper is organized as follows. Section 2 presents the theoretical model, Section 3 the empirical analysis, Section 4 discusses internal and external validity, and Section 5 concludes.

with experimental evidence (Coffman 2014) showing that women are more likely to expose their personal ideas in environments that are not male-dominated.

[13]Since 1993, the Indian mandate reservation system has increased female political representation by prescribing that a minimum of one-third of seats plus the leadership position in randomly selected villages must be reserved for women. Its consequences include a better representation of the policy preferences of the female electorate (Chattopadhyay and Duflo 2004), with higher expenditure on public goods, especially education (Clots-Figueras 2012) and public health infrastructures (Bhalotra and Clots-Figueras 2014), a stronger reduction in the gender gap in school attendance and educational attainment (Beaman et al. 2012), and a lower infant mortality due to improved prenatal and day care services significantly reduced (Bhalotra and Clots-Figueras 2014). Bagues and Campa 2021 find that the introduction of a gender quota in Spanish local elections had no effects on either the composition or size of public expenditures. Geys and Sørensen 2019 find similar evidence for Norway: the exogenous increase of female politicians due to a gender quota did not alter how local administrators use public funds.

[14]Compared to this last paper, our work differs both in the identification strategy, a Diff-in-Disc vs. an Instrumental Variable approach, and in the main outcome variable, retrieved from the Istat website and not from municipalities' balance sheets. Relatedly, exploiting close mixed gender elections, Casarico, Lattanzio, and Profeta 2022 find no significant effects of female mayors in Italy for broad expenditure categories.

2 A model of conflict and backlash

We model a sequential game in which two groups, denoted M and F, and modeled as single agents, need to decide the level of a policy variable $q \in \mathbb{R}$, over which they have different preferences, with respective bliss points b_M and b_F. Their decision may be peaceful or conflictual. First, M proposes a value $p \in \mathbb{R}$ to F. Next, F chooses whether to engage in a conflict or not, denoted as $s(p) \in \{C, N\}$, respectively. If it chooses conflict, it has a probability $\pi \in (0, 1)$ of winning, in which case it determines the outcome, so that $q = b_F$. In case of peaceful agreement, or if M wins the conflict, M's proposal is implemented, so $q = p$.[15] One can think of M as a dominant or majority group and F as an opposition or minority group. M can try to impose its will and implement its preferred policy, but F can obtain attention to its requests and some policy concessions through the threat of conflict.[16]

Engaging in a conflict has a direct cost $\gamma > 0$ for F, but it is also costly for M, as it leads to an undesired outcome with positive probability.[17] To avoid such costs, M may propose a compromise between the two bliss points, hoping to convince F to peacefully accept it. We investigate the trade-off between the cost of conflict and that of compromise.

Let preferences over q be defined by a quadratic loss function in the distance between q and each agent's bliss point, $\frac{1}{2}(q - b_i)^2$, for $i \in \{M, F\}$.[18] Without loss of generality, we let $b_F > b_M$ and normalize $b_M = 0$, so that b_F, which we rename b to simplify notation, is both F's bliss point and the distance between M and F's preferences.

Given q and $s \in \{C, N\}$, payoffs are therefore $U_M = -\frac{1}{2}q^2$ and $U_F = -\frac{1}{2}(q-b)^2 - \gamma\, \mathbb{1}\{s = C\}$. A strategy profile $(p, s(\cdot))$ specifies a proposal $p \in \mathbb{R}$ and a conflictual or non-conflictual response to any possible proposal, and generates a probability distribution over outcomes,

[15] As shown below, M proposes its bliss point b_M in equilibrium if it is willing to face conflict.

[16] If strength reflects population shares and $\pi \in \left(0, \frac{1}{2}\right)$, the majority has the right to make a proposal and also a higher probability of winning in the event of conflict.

[17] Adding direct costs of conflict for M would not generate additional insights. Hence, for the sake of simplicity, they are not included.

[18] The analysis can be easily extended to a broad class of single-peaked preferences. The quadratic form is chosen for expositional convenience.

with associated expected payoffs

$$
\begin{aligned}
U_M(p, s(\cdot)) &= -\frac{1}{2}p^2\, \mathbb{1}\{s(p) = N\} - \frac{1}{2}[(1-\pi)p^2 + \pi b^2]\, \mathbb{1}\{s(p) = C\} \\
U_F(p, s(\cdot)) &= -\frac{1}{2}(p-b)^2\, \mathbb{1}\{s(p) = N\} - \left[\frac{1}{2}(1-\pi)(p-b)^2 + \gamma\right] \mathbb{1}\{s(p) = C\}
\end{aligned}
\tag{1}
$$

We study a subgame perfect equilibrium of this game. By backward induction, let's start with F's best response $s^*(p)$. In case of loss, conflict yields the same payoff as acceptance, but in case of victory, which happens with probability π, it allows avoiding the loss associated with p. Hence, F's expected gain from conflict is $\frac{\pi}{2}(q-b)^2$ and F chooses C whenever such gain is higher than the direct cost of conflict γ.[19] The minimum accepted proposal,

$$
a \equiv b - \sqrt{\frac{2\gamma}{\pi}}, \tag{2}
$$

is thus increasing in b, decreasing in γ and, importantly for the subsequent analysis, it is an increasing and concave function of π.[20] Given this, M's best response p^* is to propose the implementation of either its bliss point or the minimum accepted proposal, so $p^* \in \{0, a\}$.

To see why, observe that if preferences are sufficiently close, M's bliss point is peacefully accepted and it is therefore the optimal proposal. This happens if $a \leq 0$, which amounts to $b \leq \sqrt{\frac{2\gamma}{\pi}}$. If b is larger than this threshold, M's bliss point would be rejected, but it remains the best option among the proposals that induce conflict, just as a is the best option among those that are peacefully accepted. M then trades off the cost of conflict, due to the chance of losing, and the cost of peace, due to the required concessions. The former is given by $\frac{\pi}{2}b^2$ and the latter by $\frac{1}{2}a^2$.

When $b = \sqrt{\frac{2\gamma}{\pi}}$ the cost of peace is zero and that of conflict is $\gamma > 0$. Both increase smoothly in b afterwards, but the cost of peace increases more steeply, because that of conflict

[19] Formally, $U_F(p, s(\cdot)|s(p) = N) < U_F(p, s(\cdot)|s(p) = C) \iff \frac{\pi}{2}(p-b)^2 > \gamma$. Hence, $s^*(p) = C$ if either $p < b - \sqrt{\frac{2\gamma}{\pi}}$ or $p > b + \sqrt{\frac{2\gamma}{\pi}}$, and $s^*(p) = N$ otherwise.

[20] These properties are not specific of quadratic preferences: they generalize to any loss function that is strictly increasing and either linear or convex in the distance from the bliss point.

is 'toned down' or 'discounted' by the fact that the undesired outcome b only materializes if M loses. As a consequence, as long as the distance in preferences is in an intermediate range, $b \in \left(\sqrt{\frac{2\gamma}{\pi}}, \frac{\sqrt{2\gamma}}{\sqrt{\pi}-\pi}\right]$, triggering conflict is more costly to M than compromising, because small concessions, whose cost is low, are enough to keep peace, and so M proposes a.[21] Yet, since the cost of peace increases faster in b than that of conflict, the concessions required to preserve peace eventually become too costly, so that if the distance in preferences is too high, $b > \frac{\sqrt{2\gamma}}{\sqrt{\pi}-\pi}$, M proposes its bliss point and triggers conflict.[22]

This characterizes equilibrium strategies.[23] Along the equilibrium path of play, for $b \leq \sqrt{\frac{2\gamma}{\pi}}$, one has peaceful agreement on M's bliss point ($q = 0$ and no coflict); for $b \in \left(\sqrt{\frac{2\gamma}{\pi}}, \frac{\sqrt{2\gamma}}{\sqrt{\pi}-\pi}\right]$, agreement is obtained through costly concessions ($q = a$ and no conflict); and for $b > \frac{\sqrt{2\gamma}}{\sqrt{\pi}-\pi}$, there is conflict, with each agent fighting to obtain its bliss point, so that $E(q) = \pi b$. Since $\pi \in (0, 1)$, the latter inequality implies

Proposition 1. *Conflict is possible along the equilibrium path of play if and only if the distance in preferences is sufficiently large relative to F's cost of conflict, $b > 4\sqrt{2\gamma}$.*

If this condition is not satisfied, $b \leq 4\sqrt{2\gamma}$, there is always agreement along the equilibrium path of play, possibly obtained through some concessions, and conflict never materializes, as detailed in

Proposition 2. *M's bliss point is always peacefully accepted if the distance in preferences is very small, $b < \sqrt{2\gamma}$. If it lies in an intermediate range, $b \in (\sqrt{2\gamma}, 4\sqrt{2\gamma}]$, peaceful agreement*

[21] Provided $b > \sqrt{\frac{2\gamma}{\pi}}$, one has $U_M(a, s^*(\cdot)) > U_M(0, s^*(\cdot)) \iff b < \frac{\sqrt{2\gamma}}{\sqrt{\pi}-\pi}$.

[22] Summing up, M's best response to $s^*(\cdot)$ is $p^* = 0$ if $b < \sqrt{\frac{2\gamma}{\pi}}$ or $b > \frac{\sqrt{2\gamma}}{\sqrt{\pi}-\pi}$ and $p^* = a$ otherwise.

[23] In a compact way, they are

$$p^* = \begin{cases} 0, & \text{if } b < \sqrt{\frac{2\gamma}{\pi}} \ \vee \ b > \frac{\sqrt{2\gamma}}{\sqrt{\pi}-\pi} \\ a, & \text{if } b \in \left[\sqrt{\frac{2\gamma}{\pi}}, \frac{\sqrt{2\gamma}}{\sqrt{\pi}-\pi}\right] \end{cases}$$

$$s^*(p) = \begin{cases} C, & \text{if } p < b - \sqrt{\frac{2\gamma}{\pi}} \ \vee \ p > b + \sqrt{\frac{2\gamma}{\pi}} \\ N, & \text{if } p \in \left[b - \sqrt{\frac{2\gamma}{\pi}}, b + \sqrt{\frac{2\gamma}{\pi}}\right] \end{cases}$$

is reached on M's bliss point if $\pi \leq \frac{2\gamma}{b^2}$*, so* $q = 0$*, and on concessions if* $\pi > \frac{2\gamma}{b^2}$*, so* $q = a > 0$*.*

The most interesting case is that in which the distance in preferences is large enough to make a conflict possible. If $b > 4\sqrt{2\gamma}$, equilibrium outcomes as a function of F's strength π are presented in

Proposition 3. *If conflict is possible, there exist thresholds* π_0*,* π_1 *and* π_2*, satisfying* $0 < \pi_0 < \pi_1 < \pi_2 < 1$*, such that, along the equilibrium path of play, there is*

- *Peaceful agreement on M's bliss point, so that* $q = 0$*, if* $\pi \leq \pi_0$*;*
- *Peaceful agreement on concessions, so that* $q = a$*, if* $\pi \in (\pi_0, \pi_1]$*;*
- *Conflict, so that* $E(q) = \pi b$*, if* $\pi \in (\pi_1, \pi_2)$*;*
- *Peaceful agreement on concessions, so that* $q = a$*, if* $\pi \geq \pi_2$*.*

The proof of these propositions immediately follows from the previous analysis.[24] The proposition highlights an interesting non-linearity. As F's power increases, in the sense of having higher chances of victory in case of conflict, M is first willing to accommodate by making concessions that are an increasing and concave function of π. However, from M's perspective, the expected utility loss in case of conflict increases linearly in π, whereas the cost of concessions increases non-linearly, first less, then more and then again less steeply than the cost of conflict, creating an intermediate range of π in which M finds it too costly to accommodate F's requests, although it eventually reverts to concessions.[25] Notice that,

[24]One has $b \leq \sqrt{\frac{2\gamma}{\pi}} \iff \pi \leq \pi_0$, where $\pi_0 = \frac{2\gamma}{b^2} > 0$ and $\pi_0 < 1 \iff b > \sqrt{2\gamma}$. Moreover, $b > \frac{\sqrt{2\gamma}}{\sqrt{\pi} - \pi} \iff \pi \in (\pi_1, \pi_2)$, where $\pi_1 = \frac{1}{2} - \frac{\sqrt{2\gamma}}{b} - \frac{1}{2}\sqrt{1 - \frac{4\sqrt{2\gamma}}{b}}$ and $\pi_2 = \frac{1}{2} - \frac{\sqrt{2\gamma}}{b} + \frac{1}{2}\sqrt{1 - \frac{4\sqrt{2\gamma}}{b}}$. Hence, conflict arises for π within this range. Inequalities are easy to establish. Interestingly, given $b > 4\sqrt{2\gamma}$, the threshold determining the end of conflict satisfies $\pi_2 < \frac{1}{2} \iff b \in (4\sqrt{2\gamma}, 2(2+\sqrt{2})\sqrt{\gamma})$, a non-empty interval.

[25]Given $b > 4\sqrt{2\gamma}$, M's cost of peace $\frac{a^2}{2}$ is convex in π for $\pi < \frac{32\gamma}{9b^2}$ and concave above this threshold, which satisfies $\pi_0 < \frac{32\gamma}{9b^2} < \pi_2$; and it is steeper than that of conflict at π_1 and less steep at π_2.

conceptually, if π reflects F's population share, then the Esteban-Ray measure of polarization is hump-shaped in π, and our model features conflict at high levels of polarization.

Figure 1 plots the expected value of the outcome variable, $E(q)$, as a function of π, along the equilibrium path of play.[26] As its strength increases, by threatening conflict F is mostly able to obtain more favorable results, but the onset of conflict determines a temporary backlash. The reason is that while concessions, if proposed, increase smoothly in π, at π_1 M stops proposing them and instead chooses to push for its bliss point and face conflict. This causes a discontinuous drop in the expected value of q and hence in F's utility (which also includes the direct cost of conflict γ).[27]

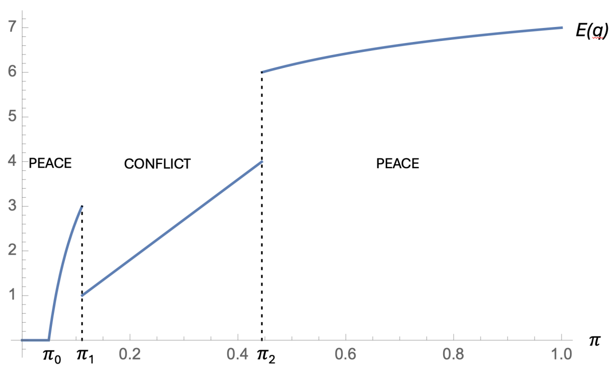


Figure 1: Expected equilibrium level of q as a function of π

This backlash may appear puzzling. When F's strength passes threshold π_1, the minimum proposal that it is willing to accept becomes more costly to M than conflict. Since F is discretely worse off under conflict, why does it not moderate its requests, promising that it will accept less generous concessions that leave M indifferent between compromise and conflict? This solution seems reasonable: it avoids a socially wasteful conflict and it is a Pareto improvement over the equilibrium. The problem is that such a promise by F would not be credible and thus, even if made, would not be believed.

[26] It is drawn for $\gamma = 2$ and $b = 9$, so that $\pi_2 < \frac{1}{2}$, but the qualitative pattern holds for any parameter combination such that conflict is possible. $E(q)$ is equal to 0 for $\pi \in (0, \pi_0]$, to a for $\pi \in (\pi_0, \pi_1]$, to πb for $\pi \in (\pi_1, \pi_2)$, and again to a for $\pi \in [\pi_2, 1)$.

[27] M's utility in turn decreases continuously in π at π_1. The discontinuous drop in $E(q)$ is the only result that depends on quadratic preferences. With linear preferences, Propositions 1 and 2 would be virtually identical (but for the specification of the thresholds and of a), and F's utility would still drop discontinuously at π_1, but $E(q)$ would be continuous.

Ultimately, therefore, conflict and the associated backlash (for intermediate values of F's strength) depend on F's lack of commitment, on the technology of conflict (which makes the costs of conflict and peace respectively linear and concave as functions of F's strength), and on a sufficiently large disagreement over desired outcomes (relative to the direct cost of conflict). Other assumptions, like perfect information, can be easily relaxed.[28]

In terms of implications, the model shows that if preferences are close enough, conflict never arises. In that case, an increase in F's strength is associated with no or monotone improvements in F's outcomes and welfare. If preferences are instead sufficiently distant to make conflict possible, a progressive increase in F's strength generates a temporary backlash precisely after a period of fast improvement, but eventually the conflict phase ends and the improvement takes back its old path. Following the interpretation of F as a minority group, we take these predictions to the data leveraging the effect of an Italian law that introduced gender quotas in some municipal councils.

3 Empirical evidence

To bring the model to the data, we leverage a reform that generated an exogenous change in female representation in some municipal councils in Italy, allowing a difference in discontinuity approach. The reform aimed precisely at increasing women's presence among Italian local politicians, in order to reduce the country's gender gap in politics.[29]

[28]If M were uncertain about F's bliss point, it would base its choices on its expected value. The analysis would be similar, with two major differences: first, there would be an additional source of conflict, namely inadequate proposals due to expectations that turn out to be incorrect ex post; second, a risk-averse M would make more generous concessions to hedge against the risk of undesired conflict.

[29]Women are under-represented in politics for a large variety of reasons, such as voter biases against women (Cella and Manzoni 2023), lack of self-confidence (Fox and Lawless 2011), gaps in political rewards (Júlio and Tavares 2017), party leaders' bias (Hessami and Fonseca 2020), no visibility within parties (Kjaer and Krook 2019), gender differences in voters' persuasion (Savio 2024). This gap in representation has called for the implementation of affirmative action measures, such as gender quotas.

3.1 The Italian electoral reform of 2012

The reform, contained in Law 215/2012, changed the rules for local elections in municipalities with more than 5,000 inhabitants. Municipal councils in Italy are made up of 12 councillors and decide by majority the allocation of public funds (resulting from their own taxes and tariffs, transfers from the central government, and revenues from fines) to a variety of expenditure categories.[30]

Law 215/2012 introduced two major changes to local elections in municipalities with more than 5,000 inhabitants. First, it imposed a gender quota, stipulating that no more than two-thirds of the candidates on a municipal council electoral list could be of the same sex. Non-complying parties were punished with the removal of same-sex candidates exceeding two-thirds of the total in their lists. Second, it allowed citizens to cast two preference votes for councillors instead of one, provided the votes were for candidates of different sexes.

By 2018, the share of female councillors in affected municipalities had roughly doubled. Through a regression discontinuity design, Baltrunaite et al. 2019 show that the law increased the percentage of female councillors by an average of 18 percentage points, without relevant effects on elected candidates' characteristics in terms of age, years of education, or previous occupation.[31]

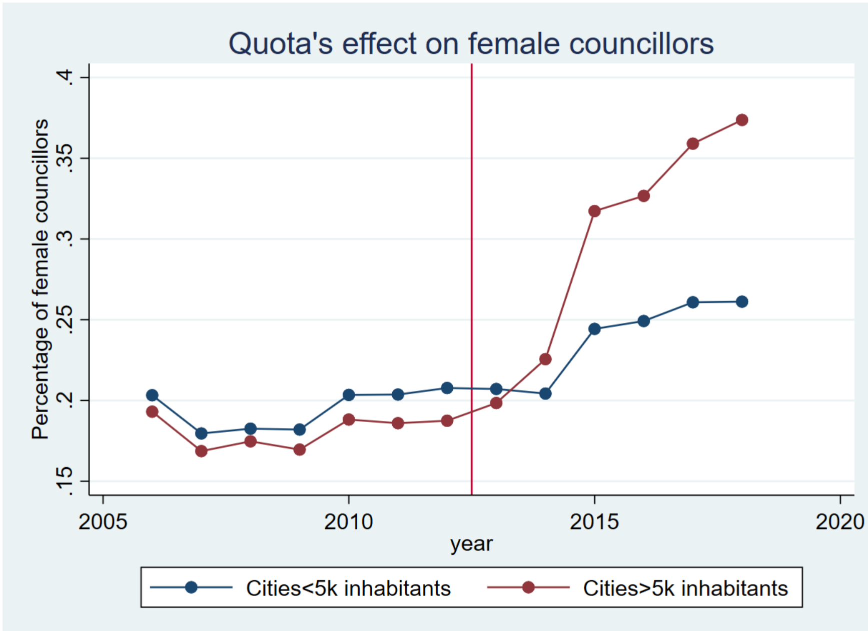
We focus on municipalities between 2,000 and 10,000 inhabitants between 2006 and 2018, which constitute a relatively homogeneous sample in terms of characteristics and electoral rules. Figure 2 plots the evolution of the share of female councillors in municipalities above

[30] In Italy, there are approximately 8100 municipalities, which represent the lowest sub-national administrative level, after regions and provinces. Each municipality has a mayor assisted by a local council ("Consiglio comunale") with legislative power and an executive committee ("Giunta comunale") with executive power. Elections are held every five years, and voters can express their preference for both the mayor and local councillors. In municipalities between 2,000 and 10,000 inhabitants, citizens over 18 can cast their vote for the mayor and, until 2013, for one local councillor, since the electoral system prescribes semi-open lists. Each mayoral candidate can be backed by one list of council candidates, and the mayoral candidate receiving the relative majority of votes obtains two-thirds of the seats that are allocated to his or her councillors. The remaining third of seats are allocated to other mayoral candidates via a proportional system. Seats are then attributed to councillors according to their vote ranking, which is relative to each party.

[31] They argue that the law's effects were mostly due to the double vote for councillors, whereas its impact on turnout, even by gender, was negligible.

and below 5,000 inhabitants.

Figure 2: **Share of Female councillors**



Notes: Percentage of female councillors over the years in Italian municipalities with a population between 2,0000 and 10,000 inhabitants. The red vertical line indicates the time when Law 215/2012 was introduced. Elections are staggered across municipalities. Data include 1,422 out of 8,100 Italian municipalities, observed between 2006 and 2018.

Municipalities affected by Law 215/2012 display a sharp increase in the share of female councillors, especially after 2013. On average, the percentage of female councillors passed from 18% in the pre-law period to almost 38% in 2018 for affected municipalities.[32] Despite this, even after the imposition of the new law, in almost all municipalities (93% of our sample) men retained the majority, which is required in order to pass motions on public expenditures: in other words, political power mostly remained in the hands of male councillors. The share of female councillors also increased in municipalities not affected by the law, which constitute our control group. This may be a spillover effect of the reform, in which case our estimates, presented below, may be interpreted as a lower bound of the true impact.[33]

[32]It is important to point out that elections are staggered across municipalities, thus different municipalities may have different election years. The big spike in the increase in female councillors that we observe in 2014 is due to the fact that most municipalities subject to the law held elections in that year, thus in 2013 their councillors were still the ones chosen with the old set of rules.

[33]Part of the increase may also be related to Law 56/2014, which mandated all municipalities with more

3.2 Data

We collected data on a total of 1424 municipalities with populations between 2000 and 10000 inhabitants, focusing on the period from 2013 to 2018. Given that Law 215/2012 is binding for municipalities with a population above 5000 inhabitants, we consider those municipalities as belonging to our treatment group, while the remaining ones constitute our control group. In Italy, local elections can be held at different moments for different municipalities: each year of observations we sampled is an electoral year for a group of municipalities.[34] Even though we collected data for the years before 2013, we decided to exclude the pre-2013 period since municipalities above the 5000 inhabitants threshold had different fiscal rules with respect to those below the threshold: they were indeed subject to the "Domestic Stability Pact", which in 2013 was extended to all municipalities.[35] Since we focus on public expenditures, our sampled municipalities need to have a homogeneous set of fiscal rules and this condition is satisfied for the period after 2013. To conduct our analysis, we merged two datasets. The first one, collected from the website of the Italian Ministry of Interior, concerns characteristics of elected councillors, mayors and aldermen, namely age, sex, party and previous job. Since the units of observation of our analysis are municipalities and not single politicians, for each municipality and year we created several variables indicating the characteristics of the councillors in charge, together with the percentage of female councillors. Together with the dataset on local politicians, we collected data on municipal expenditures, both from

than 3,000 inhabitants to have at least 40% of aldermen from each gender. This may have influenced the proportion of female councillors below 5,000 inhabitants since aldermen must be chosen among councillors. We argue that this additional quota should not bias our results, since day care expenditures are approved by the local councils, while aldermen are just in charge of implementing decisions. In addition, the year-fixed effects from our specifications should capture any potential impact from Law 56/2014 on day care, since the law is binding for both treated and control municipalities. However, we checked whether this other law influenced day care expenditures and we found no effect on the level of expenditures for municipalities above 3,000 inhabitants.

[34]To be more precise, out of the 1424 sampled municipalities, 114 held elections in 2013, 808 in 2014, 104 in 2015, 226 in 2016, 155 in 2017, and 134 in 2018.

[35]The Domestic Stability Pact introduced some restrictions on municipalities' expenditures: the overall budget balance had to be proportional to a moving average of the balances of the previous years. Lack of adherence implied sanctions for municipalities, while compliance granted a reduction in interest expenses for government loans.

Istat (Italian National Institute of Statistics) and from municipal balance sheets.[36] Since a city's local council has discretion over more than 100 spending categories, we decided to focus on the most gender-sensitive one, for which we expect to see an effect of the law: day care expenditures. Given the significant rise in female councillors, we may expect a corresponding increase in public funds allocated to day care. For this reason, we collected data on municipal day care expenditures, as well as the number of available spots and users by municipality and year.[37] The main dependent variable of our analysis, day care expenditures per capita, results from three expenditure categories added up, with the total divided by the municipality's population. These categories include municipal expenditures for directly managed day care facilities, privately managed day care facilities, and day care-related expenses.[38] All these expenditures are decided by local administrators and have to be approved by the local council.[39] Here, when we speak of day care facilities, we refer to nurseries for early childhood assistance, which in Italy are meant for children aged 0 to 3. In Table 1 we show summary statistics for the share of female councillors and day care expenditures, for the whole sample and then for treated and control municipalities.

[36] The "Certificati Consultivi" from the Ministry of Interior website.

[37] Data on municipal day care expenditures were obtained from Istat, which provides additional microdata on local government finances.

[38] Local councils can decide to directly provide this service or outsource it to privates. Day care-related expenses include all subsidies to families for day care integrated services or other related expenses, for instance the "babysitter bonus".

[39] There might be occasional state-level contributions to day care. For instance, national law 232 of the 11 December 2016 granted a 1000 euros voucher for kindergartens to all families. We argue that this effect should not bias our results, since it is not different for treated and control municipalities, and any potential bias should be absorbed by year-fixed effects.

Table 1: **Day care expenditures and shares of female councillors**

Variable	Obs	Mean	Std. Dev.	Min	Max
A. Whole sample					
Day care exp. (per capita)	8027	7.194	12.519	0	112.956
Mean share of female councillors	8027	0.258	0.135	0	0.692
B. Treatment group					
Day care exp. (per capita)	3213	8.983	13.999	0	83.943
Mean share of female councillors	3213	0.285	0.133	0	0.500
C. Control group					
Day care exp. (per capita)	4814	5999	11.269	0	112.956
Mean share of female councillors	4814	0.239	0.133	0	0.692

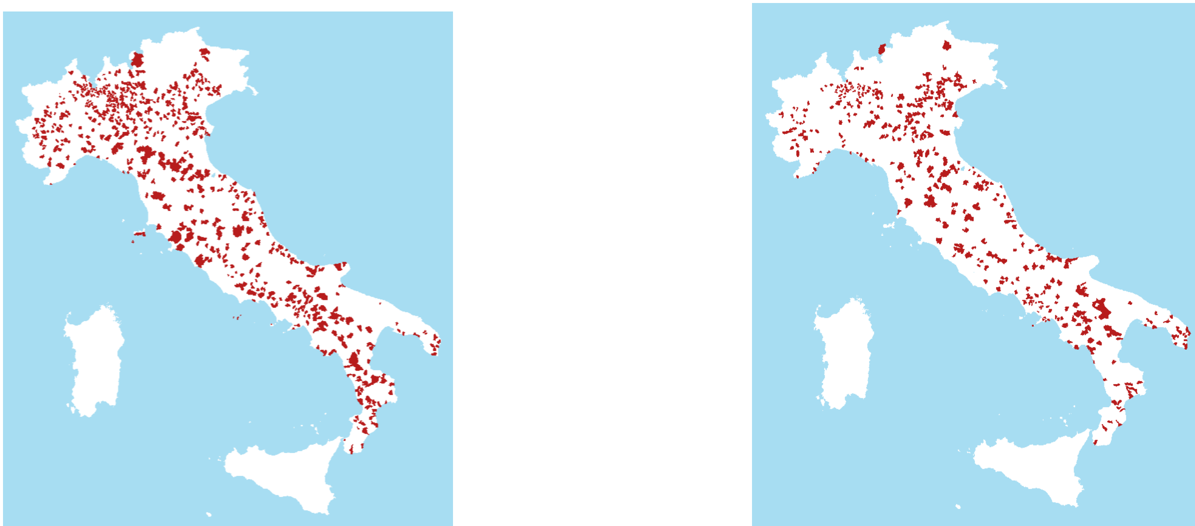
Notes: Day care expenditure per capita (in euros) sums three categories: expenditures for directly managed day care facilities, funding to privately managed day care facilities, and municipality funding for day care-related expenses. These expenses amount to roughly 0.6–1% of total annual municipal expenditures in the sample. Council characteristics include the mean share of female councillors. The treatment group includes municipalities with population above 5,000; the control group includes those below. The panel comprises 1,422 municipalities (out of about 8,100 in Italy) observed from 2013 to 2018.

Treated municipalities have a higher share of female councillors, also due to Law 215/2012, and spend more on average on day care.[40] The vast majority of the sampled politicians (95%) belong to so-called "liste civiche" (civic lists), which are local political groups that generally do not align with traditional left-right ideologies. As such, these politicians are not bound by the directives of national or regional parties. Moreover, voting in municipal councils is secret. This independence grants local councillors greater freedom to vote according to their personal preferences and local priorities, compared to their counterparts in higher-level institutions such as the national Parliament or regional councils.

In Figure 3 we present the geographical distribution of control and treated municipalities: these two maps give us confidence that there is no geographical bias in either our sampling or assignment to treatment procedure.

[40] In Table 1 we have both pre and post-treatment periods, therefore the difference in the share of female councillors is not the 18 percentage points that we cited before.

Figure 3: **Control and treated municipalities**



(a) Control municipalities

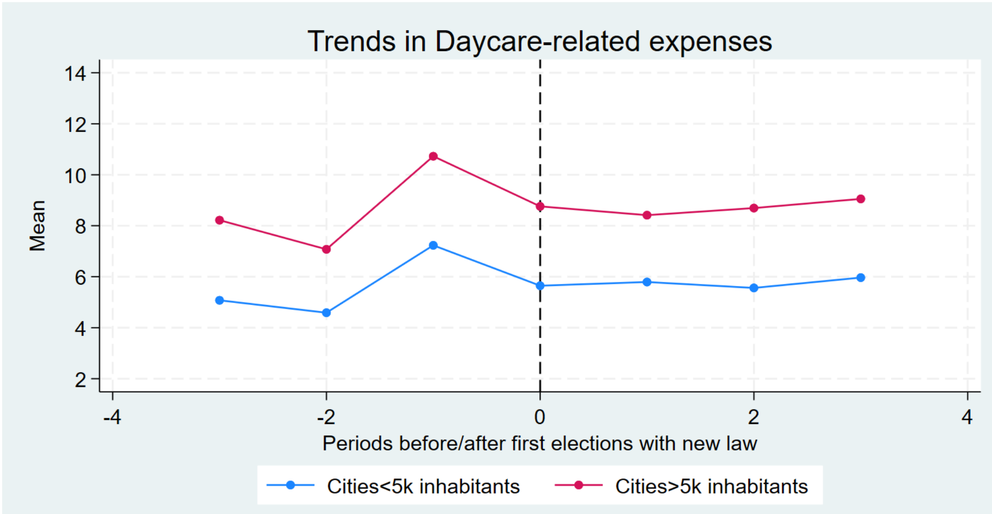
(b) Treated municipalities

Notes: Figure 4a shows all municipalities belonging to our sample with population below 5000 inhabitants, while Figure 4b shows the subsample of treated municipalities, with population higher than 5000 inhabitants. Treated municipalities were subject to Law 215/2012 since their first election after 2012. Sampled municipalities are 1424 out of 8100 Italian municipalities. Special regions' municipalities (e.g. those in the main Islands) were excluded from the sample since they are subject to different fiscal rules with respect to ordinary regions ones.

We focus on day care spending because, as discussed in Section 1, the literature documents a stronger preference for this expenditure category among female politicians compared to male politicians. In Figure 4 we can see the evolution of day care expenditures for both treated and control municipalities over the years before and after the first elections with the new law. As we can see, trends are extremely similar, even if treated municipalities exhibit higher levels of expenditures per capita on average. There is a spike in expenditure in the pre-electoral year, possibly due to reelection incentives: however, this should not create concerns in our analysis since it is almost identical for treated and control municipalities.[41]

[41] We formally test whether the dynamic difference in expenditures is significant in Appendix C, where we run a dynamic difference in difference. There are no differential either pre or post-law trends in day care expenditures between treated and control municipalities. In addition, controlling for pre-electoral year does not change our main results.

Figure 4: **Trends in day care expenditures for treated and control muncipalities**



Notes: Trends in day care-related expenditure per capita at the municipality level. The vertical line indicates the year when a municipality held its first elections under the law 215/2012, while the x-axis indicates the years before and after this time. The variable, expressed in euros, includes three expenditure categories added up: expenditures for directly managed day care facilities, funding to privately managed day care facilities and municipality's funding for day care-related expenses. These kinds of expenses represent a fraction between 0.6 and 1% of total expenses taken yearly by our sampled municipalities. The treatment group includes sampled cities with populations above 5000 inhabitants, while the control group includes those below. Sampled municipalities are 1424 out of 8100 Italian municipalities.

In the following section, we explain how our identification strategy allows assessing the effect of the introduction of Law 215/2012 on day care-related expenses.

3.3 Empirical strategy

To apply our theory to municipal councils, we use the share of female councillors as a proxy of their political strength. The theory predicts that if the preference distance between male and female politicians over an issue is small, there is no conflict along gender lines (Proposition 1). In this case, policy outcomes can improve but cannot be worse for women after an increase in their representation (Proposition 2). Most interestingly, if preferences are sufficiently distant, policy outcomes are predicted to improve for women after small increases in their representation but may worsen after large increases (Proposition 3).

To test these predictions, and in particular the last one, we make use of a Difference in Discontinuity identification strategy (Grembi, Nannicini, and Troiano 2016), which combines a Regression Discontinuity design with a Difference in Difference analysis. We focus on the change in day care expenditures per capita at the threshold of 5000 inhabitants after the first elections with the new legislation on councillors' lists and the alternate vote. Thus, our treatment is represented by having a council elected after 2012 in a municipality with a population larger than 5000 inhabitants. We evaluate the effect of this treatment on our outcome of interest, which is a municipality's day care expenditures per capita.

Our model takes the following functional form:

$$Daycarepc_{it} = \alpha + \beta * Treatment_i + \gamma * PostLaw_{it} + \delta * TreatPost_{it} + normPop_i * (\eta * Treatment_i + \theta * PostLaw_{it} + \iota * TreatPost_{it}) + Year_t + City_i + \Lambda * Mayor_{it} + \epsilon_{it}$$

where $Daycarepc$ measures day care-related expenses per capita in municipality i at time t; $Treatment$ is dummy for a municipality with more than 5000 inhabitants; $PostLaw$ is dummy for councils elected after Law 215/2012 entered into force; $TreatPost$ is the interaction between $Treatment$ and $PostLaw$; $normPop$ is normalized population, equal to population-5000; $Year$ and $City$ capture year and municipality fixed effects, respectively; and $Mayor$ is a set of controls for the current mayor (age, sex and level of education).

The treatment effect is identified by coefficient δ, which captures the effect of the introduction of the law on the treated municipalities around the threshold of 5000 inhabitants. We consider only municipalities belonging to a bandwidth $[normPop-h;normPop+h]$, with h computed using the methodology of Calonico, Cattaneo, and Farrell 2020, the standard MSE-optimal bandwidth.[42] Following Gelman and Imbens 2019, we perform only a linear fit and we don't use any higher-order polynomial to identify our treatment effect at the threshold.

A Difference in Discontinuity strategy allows to disentangle the law effect from another confounder that we have at the same 5000 inhabitants' thresholds: the mayor's wage. While mayors of municipalities under 5,000 inhabitants earn €2,170 per month, in cities above

[42] The value of h computed through its methodology is 716, meaning that the municipalities considered in our main regressions have a population belonging to the interval [4284;5716]. However, our results are robust to the adoption of another narrower bandwidth minimizing the coverage error rate.

this cutoff (up to 10000 inhabitants) mayors' monthly wage increases to 2790 euros. This confounder might have effects on both selection into politics and the kind of policies adopted (Gagliarducci and Nannicini 2013), possibly biasing our results in case we implemented a simple Regression Discontinuity design. On the other hand, a dynamic strategy such as the Difference in Discontinuity can disentangle the law effect from the wage effect. Moreover, since this strategy focuses on identifying the treatment effect at the threshold (thus, considering only municipalities belonging to a bandwidth), it guarantees a more convincing comparison between treated and control units with respect to a simple Difference in Difference.[43] In Appendices A and B, we discuss the validity assumptions that must be satisfied for this identification strategy to be implemented (Grembi, Nannicini, and Troiano 2016).

3.4 Aggregate results on day care expenditure

We first present aggregate results that serve as a benchmark. We then explore heterogeneous treatment effects with respect to our proxy for strength, namely the post-law share of female councillors.

Table 2 presents the aggregate results from the estimation of the Difference in Discontinuity model described above, with and without mayor controls. To preserve space, we present only the coefficient of the treatment effect, *TreatPost*.[44]

[43] Our main Difference-in-Discontinuity results are robust in coefficient signs and magnitudes to the use of a less local identification strategy such as the Difference-in-Difference. We did not choose a Difference-in-Difference design as our main strategy, since the more we move away from the threshold of 5000 inhabitants, the higher the chances that we are comparing cities with different institutional rules: for instance, at 3000 inhabitants there are other changes in terms of electoral laws and local government composition.

[44] The exclusion of Municipality and Year fixed effects does not impact the magnitude and significance levels of our coefficients throughout all the empirical results in the paper.

Table 2: **Difference in Discontinuity aggregate results**

VARIABLES	(1) Daycarepc	(2) Daycarepc
TreatPost	-0.576	-0.393
	(0.997)	(0.990)
Mayor's controls	NO	YES
Municipality FE	YES	YES
Year FE	YES	YES
Observations	2,548	2,548
R-squared	0.027	0.030
Number of municipalities	570	570
Pre-Law Daycarepc Mean	6.06	6.06

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Difference in discontinuity regressions. TreatPost corresponds to the treatment coefficient, the interaction between the dummy identifying the treatment group and the dummy for the post-treatment period. The dependent variable is yearly per capita expenditures on day care-related categories. Expenditures taken during election years are attributed to the council governing up to the elections' month. The bandwidth around the 5000 inhabitants' threshold is computed following one of the procedures described by Calonico, Cattaneo, and Farrell 2020, through the minimization of the Mean Squared Error. Year and Municipality fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

Table 2 reveals that the aggregate effect of Law 215/2012 on day care-related expenses was zero. In Appendix C we show that this null aggregate impact of the law on day care expenditures is confirmed when we look at the treatment effect over time, through a dynamic Difference in Discontinuity similar to the one described by Vannutelli 2021.

Proposition 3 suggests that this null result may be due to the aggregation of positive effects at low shares of female councillors, due to more favorable policy agreements, and negative effects at higher shares, due to conflict. The next section is devoted to analyzing these possible heterogeneous effects.

3.5 Heterogeneous effects by representation

To explore heterogeneous effects with respect to the post-law share of female councillors in treated municipalities, we first generate the variable *Female post law*, indicating the percentage of female councillors elected in the first election after the new law entered into force. Next, we create dummies for different parts of the distribution of this variable across treated municipalities: below the 25^{th} percentile, and above the 50^{th}, 75^{th} or 90^{th} percentile. Finally, we interact the treatment variable, $TreatPost$, with either *Female post law* or these percentile dummies. The heterogeneous effects of the new law on day care expenditures are given by the sum of $TreatPost$ and its interaction with these variables.[45] We exclude from this analysis the few treated councils with a female majority, but our findings are robust to their inclusion.[46]

Table 3 shows how the increase in female politicians affected day care expenditures differently based on each council's post-law share of female councillors.

Table 3: **Difference in Discontinuity heterogeneous results - % female councillors**

VARIABLES	Baseline interaction Daycarepc	25^{th} percentile Daycarepc	median Daycarepc	75^{th} percentile Daycarepc	90^{th} percentile Daycarepc
TreatPost	3.533	-0.624	0.167	0.096	-0.284
	(2.342)	(1.004)	(0.984)	(0.964)	(0.965)
TreatPost*Fem. post law	-10.305*				
	(6.033)				
TreatPost*Twenty-fifth p.		1.985**			
		(0.996)			
TreatPost*Median			-1.620*		
			(0.976)		
TreatPost*Seventy-fifth p.				-3.767	
				(2.846)	
TreatPost*Ninetieth p.					-2.634
					(3.301)
Observations	2,548	2,548	2,548	2,548	2,548
R-squared	0.033	0.031	0.032	0.036	0.031
Number of municipalities	570	570	570	570	570
Pre-Law Daycarepc Mean	6.06	6.06	6.06	6.06	6.06
Mayor controls	YES	YES	YES	YES	YES
Municipality FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Robust standard errors in parentheses are clustered at the municipality level
*** p<0.01, ** p<0.05, * p<0.1

Notes: Difference in discontinuity regressions. TreatPost corresponds to the treatment coefficient, the interaction between the dummy identifying the treatment group and the dummy for the post-treatment period. Interaction terms indicate the post-law share of elected female councillors, taken with respect to the overall distribution of treated cities' councils. The dependent variable is yearly per capita expenditures on day care-related categories. Expenditures taken during election years are attributed to the council governing up to the election month. The bandwidth around the 5000 inhabitants' threshold is computed following one of the procedures described by Calonico, Cattaneo, and Farrell 2020, through the minimization of the Mean Squared Error. Year and Municipality fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

In the first column, the interaction between the treatment effect and the continuous

[45] In particular, the interaction with the 25^{th} dummy captures the effect in councils with low post-law female presence. The interactions with the 50^{th}, 75^{th} and 90^{th} percentiles assess the effect in councils with high post-law shares of female councillors. To give an idea, the median is around 38%.

[46] In these councils there is an inversion of roles that does not fit the theory: female councillors in principle have the power to propose and approve their preferred policy, unless male councillors engage in a conflict and win it.

measure of *Female post law* is negative and significant: the higher the share of female councillors, the smaller the effect of the law on day care expenditures. The subsequent columns clarify that this effect turns from positive and significant in councils with few women to negative and significant in those with many (specifically, in those with a share of female councillors in the first quartile and above the median, respectively). The effect remains negative, although not significant, in municipalities with particularly high shares of female councillors (in the last quartile or decile).

To give a sense of the magnitude of the effect, in cities with low shares of female councillors the increase in expenditure was about 1,36 euros per capita, a 22% growth with respect to the pre-law period and the control municipalities. On the other hand, in cities with *Female post law* above the median, the size of the decrease was about 1,45 euros per capita, a 24% reduction with respect to the control group.[47] The effect is even stronger if we exclude municipalities that do not decide day care expenditures independently but within unions of municipalities ("Unioni di comuni").[48]

One might expect that a law that raises female representation has stronger effects on a policy preferred by women, the stronger the increase in their representation. From this point of view, our finding that, relative to control municipalities, the introduction of the law was followed by an increase in day care expenditures in treated municipalities with few female councillors and by a reduction in those with many may appear surprising. Yet, it is in line with Proposition 3, according to which small increases in strength may lead to policy gains for women but higher increases may translate into conflict and policy losses for women.[49]

In our theory, whether a change in representation produces policy gains or losses depends

[47]Using Istat data, we verified whether this effect corresponded to a change in the number of available spots or in the number of subscriptions to day care facilities. We find a significant increase in public spots and subscriptions in cities with low shares of post-law female councillors, and no effect in those with high shares.

[48]When we drop from the sample all municipalities belonging to these unions, we find a reduction of up to 6 euros per capita in municipalities with post-law share of female councillors higher than the 75th percentile.

[49]Proposition 3 also predicts that sufficiently high gains in strength should eventually generate policy benefits for women. The evidence in the last two columns suggests that, quite reasonably, even in municipalities with large minority shares of female councillors, their probability of winning in case of conflict remains well below one (formally, below π_2 in the model).

on both the starting and final share of female councillors, which may be related to one another. In our specification, pre-law shares are absorbed by municipality fixed effects. In Appendix D we explicitly control for pre-law shares (dropping municipality fixed effects), and show that our main results are preserved.

Another potential explanation for our results could be that Law 215/2012 made inexperienced female politicians enter local councils. These new politicians might then have lacked the political skills to promote their needs and could simply have adapted to male politicians' decisions. We tend to exclude this alternative explanation, since there are no significant heterogeneous results of the law with respect to either male or female councillors' average age or political experience. Moreover, as shown by Baltrunaite et al. 2019, law 215/2012 did not alter the average age of elected officials.

3.6 Mechanism: conflict

According to our theory, the backlash in day care expenditures is due to higher conflict after larger increases in female representation. Here we provide evidence of this higher conflictuality. As a measure of conflict, we take municipal council dissolutions triggered by either the resignation of more than half of local councillors or the passage of a non-confidence motion.[50] We create a corresponding *Dissolution* dummy taking value 1 in the year in which the local council was dissolved for these political reasons, as well as in the previous year, to capture mounting conflict.[51]

In relative terms, the resignation of the majority of councillors is a rare event: only 3.8% of municipalities in our sample experience this type of dissolution. Nevertheless, these dissolu-

[50] In Italy, municipal councils can be dissolved for various reasons: persistent law violations and threats to public policy (e.g., mafia infiltrations), the mayor's forfeiture because of resignation, promotion to higher office, or death, prolonged financial distress, or political breakdown. In the latter case, this typically takes the form of the resignation of the majority of councillors or the passage of a non-confidence motion. When a council is dissolved, local administrators are replaced with external commissioners, and municipal finances are placed under supervised management until new elections are held. We focus on these political dissolutions, as they signal strong disagreement over the implementation of the political agenda.

[51] Control municipalities are dissolved more frequently for these reasons than treated municipalities; however, yearly trends between the two groups are similar.

tions provide a clear signal of conflict, and, in line with our theoretical reasoning, we expect them to occur more frequently where the new law generated a larger presence of female councillors.[52]

Table 4 replicates the Difference-in-Discontinuity specification of Table 3, but using *Dissolution* as the dependent variable.

Table 4: **Heterogeneous Law's effect on councils' dissolutions**

VARIABLES	Baseline interaction Dissolution	25^{th} percentile Dissolution	median Dissolution	75^{th} percentile Dissolution	90^{th} percentile Dissolution
TreatPost	-0.082 (0.062)	0.007 (0.023)	-0.015 (0.025)	-0.005 (0.024)	-0.002 (0.024)
TreatPost*Fem. post law	0.217 (0.135)				
TreatPost*Twenty-fifth p.		-0.060 (0.063)			
TreatPost*Median			0.043** (0.021)		
TreatPost*Seventy-fifth p.				0.041*** (0.013)	
TreatPost*Ninetieth p.					0.046*** (0.014)
Observations	2,548	2,548	2,548	2,548	2,548
R-squared	0.042	0.042	0.043	0.039	0.038
Number of municipalities	570	570	570	570	570
Pre-Law Dissolution Mean	.021	.021	.021	.021	.021

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Difference in discontinuity regressions with council's dissolution because of majority of councillors' resignation as main outcome. TreatPost corresponds to the treatment coefficient, the interaction between the dummy identifying the treatment group and the dummy for the post-treatment period. Interaction terms indicate the post-law share of elected female councillors, taken with respect to the overall distribution of treated cities' councils. The bandwidth around the 5000 inhabitants' threshold is computed following one of the procedures described by Calonico, Cattaneo, and Farrell 2020, through the minimization of the Mean Squared Error. Year and Municipality fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

From Table 4 we see that in councils where the law generated a larger increase in female councillors, this led to a higher probability of dissolution due to internal disagreement, compared to the control group: columns 3, 4 and 5 confirm our theoretical prediction.[53] On the other hand, in councils with low levels of post-law female politicians, this probability

[52]One might argue that for dissolution to reflect conflict there should be a gender divide in voting. However, persistent disagreement—such as over the administration of public funds—can itself lead to a non-confidence motion, which may be presented even by a minority of councillors and eventually result in dissolution.

[53]This evidence is in line with Gagliarducci and Paserman 2012, who show that Italian municipalities with female mayors are more likely to be dissolved, and that this probability increases in male-dominated councils. Their evidence, consistent with ours, suggests that an important determinant of group dynamics is its gender composition.

was not affected (column 2).[54] Since this is a reduced form effect, we cannot exclude other possible factors leading to council dissolution. However, this additional piece of evidence supports our claim that the law increased the probability of political conflict.

3.7 Other expenditure categories

Our main analysis focuses on child care, a policy over which men and women have substantially different preferences. Thus, it allows us to test Proposition 3, which assumes sufficiently different preferences. Many other policies are less gender-sensitive, and preferences do not differ systematically across gender lines. Proposition 2 predicts that, if preferences are sufficiently aligned, the majority will propose its bliss point and the minority, irrespective of its strength, will accept it. In this case, changes in female representation should not translate into policy changes.

To test this prediction, we consider a range of policies that are not gender-related, so that preferences should not differ between male and female councillors. Given that each municipality's balance sheet contains more than 100 expenditure categories, we focus on the 10 most relevant ones in terms of share of total expenditures, with the condition that they are decided by the local council:[55] Sport, Police, Economic development, Tourism, Parks, Public viability (public works/roads), Civil protection, Cultural events, Retirement centers, Charity.[56]

We investigate whether Law 215/2012 had an effect on the corresponding municipal expenditure, and whether such effect varies with the post quota share of female councillors.

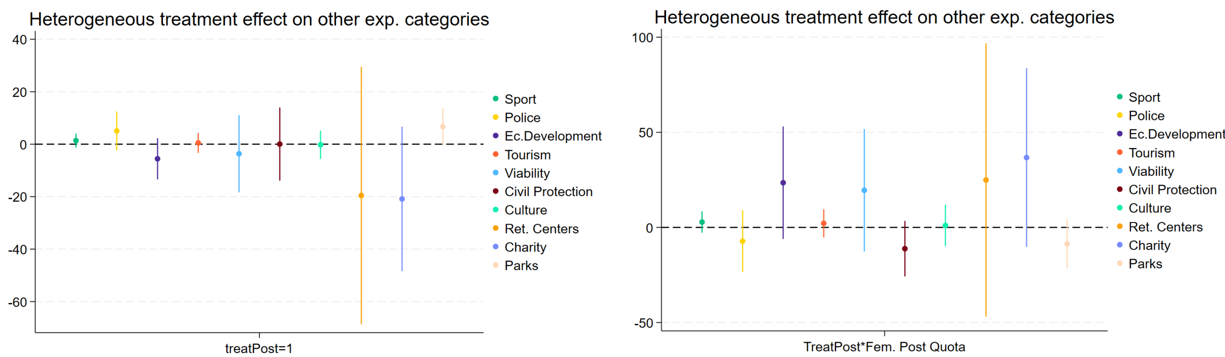
[54] To give a sense of the magnitude of the effect, for levels of post-law female councillors higher than the median, the law more than doubled the probability of dissolution due to conflict (it induced an increase of 4.3 percentage points, out of a baseline probability of 3.8%.)

[55] The national government contributes through its central funds to municipalities' expenditures, but funds are usually not restricted to a specific service or infrastructure. The local government can decide how to administer the funds and which public service needs most financial support. All expenditure categories in this exercise are under the responsibility of local administrators. A major exception is healthcare, which is the responsibility of the regional government and is not included in the exercise.

[56] The source of data for these expenditures, "Certificati Consultivi", is different from that of Day care expenditures, which is Istat. We focus on current, and not capital, expenditures per capita since they are more likely to be affected by sudden changes in council composition. However, results with capital expenditures are not different in terms of magnitude/significance level of coefficients.

To this end, for each expenditure category, we run a Difference in Discontinuity regression with the same specification as in column 1 of Table 3. For each regression, Figure 5 reports on the left the baseline *TreatPost* coefficient and on the right its interaction with *Female post-law*.

Figure 5: **Law's effect on other expenditure categories**



(a) *TreatPost* (b) *TreatPost*Female post-law*

Notes: Difference in Discontinuity heterogeneous effects with respect to the city council's post law shares of female councillors. For each of the expenditure categories indicated in the legend, we run a Difference in Discontinuity regression with the specification of column 1 of Table 3. Panel 5a reports the baseline *TreatPost* coefficient and Panel 5b its interaction with *Female post-law*. Confidence intervals are at the 95% level.

Figure 5 supports Proposition 2: when we consider policies for which preferences are not expected to differ by gender, we find that the Law had no effect on any of them (except for a marginally significant positive effect on park expenditures), and, importantly, that this null effect is independent of the level of *Female Post law*. In other words, as predicted by the theory, the increase in female representation produced by the Law did not affect policy choices on non-gender-sensitive issues.

This exercise is also useful for reducing concerns over other potential confounding policies (e.g., the Domestic Stability Pact), which could have driven the differential trends in day care expenditures between treated and control municipalities.

In Appendix E we perform two additional placebo tests, in which the dependent variable is day care expenditure but we apply artificial population cutoffs and year of law implementation. The heterogeneous effect we document in Table 3 does not appear in the placebo exercises, confirming that our main result is not spurious.

4 Internal and external validity

In this section we explore the internal and external validity of our findings, and provide evidence supporting the broader validity of the theory. We first present two robustness checks related to our main explanatory variable and to the identification strategy: we compare municipalities with high and low shares of post-law female councillors, and we adopt an instrumental variable strategy to explore the effect of reform-induced changes in female representation. Next, we explore the broader validity of the theory, and the external validity of the empirical findings, by exploiting the introduction of a gender quota in 2011 in Spanish administrative elections (Bagues and Campa 2021). As mentioned in Section 1, this adds to the evidence produced by Baskaran and Hessami 2025 for the Bavarian context, where an additional female councillor accelerates the expansion of public child care only in councils with few women, again consistent with our theory.

4.1 Robustness on representation

A major concern for our analysis is that what we are observing is not the heterogeneous effect with respect to the post-law share of female councillors but with respect to some other factor correlated with this variable. To exclude this hypothesis, we compare the average levels of some representative variables for treated cities with high and low shares of female councillors (respectively, post-law female councillors higher than the 75^{th} percentile and lower than the 25^{th} percentile), and check if the differences in means are statistically significant.[57] In Table 5 we can observe these comparisons.

[57] We look at variables referring to municipality balance sheets (e.g., total expenditures), population characteristics (e.g., years of education or age structure), geographical characteristics and local politicians' attributes.

Table 5: **Test for difference in means, cities with low and high levels of post-law female councillors**

Variable	High_Share	Low_Share	Difference	p-value
Tot. Expenditures	2941032	2574264	-366768	0.006**
Surface (km^2)	33.728	31.384	-2.344	0.582
Mean altitude	222.419	222.769	0.351	0.988
Population density	365.441	339.519	-25.922	0.508
South	0.338	0.277	-0.061	0.275
Gender participation gap 2011	20.249	19.868	-0.381	0.486
Female labor force participation 2011	41.544	42.725	1.181	0.172
Female unemployment rate 2011	12.623	12.181	-0.442	0.603
Female mayor	0.147	0.169	0.022	0.612
Y.o.E. mayor	15.762	15.795	0.033	0.920
Mean Y.o.E. councillors	14.346	14.279	-0.068	0.667
Mean age councillors	49.628	50.773	1.145	0.013**
Mean Y.o.E. aldermen	14.438	14.530	0.092	0.812
Mean age aldermen	45.964	45.331	-0.633	0.524
Mean Y.o.E. female councillors	14.451	14.289	-0.162	0.724
Mean age female councillors	43.018	43.404	0.386	0.556
Mean Y.o.E. male councillors	13.892	13.972	0.080	0.683
Mean age male councillors	48.063	48.061	-0.002	0.997
Union of municipalities dummy	0.321	0.373	0.052	0.352
Daycare p.c.	7.067	5.826	-1.241	0.348
Age structure (M)	1.343	1.324	-0.019	0.675
Age structure (F)	1.762	1.706	-0.057	0.382
Age structure (Tot.)	1.545	1.508	-0.038	0.485
Fertility rate	40.011	39.121	-0.891	0.570
River	0.426	0.410	-0.017	0.776
Lake	0.320	0.361	0.042	0.457
Sea distance	69.249	79.817	10.568	0.108

Notes: Test for statistical differences in means of variables between group of treated municipalities with low and high shares of post-law female councillors (post-law female councillors$<25^{th}$ percentile and post-law female councillors$>75^{th}$ percentile). Demographic variables refer to the whole municipality's population, or to the whole population by gender. Age ratios correspond to the ratio between the population over 60 and the population under 20 years old. Information on revenues and expenditures comes from municipalities' balance sheets. South is a dummy with value 1 in case that the municipality is located in Southern Italy. Only treated municipalities belonging to our optimal bandwidth are included.

The table shows a few statistically significant differences between the two groups of cities. Specifically, cities with high levels of post-law female councillors tend to spend more on average and have slightly younger councillors. The latter difference is not problematic, since it is only one year of difference in means and it should not bias our results. Regarding the other difference, if we include Total Expenditures as a control in our Difference in Discontinuity results, the main coefficients' magnitude, size and significance levels are preserved. Thus, we can be more confident that our main results are driven by a combination of the introduction of the new law and the subsequent shares of female councillors generated in the treated municipalities' local councils.[58] In particular, we underline how fertility rates, which are a potential huge driver of the demand for day care, are comparable between the two groups of cities.

4.2 Robustness on identification: 2SLS panel regression

We perform a further robustness analysis by changing the identification strategy and resorting to an instrumental variable approach. One might suspect that the Female post law variable is endogenous to some unobserved municipal characteristics: the shares of elected female politicians can depend on a large variety of factors intrinsic to the electorate's and city's characteristics (Hessami and Fonseca 2020). Thus, given the possible endogeneity of this variable, in this section we instrument the percentage of female councillors with *Treat-post*: in this way, we can evaluate the effect on day care expenditure of the exogenous increase in female politicians driven by the introduction of the new law. The model we adopt in this Section takes the following specification:

1st stage:

$$FemaleCouncillors_{it} =$$
$$\alpha + \beta * Treatment_i + \gamma * PostLaw_{it} + \delta * Treatpost_{it} + Year_t + City_i + \Lambda * Mayor_{it} + \epsilon_{it}$$

2nd stage:

$$Daycarepc_{it} =$$
$$\zeta + \eta * Treatment_i + \theta * PostLaw_{it} + \iota * \widehat{FemaleCouncillors}_{it} + Year_t + City_i + \kappa * Mayor_{it} + \mu_{it}$$

The coefficient ι captures the Local Average Treatment Effect of female councillors on day care expenditures, that is the effect of the law-induced increase in the share of female

[58] We performed also the test in differences (i.e. comparing municipalities with high post-law increases in female councillors and municipalities with low increases in female councillors), and the results were in line with this test. Few variables exhibited statistically significant differences in means, for instance total revenues, but not so large in magnitude to generate concerns.

councillors on funding for day care.[59] Table 6 presents the results from our instrumental variable analysis, with first and second-stage results, together with the OLS baseline. Column 1 reports the first-stage regression, showing that the introduction of Law 215/2012 significantly increased the share of female councillors in treated municipalities. Columns 2 to 4 display the second-stage regressions alongside the baseline OLS estimate. We report the 2SLS estimates both with the inclusion and exclusion of the variables $Treatment$ and $Postquota$, to provide a more comprehensive overview of the results.

Table 6: **First and second stage regressions (2SLS), and OLS baseline**

VARIABLES	2SLS 1st stage Female councillors	OLS Daycarepc	2SLS 2nd stage Daycarepc	2SLS 2nd stage Daycarepc
TreatPost	0.100***			
	(0.008)			
Female councillors		0.043	-5.402***	-5.120*
		(0.766)	(1.915)	(2.783)
Treatment		-2.152***		-2.030***
		(0.718)		(0.711)
Postlaw		-0.184		0.386
		(0.203)		(0.365)
Municipality FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Observations	8,027	8,027	8,027	8,027
Number of municipalities	1,424	1,424	1,424	1,424
First stage F test	120.0			

Robust standard errors in parentheses, clustered at the municipality level
*** p<0.01, ** p<0.05, * p<0.1

Notes: Column (1) reports the first stage of the 2SLS setting, where the dependent variable is the share of female councillors. TreatPost is the interaction between a dummy for treated municipalities and a post-law dummy. Columns (2) to (4) report the OLS and 2SLS second stage results, where the dependent variable is per capita day care expenditure. Female councillors is instrumented using TreatPost. All regressions include municipality and year fixed effects. Robust standard errors are clustered at the municipality level.

Results from Table 6 indicate that a higher share of female councillors leads to decreased per capita spending on day care, with the 2SLS estimates being substantially larger in magnitude than the OLS counterpart, suggesting the presence of endogeneity in the simple

[59]This effect is "Local" in the sense that it concerns only those treated municipalities that had their shares of female councillors increased by the new law. For this identification strategy to be valid, we need the instrument to influence the outcome only through its effect on councillors' sex ratio (exclusion restriction assumption). We argue that this is the case: the new law was imposed on a set of municipalities irrespective of their expenditure composition, it did not change how the local administrators could spend public money, nor did it affect their revenues. Moreover, it's hard to think that the new law generated any effect on the demand for day care, since we saw how other variables such as the number of spots did not change. Thus, we argue that being subject to the new law did not alter municipal expenditures for day care directly, but it affected day care only through its effect on female councillors.

OLS specification.[60] On the other hand, the correlation between *Female councillors* and day care expenditures is almost null and non-significant in the OLS results: this is not particularly surprising, since this correlation also takes into account control municipalities and the period before the introduction of the law.

The magnitude of the last column's effect is larger than our Diff-in-Disc results: the decrease in 2SLS results is roughly 5 euros per capita, compared to the 1,45 euros per capita in the Diff-in-Disc setting (municipalities with polarization levels higher than the median). There might be several reasons for this: first of all, Table 6 presents a LATE effect, which applies to a subset of treated municipalities, the compliers, whose share of female councillors was increased by the law: this effect might be different from the effect at the threshold observed in the Diff-in-Disc. Moreover, we include here all municipalities in our sample, and not exclusively those belonging to a narrow bandwidth around the cutoff. Plus, the 2SLS setting allows us to overcome possible biases due to endogeneity in the *Female post-law* variable: it might be that these biases shrink the coefficients in the Diff-in-Disc framework. However, this last piece of evidence confirms that the law-induced increase in female councillors had on average negative effects on the funding for day care, in line with our theoretical framework.[61]

[60] When we substitute the share of female councillors with the increase in female councillors, 2SLS results are in line. To be more specific, the instrumented variable's coefficient has a magnitude of -5.432 and it is significant at the 1% level in the most complete specification.

[61] Given the results in Table 6, it is important to identify the compliers—i.e., the municipalities in which the increase in female councillors was driven by the introduction of the law. While it is not possible to determine with certainty which municipalities experienced this increase solely because of the reform, we approximate the set of compliers as treated municipalities that registered a rise in the share of female councillors after 2012. Conversely, likely non-compliers are treated municipalities that did not exhibit such an increase. Compliers constitute the majority of municipalities in our sample (around 80%). They are more likely to be located in the South and tend to have lower average day care expenditures. This suggests that Southern municipalities may have been major drivers of our results: starting from lower expenditure levels, child care spending may have been a more salient topic of council debate. Moreover, these municipalities also began with lower shares of female councillors, so the increase induced by the law may have disrupted prior political equilibria, potentially leading to greater conflict within newly elected councils.

4.3 External validity: Spanish elections

In the context of Spanish local elections, a gender quota on candidates' lists was introduced in 2007 for all municipalities with more than 5000 inhabitants and in 2011 it was extended to municipalities with at least 3000 inhabitants. The quota provision mandated that each electoral list include no less than 40% of candidates from each sex. Bagues and Campa 2021 study the effect of this quota on a wide variety of outcomes, finding mainly an increase in the percentage of female councillors by 4 percentage points, but no effects on politicians' quality or on the size and composition of public finances. Overall, the setting is quite similar to the Italian one, with a sudden increase in the share of female councillors affecting only a subsample of municipalities. Therefore, we might think that this could have triggered a potential heterogeneous effect on expenditures, close to the one we present in this paper. Consequently, we try to replicate our result with the data from Bagues and Campa 2021: we look at the differential effect of the quota with respect to the generated increase in female councillors across treated Spanish municipalities.

We focus only on the 3000 population threshold,[62] and we look at the quota's effect on expenditures across 3 different electoral cycles (2007, 2011, 2015) with respect to the post-quota share of female councillors. The main outcome of our regressions is the "female expenditures per capita" variable constructed by the paper's authors, including all expenditures for which there is a proven gender difference in preferences according to surveys conducted among the Spanish population.[63] We start by replicating our Difference-in-Discontinuity results from Table 3, and we present them in Table 7.

[62] This is because at 5000 inhabitants there is another confounder, which is a higher level of transfers received from the central government.

[63] The authors categorize as female expenditures Social security and protection, Education, Social promotion and Health, while male expenditures include Housing and urbanism, Basic infrastructure and transport, Agricultural infrastructure, and Agriculture, hunting and fishing.

Table 7: **Difference in Discontinuity - Spanish municipalities**

VARIABLES	Baseline interaction F. Expenditures	25^{th} percentile F. Expenditures	median F. Expenditures	75^{th} percentile F. Expenditures	90^{th} percentile F. Expenditures
TreatPost	37.151	18.740	19.888	17.663	17.678
	(41.892)	(14.516)	(15.353)	(14.167)	(14.172)
TreatPost*Fem. post quota	-64.609				
	(102.349)				
TreatPost*Twenty-fifth p.		-6.788			
		(13.301)			
TreatPost*Median			-6.473		
			(15.528)		
TreatPost*Seventy-fifth p.				-66.693***	
				(12.569)	
TreatPost*Ninetieth p.					-66.601***
					(12.566)
Observations	4,373	7,026	7,026	7,026	7,026
R-squared	0.122	0.108	0.108	0.108	0.108
Number of municipalities	809	862	862	862	862
Municipality FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Robust standard errors in parentheses are clustered at the municipality level
*** p<0.01, ** p<0.05, * p<0.1

Notes: Difference in discontinuity regressions with Spanish municipalities. The gender quota was imposed in 2011 on municipalities with more than 3000 inhabitants, which represent our treated sample. TreatPost corresponds to the treatment coefficient, the interaction between the dummy identifying the treatment group and the dummy for the post-treatment period. Interaction terms indicate the post-quota share of elected female councillors, taken with respect to the overall distribution of treated cities' councils. The dependent variable is the sum of all "female expenditures", divided by the municipality's population. Expenditures taken on election years are attributed to the council governing up to the elections' month. The bandwidth around the 3000 inhabitants' threshold is computed following one of the procedures described by Calonico, Cattaneo, and Farrell 2020, through the minimization of the Mean Squared Error. Year and Municipality fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

As we can see from Table 7, in municipalities with high levels of post-quota female politicians (last two columns), expenditures for female-related categories were reduced with respect to the control group. On the other hand, cities with low levels did not observe any significant effect, as shown in the column related to the 25^{th} percentile.

We can extend this analysis by replicating another result of our study, the 2SLS regressions in Table 6. In brief, we instrument the percentage of female councillors with the variable $TreatPost$, and thus assess the effect of the quota-related increase in female councillors on female expenditures. In Table 8 we present these additional results.

Table 8: **First stage, OLS, and second stage regression (Spain)**

VARIABLES	First stage Sh. female councillors	OLS F. Expenditures	2SLS F. Expenditures	2SLS F. Expenditures
TreatPost	0.032***			
	(0.006)			
Treatment	-0.014*	-42.622***		-43.939***
	(0.008)	(8.343)		(9.290)
Sh. female councillors		23.491	-734.717***	-510.063***
		(16.166)	(216.847)	(185.687)
Observations	32,978	32,767	32,767	32,767
R-squared	0.087	0.057		
Number of municipalities	4,371	4,365	4,365	4,365
Muni FE	yes	yes	yes	yes
Year FE	yes	yes	yes	yes

Robust standard errors are clustered at the municipality level
*** p<0.01, ** p<0.05, * p<0.1

Notes: This table reports the first stage, OLS, and second stage regressions for the Spanish sample. The dependent variable in the second to fourth columns is per capita expenditures on policies targeted to women. In the first column, the dependent variable is the share of female councillors. The variable *TreatPost* is the interaction between a treatment group dummy (municipalities with more than 3,000 inhabitants) and a post-treatment period dummy. Municipalities with more than 5,000 inhabitants are excluded, as they have been treated for more than one electoral cycle. Year and municipality fixed effects are included in all specifications. Robust standard errors, clustered at the municipality level, are reported in parentheses.

We can see from Table 8 that the quota-induced increase in female councillors had a negative average effect on treated municipalities compared to control ones. The effect is sizeable, as in the most complete specification one percentage point increase in female councillors leads to a reduction of 500 euros per capita in female expenditures in treated municipalities, relative to control ones.[64] Moreover, the first stage results confirm that the quota increased the share of female councillors across treated municipalities by 3.2 percentage points.

Tables 7 and 8 confirm that the main result of our paper is not a peculiarity of the Italian context: Spanish data show another case in which a sudden increase in the share of female politicians caused a reduction in expenditures in the most "gender-sensitive" categories. Once again, while possibly surprising, this is in line with our theory of conflict and backlash, and also adds to the evidence presented by Baskaran and Hessami 2025 for Bavarian elections, which we discussed above. Overall, we observe similar non-obvious patterns in Italy, Spain and Germany, which are well explained by our theory.

[64] To be more precise, this is a LATE effect and therefore it is only exclusive to the compliers, namely the municipalities for which the percentage of female councillors was increased by the quota.

5 Conclusion

Western societies have long tried to increase the political power and representation of social groups that were unable to influence political decisions. Measures such as the extension of suffrage to Afro-Americans and the introduction of gender quotas have typically provided long-term benefits to these groups, for example by granting them better social rights and access to welfare programs (Miller 2008, Fujiwara 2015, Chattopadhyay and Duflo 2004). However, the path to these gains has often been rough and marked by conflict and temporary backlash. In this paper, we reflect on these short-term tensions and difficulties. We document empirically and explain theoretically how a sudden increase in political power may trigger conflict and translate into a temporary worsening of policy outcomes.

In a stylized model featuring a group with the power to set a policy and another group that may accept it peacefully or not, we show that conflict emerges in equilibrium if preferences are sufficiently different and the second group becomes sufficiently strong (although not too strong).

In the empirical part, we exploit the introduction in 2012 of a law that substantially raised the share of female politicians at the local level in a subsample of Italian municipalities. Through a difference-in-discontinuity identification strategy, we document that a sufficiently strong increase in female representation led to a significant and sizable reduction in municipal spending on day care, a highly gender-sensitive issue, on which male and female politicians have different preferences.

This non-obvious finding is in line with our theory of conflict and backlash. We provide evidence of the mechanism by documenting an increase in council dissolution due to disagreement among councillors. We also show that, again in line with the theory, the reform had positive effects on day care expenditure where the increase in female representation was small, and no effect on issues that are not gender-sensitive.

We support the internal validity of our findings by running several robustness checks: we show that that are no substantial differences between municipalities with high and low post-law shares of female councillors, we run different placebo tests, and also exploit an alternative instrumental variable identification strategy. We also support external validity by exploiting a 2011 reform to Spanish local elections that similarly generated a sudden increase in the share of female councillors in a subsample of municipalities. Again, we show that municipalities experiencing larger increases in female representation experienced a significant reduction in expenditures on “female-preferred” policy categories. Baskaran and Hessami 2025 report similar patterns for child care expenditure in Bavarian municipalities, which further confirms the external validity.

We thus document that the short-run consequences of enhancing a minority’s political representation can be more complex than expected, and may be marked by conflict and a backlash. Our theory provides a simple and parsimonious explanation for this nontrivial pattern. It suggests different strategies to reduce the chances of conflict: providing sufficient political power to the initially underrepresented group, favoring its commitment capacity, or reducing the distance in preferences between the two groups. If none of these options is feasible, in the short term an increase in the political power of an underrepresented group may be accompanied by conflict and less favorable policies. However, since increasing the political power of minorities typically grants consistent long-term benefits, we are certainly not advocating limiting their enfranchisement.

References

Acemoglu, Daron and James A Robinson (2000). “Why did the West extend the franchise? Democracy, inequality, and growth in historical perspective”. In: *The Quarterly Journal of Economics* 115.4, pp. 1167–1199.

— (2001). “A theory of political transitions”. In: *American Economic Review* 91.4, pp. 938–963.

Acemoglu, Daron and Alexander Wolitzky (2024). "Mistrust, misperception, and misunderstanding: Imperfect information and conflict dynamics". In: *Handbook of the Economics of Conflict*. Vol. 1. Elsevier, pp. 59–104.

Andreoli, Francesco, Elena Manzoni, and Margherita Margotti (2022). "Women at work: Gender quotas, municipal elections and local spending". In: *European Journal of Political Economy* 75, p. 102175.

Aneja, Abhay P and Carlos F Avenancio-León (2019). "Disenfranchisement and economic inequality: Downstream effects of shelby county v. holder". In: *AEA Papers and Proceedings*. Vol. 109. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, pp. 161–165.

Bagues, Manuel and Pamela Campa (2021). "Can gender quotas in candidate lists empower women? Evidence from a regression discontinuity design". In: *Journal of Public Economics* 194, p. 104315.

Baliga, Sandeep and Tomas Sjöström (2020). "The strategy and technology of conflict". In: *Journal of Political Economy* 128.8, pp. 3186–3219.

— (2024). "Causes of war". In: *Handbook of the Economics of Conflict*. Vol. 1. Elsevier, pp. 1–58.

Baltrunaite, Audinga et al. (2019). "Let the voters choose women". In: *Journal of Public Economics* 180, p. 104085.

Barigozzi, Francesca, Helmuth Cremer, Chiara Monfardini, et al. (2019). "The gender gap in informal child care: theory and some evidence from Italy". In: *Economia Italiana* 2019, pp. 37–71.

Baron, David P and John A Ferejohn (1989). "Bargaining in legislatures". In: *American political science review* 83.4, pp. 1181–1206.

Baskaran, Thushyanthan and Zohal Hessami (2025). "Women in political bodies as policymakers". In: *Review of Economics and Statistics*, forthcoming.

Beaman, Lori et al. (2012). "Female leadership raises aspirations and educational attainment for girls: A policy experiment in India". In: *science* 335.6068, pp. 582–586.

Besley, Timothy and Anne Case (1995). "Does electoral accountability affect economic policy choices? Evidence from gubernatorial term limits". In: *The Quarterly Journal of Economics* 110.3, pp. 769–798.

Besley, Timothy and Stephen Coate (1997). "An economic model of representative democracy". In: *The quarterly journal of economics* 112.1, pp. 85–114.

Bhalotra, Sonia and Irma Clots-Figueras (2014). "Health and the political agency of women". In: *American Economic Journal: Economic Policy* 6.2, pp. 164–97.

Bonomi, Giampaolo, Nicola Gennaioli, and Guido Tabellini (2021). "Identity, beliefs, and political conflict". In: *The Quarterly Journal of Economics* 136.4, pp. 2371–2411.

Calonico, Sebastian, Matias D Cattaneo, and Max H Farrell (2020). "Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs". In: *The Econometrics Journal* 23.2, pp. 192–210.

Casarico, Alessandra, Salvatore Lattanzio, and Paola Profeta (2022). "Women and local public finance". In: *European Journal of Political Economy* 72, p. 102096.

Cella, Michela and Elena Manzoni (2023). "Gender bias and women's political performance". In: *European Journal of Political Economy* 77, p. 102314.

Chattopadhyay, Raghabendra and Esther Duflo (2004). "Women as policy makers: Evidence from a randomized policy experiment in India". In: *Econometrica* 72.5, pp. 1409–1443.

Clots-Figueras, Irma (2012). "Are female leaders good for education? Evidence from India". In: *American Economic Journal: Applied Economics* 4.1, pp. 212–44.

Coffman, Katherine Baldiga (2014). "Evidence on self-stereotyping and the contribution of ideas". In: *The Quarterly Journal of Economics* 129.4, pp. 1625–1660.

Crawford, Vincent P (1982). "A theory of disagreement in bargaining". In: *Econometrica* 50.3, pp. 607–637.

Downs, Anthony (1957). "An economic theory of democracy". In: *Harper and Row* 28.

Ellingsen, Tore and Topi Miettinen (2008). “Commitment and conflict in bilateral bargaining”. In: *American Economic Review* 98.4, pp. 1629–1635.

Esteban, Joan and Debraj Ray (1994). “On the measurement of polarization”. In: *Econometrica* 620.4, pp. 819–851.

— (2011). “Linking conflict to inequality and polarization”. In: *American Economic Review* 101.4, pp. 1345–1374.

Fearon, James D (1995). “Rationalist explanations for war”. In: *International organization* 49.3, pp. 379–414.

— (1998). “Bargaining, enforcement, and international cooperation”. In: *International organization* 52.2, pp. 269–305.

— (2004). “Why do some civil wars last so much longer than others?” In: *Journal of Peace Research* 41.3, pp. 275–301.

Ferreira, Fernando and Joseph Gyourko (2014). “Does gender matter for political leadership? The case of US mayors”. In: *Journal of Public Economics* 112, pp. 24–39.

Fox, Richard L and Jennifer L Lawless (2011). “Gendered perceptions and political candidacies: A central barrier to women’s equality in electoral politics”. In: *American journal of political science* 55.1, pp. 59–73.

Fujiwara, Thomas (2015). “Voting technology, political responsiveness, and infant health: Evidence from Brazil”. In: *Econometrica* 83.2, pp. 423–464.

Funk, Patricia and Christina Gathmann (2015). “Gender gaps in policy making: Evidence from direct democracy in Switzerland”. In: *Economic Policy* 30.81, pp. 141–181.

Gagliarducci, Stefano and Tommaso Nannicini (2013). “Do better paid politicians perform better? Disentangling incentives from selection”. In: *Journal of the European Economic Association* 11.2, pp. 369–398.

Gagliarducci, Stefano and M Daniele Paserman (2012). “Gender interactions within hierarchies: evidence from the political arena”. In: *The Review of Economic Studies* 79.3, pp. 1021–1052.

Gelman, Andrew and Guido Imbens (2019). “Why high-order polynomials should not be used in regression discontinuity designs”. In: *Journal of Business & Economic Statistics* 37.3, pp. 447–456.

Genicot, Garance and Debraj Ray (2024). “Identities in conflict”. In: *Handbook of the Economics of Conflict*. Vol. 1. Elsevier, pp. 201–247.

Geys, Benny and Rune J Sørensen (2019). “The impact of women above the political glass ceiling: Evidence from a Norwegian executive gender quota reform”. In: *Electoral Studies* 60, p. 102050.

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano (2016). “Do fiscal rules matter?” In: *American Economic Journal: Applied Economics*, pp. 1–30.

Grossman, Gene M and Elhanan Helpman (2021). “Identity politics and trade policy”. In: *The Review of Economic Studies* 88.3, pp. 1101–1126.

Güth, Werner and Martin G Kocher (2014). “More than thirty years of ultimatum bargaining experiments: Motives, variations, and a survey of the recent literature”. In: *Journal of Economic Behavior & Organization* 108, pp. 396–409.

Güth, Werner, Rolf Schmittberger, and Bernd Schwarze (1982). “An experimental analysis of ultimatum bargaining”. In: *Journal of Economic Behavior & Organization* 3.4, pp. 367–388.

Hessami, Zohal and Mariana Lopes da Fonseca (2020). “Female political representation and substantive effects on policies: A literature review”. In: *European Journal of Political Economy* 63, p. 101896.

Júlio, Paulo and José Tavares (2017). “The good, the bad and the different: can gender quotas raise the quality of politicians?” In: *Economica* 84.335, pp. 454–479.

Kjaer, Ulrik and Mona Lena Krook (2019). “The blame game: analyzing gender bias in Danish local elections”. In: *Politics, Groups, and Identities*.

Marchese, Alda, Paola Profeta, and Giulia Savio (2025). "Who cares about childcare? Covid-19 and gender differences in local public spending". In: *International Tax and Public Finance*, pp. 1–21.

Miller, Grant (2008). "Women's suffrage, political responsiveness, and child survival in American history". In: *The Quarterly Journal of Economics* 123.3, pp. 1287–1327.

Nash, John (1953). "Two-person cooperative games". In: *Econometrica*, pp. 128–140.

Powell, Robert (1999). *In the shadow of power: States and strategies in international politics*. Princeton University Press.

— (2004). "The inefficient use of power: Costly conflict with complete information". In: *American Political Science Review* 98.2, pp. 231–241.

— (2006). "War as a commitment problem". In: *International Organization* 60.1, pp. 169–203.

Putnam, Robert D, Robert Leonardi, and Raffaella Y Nanetti (1994). *Making democracy work*. Princeton university press.

Rubinstein, Ariel (1982). "Perfect equilibrium in a bargaining model". In: *Econometrica*, pp. 97–109.

Savio, Giulia (2024). "Beyond the party push: Gender differences in voters' persuasion". In: *European Journal of Political Economy* 85, p. 102613. ISSN: 0176-2680. DOI: https://doi.org/10.1016/j.ejpoleco.2024.102613. URL: https://www.sciencedirect.com/science/article/pii/S0176268024001150.

Schelling, Thomas C (1960). *The Strategy of Conflict*. Harvard University Press: Boston, MA.

Slegten, Caroline and Bruno Heyndels (2020). "Within-party sex gaps in expenditure preferences among flemish local politicians". In: *Politics & Gender* 16.3, pp. 768–791.

Vannutelli, Silvia (2021). "From lapdogs to watchdogs: Random auditor assignment and municipal fiscal performance in Italy". In: *Job Market Paper: Yale University*.

Appendices

A Difference in Discontinuity: validity assumptions

For our identification strategy to correctly assess the true effect of the Law 215/2012 on day care expenditures, three validity assumptions need to be satisfied (Grembi, Nannicini, and Troiano 2016). The assumptions are the following:

1. All potential outcomes are continuous in the running variable at the threshold
2. The effect of the confounding policy is constant over time
3. The effect of the treatment at the cutoff does not depend on the confounding policy

The first condition is that at our threshold of 5000 inhabitants, all potential outcomes are continuous in the running variable, which is the normalized population.[65] To check compliance with this condition, we test whether a wide set of covariates is balanced at the threshold by performing a series of pre-treatment Regression Discontinuity designs, using these covariates as outcomes. The set of covariates includes both geographic time-invariant characteristics[66] and demographic time-varying factors.[67] In Tables 9 and 10 we present our results.

[65] Normalized population here is defined as the municipality's population - 5000.

[66] Among the time-invariant characteristics, we have South (dummy for municipality located in the south), River (dummy for river presence), Lake (dummy for lake presence), distance from the sea, surface extension, mean altitude level and dummy rural

[67] The set of demographic factors includes fertility rate, gender participation gap, female unemployment rate, percentage of graduated men and women and age structure (ratio of over 65 over under 15 years old residents)

Table 9: **RDD with time-invariant characteristics**

VARIABLES	(1) South	(2) River	(3) Lake	(4) Surface	(5) Sea distance	(6) Altitude	(7) Rural
RD_Estimate	0.084**	0.021	-0.033	0.821	-3.074	42.099*	0.016
	(0.041)	(0.066)	(0.055)	(4.442)	(7.586)	(24.130)	(0.052)
Mean Dep.Var.	.276	.449	.36	35.847	72.036	266.073	.374
Observations	3,634	3,599	3,599	3,599	3,599	3,548	3,634

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Regression discontinuity designs with municipalities' time-invariant covariates as main outcomes. The running variable is the normalized population, with 5000 inhabitants (corresponding to zero in the normalized population) as the discontinuity cutoff. Outcomes include South (dummy for municipality located in the south), River (dummy for river presence), Lake (dummy for lake presence), distance from the sea, surface extension, mean altitude level and a dummy for rural area (i.e. density of population under 150 inhabitants/sqkm). Regressions are performed over the pre-treatment period, in other words over the period before the first elections with Law 215/2012 in force. Sampled municipalities are 1424 out of 8100 Italian municipalities, observed over the period between 2013 and 2018.

Table 10: **RDD with time-variant characteristics**

VARIABLES	(1) Fertility	(2) Gender Part.gap	(3) FPLF	(4) F.Unempl.	(5) % Grad.women	(6) % Grad.men	(7) Age structure
RD Estimate	0.570	-0.669	0.804	0.949	0.373	0.325	0.025
	(1.682)	(0.491)	(0.788)	(0.684)	(0.292)	(0.299)	(0.055)
Mean Dep.Var.	41.74	19.867	42.29	12.488	8.399	6.853	1.594
Observations	2,723	3,589	3,589	3,577	3,628	3,628	3,575

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Regression discontinuity designs with municipalities' time-variant covariates as main outcomes. The running variable is the normalized population, with 5000 inhabitants (corresponding to zero in the normalized population) as the discontinuity cutoff. Outcomes include fertility rates (average number of children per woman), gender participation gap (difference between male and female labor force participation), female labor force participation, female unemployment rate, percentage of graduated women, percentage of graduated men and age structure. (ratio between share of population over 65 and under 15). Regressions are performed over the pre-treatment period, in other words over the period before the first elections with Law 215/2012 in force. Sampled municipalities are 1424 out of 8100 Italian municipalities, observed over the period between 2013 and 2018.

As we can observe, there are no jumps in the majority of municipalities' characteristics at the 5000 inhabitants' threshold, thus it seems that these potential outcomes are continuous in the running variable at this cutoff. The only concerning discontinuity is for Southern municipalities: above the threshold, there are more municipalities in the South of Italy. This could be worrying given the well documented differences in terms of economic development and social capital across Italian regions (Putnam, Leonardi, and Nanetti 1994). However, in terms of gender equality measures, which could be the main confounders in our analysis linked to both treatment and geographical location, there are no jumps, as we can see from

Table 10. Therefore, we argue that this confounder should not bias our results.

In addition, we argue that any manipulation in the running variable (for instance, by mayors seeking a higher salary) is unlikely. First, because the set of laws that becomes binding at population thresholds, such as Law 215/2012, takes as reference value the population measured by the last census, in this case 2011.[68] Second, population is measured by independent employees from Istat,[69] with no presumable interest in manipulating the true value. However, in Appendix B we perform the McCrary test to verify the presence of manipulation in our running variable at the threshold, confirming that its observational density is continuous.

Turning to the second validity assumption, we need the confounding policy, the mayor's salary, to have constant effects over time. We know that a higher salary can influence both the selection of citizens into politics and the kind of policies implemented (Gagliarducci and Nannicini 2013). The authors who studied the effects of Law 215/2012 found no change in candidates' lists composition, in terms of age, education, or previous job (Baltrunaite et al. 2019) after the novelties were introduced: thus, we can state that the selection effect was constant over time. Speaking instead of the effects of having different kinds of mayors on the policies implemented, our solution is to control for mayors' characteristics in our set of regressions. Each of the regressions we show in the Main Results' section includes controls for the current mayor's sex, age, and years of education, which should capture this potential confounding effect on policies.

Finally, we need to discuss the third validity assumption: there should be no interaction between the new law's effects and the confounding policy. To check whether the assumption is satisfied, we perform a series of Difference in Discontinuity regressions interacting the treatment variable with a set of variables for each mayor's characteristic. Indeed, we know that a higher salary, the confounding policy, creates incentives for different individuals to enter into politics: thus, if our treatment effect is different because of this confounder, we

[68]Censuses in Italy have been performed every 10 years since 1861 until 2019, when they became yearly.
[69]Istat is the Italian National Institute of Statistics

should observe significant interaction effects between the treatment coefficient and mayors' characteristics. In particular, higher salaries might influence the type of individuals who enter politics, potentially shaping how they interact with newly elected councillors. Thus, if our treatment effect differs because of this confounder, we should observe significant interaction effects between the treatment coefficient and mayors' characteristics. The characteristics we consider are the ones for which we have data on the Ministry of Interior's website: the mayor's sex, education and age. In Table 11 we check whether any of these interaction effects is relevant or present.

Table 11: **Interactions treatment coefficient and mayor's characteristics**

VARIABLES	(1) Daycarepc	(2) Daycarepc	(3) Daycarepc
TreatPost	-0.056	-2.446	2.182
	(0.926)	(2.111)	(2.639)
Mayor_fem	0.504		
	(0.505)		
TreatPost*Mayor_fem	-3.698**		
	(1.777)		
Mayor_age		-0.008	
		(0.024)	
TreatPost*Mayor_age		0.035	
		(0.039)	
Mayor_Ed			-0.090
			(0.087)
TreatPost*Mayor_Ed			-0.173
			(0.156)
Municipality FE	YES	YES	YES
	YES	YES	YES
Year FE	YES	YES	YES
	YES	YES	YES
Observations	2,623	2,591	2,596
R-squared	0.035	0.029	0.031
Number of municipalities	581	579	578

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Notes: Difference in discontinuity regressions. TreatPost corresponds to the treatment coefficient, the interaction between the dummy identifying the treatment group and the dummy for the post-treatment period. TreatPost is interacted with each one of the current mayor's characteristics, namely sex, age and education. The dependent variable is yearly per capita expenditures on day care related categories. The bandwidth around the 5000 inhabitants' threshold is computed following one of the procedures described by Calonico, Cattaneo, and Farrell 2020, through the minimization of the Mean Squared Error. Year and Municipality fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

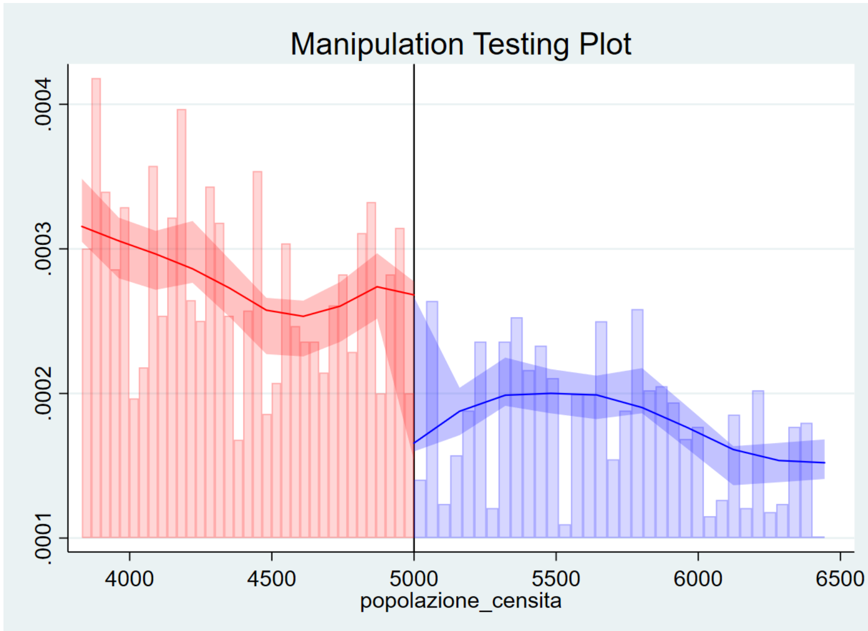
As we can observe from Table 11, there are no significant interaction effects between our treatment coefficient and the mayor's age or education. However, we observe a heterogeneous and negative treatment effect in the case of a female mayor. This is not a concern for our analysis, for two reasons: first, we control for the mayor's gender in each of our regressions. Second, even if there were an additional negative effect of having a female mayor on day care expenditures that we are not controlling for, it would mean that the effect we present in this paper is only a lower bound of the true effect of Law 215/2012 on the outcome. In addition, even if there is no perfect compliance with the third validity assumption, this would simply mean that the one we are estimating is a Local Average Treatment Effect and not an Average Treatment Effect: in this case, the effect would be valid only for the treated municipalities. Apart from the mayor's gender, no other interaction terms are significant, which supports our reasoning. Therefore, we can also be confident about compliance with the third validity assumption.[70]

B McCrary test

We want to test whether there could have been manipulation in the running variable at the threshold of 5000 inhabitants. Figure 6 shows the McCrary test, which analyzes whether there is a significant difference in observational densities at the cutoff.

[70] If the third validity assumption holds, the Difference in Discontinuity identifies the Average Treatment Effect, which can be generalized to the whole sample of observations. On the other hand, in case this assumption did not hold, we would have a Local Average Treatment Effect: the effect that we observe would be valid only for the treated municipalities.

Figure 6: **% McCrary test for manipulation of running variable**



Notes: McCrary test assessing the manipulation of the running variable. The null hypothesis is that the two observational densities, at the right and left of the 5000 inhabitants threshold, are equal. P-value for this hypothesis is 0.95, not supporting the null hypothesis' rejection. Sampled municipalities are 1424 out of 8100 Italian municipalities, observed over the period between 2013 and 2018.

The McCrary test does not reject the null hypothesis of no manipulation in the running variable, thus we can conclude that there is no reason to believe that mayors distorted their cities' population to move it either above or below the threshold.[71]

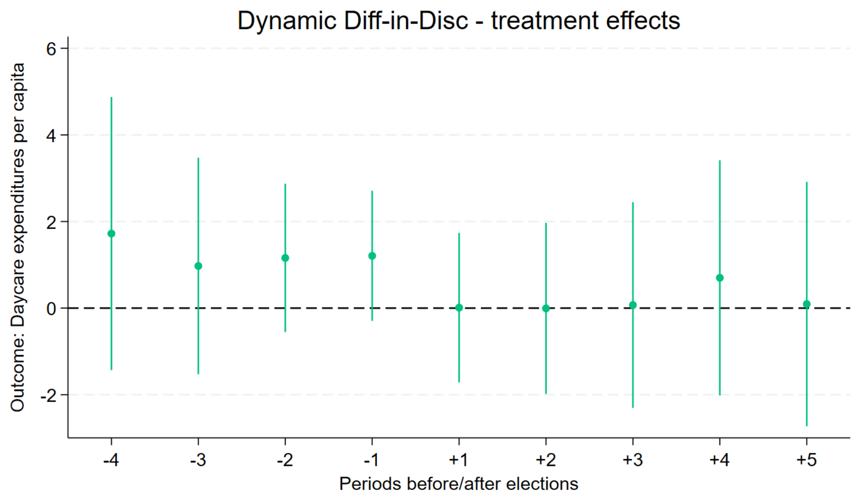
C Dynamic Difference in Discontinuity

The dynamic Difference-in-Discontinuity approach (Vannutelli 2021) used here interacts the dummy identifying the treatment group with indicators for each period observed: 1/2/3/4 years before, and 1/2/3/4 years after, the first elections under the new law. This specification allows us to observe the treatment effect over time and to conduct a placebo test by verifying whether the pre-treatment periods display null treatment effects. In other words, we are empirically assessing a parallel trends assumption. As in the main analysis, we restrict

[71]The p-value of the formal test is 0.71

the sample to municipalities within the previously computed bandwidth. The omitted time period is the election year, so all interacted coefficients measure the difference relative to this reference year.Figure 7 presents the dynamic Difference-in-Discontinuity results graphically.

Figure 7: **Dynamic Difference in Discontinuity results**



Notes: Dynamic Difference in Discontinuity results. The coefficients' size and confidence intervals shown here correspond to the variables identifying the dynamic treatment effect, obtained by interacting the dummy $Treatment$ with several dummies identifying the periods before and after the first election with Law 215/2012 in force. The dependent variable is yearly per capita expenditures on day care-related categories. Expenditures taken during election years are attributed to the council governing up to the election month. The bandwidth around the 5000 inhabitants' threshold is computed following one of the procedures described by Calonico, Cattaneo, and Farrell 2020, through the minimization of the Mean Squared Error. Year and Municipality fixed effects are included in each regression.

Looking at Figure 7, we can conclude that the aggregate treatment effect is never significantly different from zero, even when we look at it in a dynamic sense. We can also exclude differences between treated and control municipalities in political budget cycles: we do not observe recurring patterns in the periods before or after the elections. This absence of differences in spending cycles supports the idea of comparability between treated and control municipalities.

D Different level of pre-law female councillors

A potential concern we need to discuss comes from the fact that the heterogeneity we found might depend on pre-law levels of female councillors, which in turn might be driven by other municipalities' characteristics related to day care services. Even if this particular heterogeneity is captured by municipality fixed effects, we perform an additional robustness and introduce the pre-law level of female councillors as a control in our Difference in Discontinuity regressions, to exclude this potential confounder from our analysis.[72] Table 12 presents this additional evidence.

Table 12: **Difference in Discontinuity heterogeneous results (controlling for pre-law shares)**

VARIABLES	Baseline interaction Daycarepc	25^{th} percentile Daycarepc	median Daycarepc	75^{th} percentile Daycarepc	90^{th} percentile Daycarepc
TreatPost	3.554	-0.885	0.062	-0.097	-0.433
	(2.247)	(0.965)	(0.954)	(0.933)	(0.926)
TreatPost*Fem. post law	-10.506*				
	(5.808)				
TreatPost*Twenty-fifth p.		2.309**			
		(0.959)			
TreatPost*Median			-1.716*		
			(0.969)		
TreatPost*Seventy-fifth p.				-3.355	
				(2.754)	
TreatPost*Ninetieth p.					-2.604
					(3.127)
Fem. pre-law	16.157***	16.918***	18.466***	19.314***	20.059***
	(4.258)	(4.314)	(4.239)	(4.344)	(4.352)
Observations	2,473	2,494	2,494	2,494	2,494
R-squared	0.034	0.029	0.031	0.033	0.028
Number of municipalities	554	560	560	560	560
Mayor controls	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Robust standard errors in parentheses are clustered at the municipality level
*** p<0.01, ** p<0.05, * p<0.1

Notes: Difference in discontinuity regressions. TreatPost corresponds to the treatment coefficient, the interaction between the dummy identifying the treatment group and the dummy for the post-treatment period. Interaction terms indicate the post-law share of elected female councillors, taken with respect to the overall distribution of treated cities' councils. The dependent variable is yearly per capita expenditures on day care related categories. Fem. pre-law indicates the share of female councillors before the first elections with Law 215/2012 in force. Expenditures taken during election years are attributed to the council governing up to the elections month. The bandwidth around the 5000 inhabitants' threshold is computed following one of the procedures described by Calonico, Cattaneo, and Farrell 2020, through the minimization of the Mean Squared Error. Year fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

What we see from Table 12 is that our main coefficients are not modified in magnitude or significance level. Therefore, the inclusion of the share of pre-law female councillors confirms that this variable is potentially increasing the provision of day care across municipalities, but also that it should not be a source of bias for our results.

[72] In order for this control not to be absorbed by municipality fixed effects, we dropped them in the regressions in Table 12.

E Placebo tests - artificial cutoffs/year of implementation

To further validate the robustness of our results, we conduct two additional placebo analyses aimed at ruling out spurious correlations or confounding trends unrelated to Law 215/2012. These exercises test whether our observed heterogeneous effects on day care-related expenditures could emerge under arbitrary conditions unrelated to the law. In the first placebo analysis, we focus on population thresholds. Recall that the actual legal cutoff for the law's enforcement was set at 5,000 inhabitants. We artificially shift this threshold to 4,000 and 6,000 inhabitants and re-estimate our main specifications. To ensure the integrity of the test, we apply the 4,000 threshold to a sample composed only of control municipalities (i.e., those below the real threshold and unaffected by the reform), and the 6,000 threshold only to treated municipalities (i.e., those above 5,000). If our main findings were driven by generic discontinuities around population size—or by unobserved variables correlated with size—we would expect similar heterogeneous effects to emerge around these artificial thresholds.

Table 13: **Placebo test with artificial population cutoff - control municipalities**

VARIABLES	Baseline interaction Daycarepc	25^{th} percentile Daycarepc	median Daycarepc	75^{th} percentile Daycarepc	90^{th} percentile Daycarepc
TreatPost*Fem. post law	-0.487				
	(4.746)				
TreatPost*Twenty-fifth p.		1.369			
		(1.699)			
TreatPost*Median			-0.335		
			(0.817)		
TreatPost*Seventy-fifth p.				0.202	
				(0.974)	
TreatPost*Ninetieth p.					3.342**
					(1.599)
Observations	1,644	1,644	1,644	1,644	1,644
R-squared	0.020	0.021	0.020	0.020	0.025
Number of municipalities	399	399	399	399	399
Mayor controls	YES	YES	YES	YES	YES
Municipality FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Robust standard errors in parentheses are clustered at the municipality level
*** p<0.01, ** p<0.05, * p<0.1

Notes: Difference in discontinuity regressions with artificial cutoff for Law 215/2012 implementation. In this case, the sample includes only municipalities with less than 5000 inhabitants and the cutoff has been moved to 4000 inhabitants. TreatPost corresponds to the treatment coefficient, the interaction between the dummy identifying the treatment group and the dummy for the post-treatment period. Interaction terms indicate the post-law share of elected female councillors, taken with respect to the overall distribution of treated cities' councils. The dependent variable is yearly per capita expenditures on day care-related categories. Expenditures taken during election years are attributed to the council governing up to the elections' month. The bandwidth around the artificial 4000 inhabitants' threshold is computed following one of the procedures described by Calonico, Cattaneo, and Farrell 2020, through the minimization of the Mean Squared Error. Year and Municipality fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

Table 14: **Placebo test with artificial population cutoff - treatment municipalities**

VARIABLES	Baseline interaction Daycarepc	25^{th} percentile Daycarepc	median Daycarepc	75^{th} percentile Daycarepc	90^{th} percentile Daycarepc
TreatPost	1.288	-1.823	-0.827	-1.685	-1.582
	(2.111)	(1.937)	(1.595)	(1.863)	(1.829)
TreatPost*Fem. post law	-7.623				
	(7.084)				
TreatPost*Twenty-fifth p.		1.520			
		(1.229)			
TreatPost*Median			-2.235		
			(1.781)		
TreatPost*Seventy-fifth p.				1.448	
				(1.276)	
TreatPost*Ninetieth p.					-0.283
					(1.497)
Observations	722	722	722	722	722
R-squared	0.031	0.030	0.036	0.030	0.029
Number of municipalities	212	212	212	212	212
Mayor controls	YES	YES	YES	YES	YES
Municipality FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Robust standard errors in parentheses are clustered at the municipality level
*** p<0.01, ** p<0.05, * p<0.1

Notes: Difference in discontinuity regressions with artificial cutoff for Law 215/2012 implementation. In this case, the sample includes only municipalities with more than 5000 inhabitants and the cutoff has been moved to 6000 inhabitants. TreatPost corresponds to the treatment coefficient, the interaction between the dummy identifying the treatment group and the dummy for the post-treatment period. Interaction terms indicate the post-law share of elected female councillors, taken with respect to the overall distribution of treated cities' councils. The dependent variable is yearly per capita expenditures on day care-related categories. Expenditures taken during election years are attributed to the council governing up to the elections' month. The bandwidth around the artificial 6000 inhabitants' threshold is computed following one of the procedures described by Calonico, Cattaneo, and Farrell 2020, through the minimization of the Mean Squared Error. The dummy for the twenty-fifth percentile has been dropped due to multicollinearity. Year and Municipality fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

Tables 13 and 14 report the results of these placebo tests. As expected, none of the interaction terms capturing heterogeneity in female representation yields statistically significant effects. This provides strong reassurance that our original results are not due to arbitrary discontinuities in the data or pre-existing trends around municipality size.

The second placebo test examines the timing of the reform. Here, we falsely assign the law's implementation year to 2015 (rather than 2012), designating 2016–2018 as the "post-treatment" period. This test assesses whether the observed effects could be attributed to unrelated time trends, such as broader changes in child care policy or macroeconomic fluctuations. Table 15 shows that the placebo interaction effects and treatment coefficients are again statistically insignificant, reinforcing the interpretation that the actual policy reform in 2012—not coincidental year effects—drives our findings.

Table 15: **Placebo test with artificial year of law's introduction**

VARIABLES	Baseline interaction Daycarepc	25^{th} percentile Daycarepc	median Daycarepc	75^{th} percentile Daycarepc	90^{th} percentile Daycarepc
TreatPost	1.909 (1.862)	0.378 (1.088)	0.796 (1.122)	0.673 (1.111)	0.567 (1.082)
TreatPost*Fem. post law	-3.710 (4.039)				
TreatPost*Twenty-fifth p.		0.967 (0.973)			
TreatPost*Median			-0.793 (0.733)		
TreatPost*Seventy-fifth p.				-1.233 (1.895)	
TreatPost*Ninetieth p.					-1.086 (2.034)
Observations	2,548	2,548	2,548	2,548	2,548
R-squared	0.030	0.030	0.030	0.031	0.030
Number of municipalities	570	570	570	570	570
Mayor controls	YES	YES	YES	YES	YES
Municipality FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Robust standard errors in parentheses are clustered at the municipality level
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Notes: Difference in discontinuity regressions with artificial year of introduction for Law 215/2012. In this case, we consider as councils governing under the new legislation those elected after 2015. TreatPost corresponds to the treatment coefficient, the interaction between the dummy identifying the treatment group and the dummy for the post-treatment period. Interaction terms indicate the post-law share of elected female councillors, taken with respect to the overall distribution of treated cities' councils. The dependent variable is yearly per capita expenditures on day care-related categories. Expenditures taken on election years are attributed to the council governing up to the elections' month. The bandwidth around the 5000 inhabitants' threshold is computed following one of the procedures described by Calonico, Cattaneo, and Farrell 2020, through the minimization of the Mean Squared Error. Year and Municipality fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

Taken together, these placebo analyses support the internal validity of our empirical strategy. The absence of significant effects under artificial conditions strengthens our claim that Law 215/2012 is responsible for the heterogeneous impacts we document across municipalities.

Alma Mater Studiorum - Università di Bologna
DEPARTMENT OF ECONOMICS

Strada Maggiore 45
40125 Bologna - Italy
Tel. +39 051 2092604
Fax +39 051 2092664
http://www.dse.unibo.it