Political power and the influence of minorities: theory and evidence from Italy

Giovanni Righetto[1]

February 27, 2024

Abstract

In this paper, we analyze the relationship between minority and majority in politics, and how it can influence policy outcomes and potential conflict between parties. In particular, we focus on the consequences of a sudden increase in the political power of a minority (e.g. female politicians after a gender quota), and its potential effects on the relationship with the long-standing majority. We first present a theoretical model describing the possible consequences of such an increase in a minority's political power and show how it can increase difficulties in reaching a compromise on policy outcomes between parties. In the case of a high increase in minority's power, its demands in terms of policy outcomes increase and make the compromise costlier for the majority, which might prefer to engage in conflict. Furthermore, we empirically test these implications by exploiting the introduction in 2012 of a gender quota in Italian local elections. By means of a Difference-in-Discontinuity strategy, we show how the generated increase in female politicians had heterogeneous effects on the level of funding for daycare, based on its differential effects on the share of post-quota women councillors. For high shares of female councillors, a decrease in expenditure for day care was observed with respect to control municipalities, while in municipalities with low shares the quota was followed by a relative increase in funding for day care.
JEL classification numbers: D71, H53, I38
Keywords: Conflict, committee, gender quotas, day care investments

[1] Post-Doctoral researcher, University of Milan. Email: giovannirighetto7@gmail.com

I would like to thank Paolo Vanin, my PhD supervisor. I am also grateful to participants in all conferences and seminars I attended for all the comments and constructive feedback. In addition, I acknowledge the financial support of the University of Bologna and the University of Milan.

1 Introduction

During the last two centuries, the democratization process in many countries has helped the enfranchisement of social groups which were still unable to influence the political process of decision-making. We can think for instance of the extension of suffrage to women or Afro-Americans or the recent introduction of gender quotas in politics both in developed and developing countries. The enfranchisement process has been demonstrated to benefit the under-represented social groups in the long run from many different perspectives such as employment (Aneja and Avenancio-Leon (2019)), public investments (Miller (2008)), and political responsiveness to these groups' needs (Fujiwara (2015)). On the other hand, after a minority receives larger political power there might be some other mechanisms triggered in the short run, depending on how the long-standing majority reacts to this sudden decrease in its capacity to influence the political decision-making process. For instance, while a majority group might accept to "give in" to minority's instances when the balances of power are disproportionate, it might feel threatened when the minority gains more power and demands more favourable treatment in terms of policy decisions. Indeed, growth in the representation of minorities in the political class could encourage them to increase their political demands, thus making it more difficult from the point of view of the long-standing majority to accept a compromise. In this paper, we first try to model this mechanism by sketching the preferences and behaviours of two groups of individuals in a commitee, a majority and a minority group, with possible changes in the balances of political power: think for instance of female politicians after the introduction of a gender quota. The two groups of individuals have to decide on the level of a policy: in the sequential game, the majority decides on a level, and then the minority chooses whether to engage in political conflict or not. Conflict can be won by each of the two groups with a probability proportional to the group's size. The chances of either conflict or compromise are influenced by the relative groups' shares: when the outgroup share grows, also its respective demands in terms of the implemented policy level move closer to its bliss point. In this setting, a shock in the minority's political power increases the probability that the majority decides not to compromise on policy decisions, since minority claims are too expensive. In the case of no compromise, the majority decides to propose its bliss point on the policy level, with the minority that might end up being worse off after the increase in representation. We test these predictions in our empirical setting by exploiting the introduction in 2012 of a new law on Italian local elections, which increased the presence of female local politicians in a subsample of small municipalities. By the means of a Difference in Discontinuity strategy, we look at the effect of this sudden increase in female politicians' representation levels and subsequently on day care-related expenditures, possibly the most "gender-sensitive" among all kinds of policies. We find an extremely interesting result: in line with our argument, the effect on day care expenditures depended on the post-law share of female councillors in treated municipalities' councils. While in municipalities with low shares of female councillors we observe a positive change in funding for day care by 9.5% with respect to the control group, in cities with large increases in female politicians, an 18% reduction in this kind of expenditure followed the introduction of the law, always in relation to control municipalities. Thus, this increase in a minority's political representation had a heterogeneous effect depending on whether the following increase in the minority group's share was such to prevent compromise with the majority. We deepen the analysis of this effect by describing differences between cities with high and low shares of post-law female councillors, to understand whether there might have been other structural elements driving our results: however, empirical evidence shows how our heterogeneous effect indeed depended on the law-induced increase in female politicians. One of our major robustness consists in assessing the heterogeneous new law's effect on other expenditure categories: in line with our reasoning, we observe a significant heterogeneous effect only on day care and not on non-gender-related categories such as transportation or tourism. This last piece of evidence is in line with our model in two different ways: first, because day care is a service for which men and women have distinct preferences (Barigozzi et al. (2019), Baskaran and Hessami (2023)), and second because gender quotas increased the size of women group of politicians, thus raising the probability of conflict. In addition, to support our empirical results, we conduct other robustness checks: first, a placebo analysis with artificial timing/threshold of the quotas' introduction to show that we are not capturing spurious correlations. Moreover, to reinforce our argument that the increase in women politicians was responsible for this heterogeneous effect, we adopt another identification strategy to demonstrate our main result. By using a 2SLS panel regression, we show how the exogenous increase in female councillors caused by the quota was followed by a decrease in day care-related expenses in the subsample of treated municipalities. Finally, we add some evidence to support our idea that the increase in the minority group size triggered political conflict. Indeed, in the municipalities with larger post-quota shares of female councillors, we also observe an increase in the probability of dissolution of the local council.

The paper is organized as follows: in Section 2 we describe the literature that we refer to, together with our contribution. After that, in Section 3 we present our theoretical model highlighting how group shares shape relationships and decision outcomes among committee members. In Section 4 we describe the new Law introduced in Italian local elections, which caused an increase in female politicians starting in 2012. In Section 5 we describe the data used and our empirical strategy, while in Section 6 we show our main results. Section 7 conducts a set of robustness checks together with additional evidence supporting our empirical findings. Finally, Section 8 concludes.

2 Literature contribution

According to the literature on increasing minorities' and politically under-represented social groups' rights, the existing evidence found a variety of positive long-term benefits, not only exclusive to these groups. For instance, women's suffrage in the US increased public health expenditures and helped the promotion of a large public hygiene campaign, leading to a reduction in child mortality (Miller (2008)). Moreover, the enfranchisement of the less educated electorate through the introduction of the electronic vote increased public health care spending in Brazil, facilitating in particular uneducated mothers' access to prenatal visits (Fujiwara (2015)). Another example is provided by the passing of the 1965 Voting Rights Act in the US which, by increasing the voting rights of racial minorities, reduced the wage gap between black and white workers by 5.5 pp between 1950 and 1980 (Aneja and Avenancio-Leon (2019)), mainly because of major employment of black people in the public sector. However, our main idea is that a positive shock in the rights of an under-represented group can have a short-term effect on group dynamics, leading to reduced margins of compromise between majority and minority. In this regard, our paper speaks to the strand of literature on conflicts and identity. The main paper we refer to is the work by Bonomi et al. (2021), which describes how social groups align according to their identities, generating a polarization of beliefs and conflicting views on policies' effects. An external shock, such as a migration wave or globalization, can change the dimension of conflict, for instance from an economic to a cultural conflict, leading people to align on two factions depending on their views. In this paper, we adapt their intuition to a simpler model with only one identification dimension, in which the reshaping of groups' shares can possibly trigger a political conflict. This intuition is also present in the recent work by Grossman and Helpman (2021), which shows how changes in the identification patterns of voters can alter their policy preferences. Furthermore, in this paper, we also make reference to the studies on polarization by Esteban and Ray (Esteban and Ray (1994), Esteban and Ray (1999)), who described how polarization can be measured and how it might generate conflict in society. One of their major conclusions is that intra-group homogeneity and inter-group heterogeneity raise polarization levels in society. Here, our contribution is to offer a novel view of how altering the sizes of political groups can influence polarization and conflict, and ultimately even impair the investment in a publicly-funded good. Moving to our empirical part, when we focus on the potential effects of increasing the share of female politicians on the kind of policies implemented, the literature offers mixed conclusions. The theoretical predictions in these cases are not univocal: the median voter theory (Downs et al. (1957)) states that the personal characteristics of elected politicians do not matter for policy making, while the citizen-candidate model (Besley and Coate (1997)) reaches the conclusion that individual politicians' preferences matter for the kind of policies implemented. In some empirical works, the citizen-candidate model has been proven to be more plausible: for instance, there is evidence that in the US policies differ depending on whether a Republican or a Democrat wins the elections (Besley and Case (1995)). Consequently, if we think of gender quotas, we might suspect that such a sudden change in the political class' composition might have consequences in terms of policies.

Looking at the empirical literature closest to this paper, many works have exploited the exogenous imposition of gender quotas to understand whether changing the sex ratio of politicians in charge might have a causal effect on policy choices. An interesting setting in this sense is offered by the Indian mandate reservation system: since 1993, a minimum of one-third of seats plus the leadership position in randomly selected villages must be reserved for women. Policy and social consequences of this increase in female politicians have been wide, from a better representation of the policy preferences of the female electorate (Chattopadhyay and Duflo (2004)), with a larger expenditure on public goods, to a stronger reduction in the gender gap in school attendance and educational attainment (Beaman et al. (2012)). In addition, there is evidence that Indian female politicians invest more in education (Clots-Figueras (2012)) and public health infrastructures(Bhalotra and Clots-Figueras (2014)). Lastly, the generated increase in prenatal and childcare services was found to significantly reduce infant mortality (Bhalotra and Clots-Figueras (2014)). However, all these effects are not necessarily valid for other contexts: indeed, many other studies found no consequences of exogenous increases in female politicians. For instance, Bagues and Campa (2021) study the introduction of a gender quota in Spanish local elections and find no effects on either the composition or size of public expenditures. Also in Norway, the exogenous increase of female politicians due to a gender quota did not alter how local administrators were using public funds (Geys and Sørensen (2019)). Other researchers exploit close mixed-gender races to study the effects of the election of a female mayor on policy choices. Also in this case, empirical results offered by the literature are mixed. For instance, it was found that in the US electing a female mayor did not affect the size and composition of public expenditures (Ferreira and Gyourko (2014)). On the other hand, a female victory in the Bavarian local mixed-gender race was demonstrated to lead to a public child care expansion between 40% and 50% (Hessami and Baskaran (2019)). Thus, the empirical literature still did not find a consensus on whether an increase in female politicians can alter policy decisions. Hence, our contribution in this sense is to underline the importance of evaluating the post-quota interaction between male and female politicians and their respective group shares, since they can generate different policy outcomes. In other words, we demonstrate in this paper how looking at heterogeneous instead of overall effects can completely change the answer to this important question.

3 The model

We model a sequential game in which a committee has to decide the level of a policy. The committee is composed of members who belong to either one of two groups $G = \{M, F\}$: the groups' shares are $(\pi, 1-\pi)$, with the minority group F having a share equal to $\pi \leq 1/2$. The game is structured as follows:

1. The majority M proposes a level of the policy q_P to implement, with $q_P \in \mathbb{R}$
2. The minority F decides whether to engage in political conflict or not (cases $J \in \{C, N\}$)

Conflict changes the expected policy outcomes, but it generates costs for the minority group, which we interpret as "mobilization" costs. In the absence of conflict, the majority can implement the policy level that maximizes its utility: we call this implemented policy level q_I, with $q_I \in \mathbb{R}$[2].

Preferences over the implemented policy level are given by the quadratic loss function $-\frac{1}{2}(q_I - q_G)^2$, where q_G is group G's bliss point. We assume without loss of generality that $q_F > q_M$, and we normalize $q_M = 0$, so that we can interpret q_F as the distance in preferences. Furthermore, we make the assumption that in case of conflict, this is won with probability π by the minority, which then can implement its desired policy level. On the other hand, if the majority wins, it can impose its proposed level q_P. Therefore, conflict changes the expected outcome in this way:

$$E(q_I|C) = (1-\pi)q_P + \pi q_F$$

In other words, conflict makes the expected implemented policy level equal to a weighted average of the majority's proposal q_P and the minority's bliss point q_F, and the weights are the two groups' sizes. Therefore, growing a group size increases the chances that it wins the conflict and that it can impose its most preferred policy level. On the other hand, conflict entails also a cost for the minority that we assume to be fixed and equal to γ. Therefore,

[2] Thus, in case of conflict, $q_I = q_P$

the utility U_G of group G depends on the policy level implemented and on the minority's decision, which we call $S_F(q)$ with $S_F(q) \in \{N, C\}$. We have that the two groups' utilities are defined in the following way:

$$U_M(q_I, J) = -\tfrac{1}{2}(q_I - q_M)^2 = -\tfrac{1}{2}(q_I)^2$$
$$U_F(q_I, J) = -\tfrac{1}{2}(q_I - q_F)^2 - \gamma 1(J = C)$$

On the other hand, expected utilities are defined by the two groups' strategies $(q_P, S_F(q))$:

$$E(U_M(q_P, S_F(\cdot))) =$$
$$= -\tfrac{1}{2}1(S_F(q_P) = N)(q_P - q_M)^2 - \tfrac{1}{2}1(S_F(q_P) = C)[(1-\pi)(q_P - q_M)^2 + (\pi)(q_F - q_M)^2] =$$
$$= -\tfrac{1}{2}1(S_F(q_P) = N)(q_P)^2 - \tfrac{1}{2}1(S_F(q_P) = C)[(1-\pi)(q_P)^2 + (\pi)(q_F)^2]$$
$$E(U_F(q_P, S_F(\cdot))) =$$
$$= -\tfrac{1}{2}1(S_F(q_P) = N)(q_P - q_F)^2 - \tfrac{1}{2}1(S_F(q_P) = C)[(1-\pi)(q_P - q_F)^2 + (\pi)(q_F - q_F)^2 - \gamma] =$$
$$= -\tfrac{1}{2}1(S_F(q_P) = N)(q_P - q_F)^2 - \tfrac{1}{2}1(S_F(q_P) = C)[(1-\pi)(q_P - q_F)^2 - \gamma]$$

We solve the game by backward induction, starting from the best response $S_F^*(q_P)$ of F and then moving to the optimal level q^* of q_P proposed by M. F's optimal strategy is to choose conflict if $E(U_F(q_P, S_F(\cdot)|S_F(q_P) = N)) < E(U_F(q_P, S_F(\cdot)|S_F(q_P) = C))$, that is

$$S_F^*(q_P) = C \iff q_P < q_F - \sqrt{\frac{2\gamma}{\pi}} = \overline{q}_M \tag{1}$$

The intuition here is that when the majority proposes a policy level q_P lower than a certain threshold $\overline{q}_M$, for the minority it is convenient to engage in conflict. Indeed, conflict has the potential to change the policy outcome to the minority's bliss point q_F with probability π. We focus on the case in which $q_P < q_F$, as the case $q_P > q_F$ does not happen in equilibrium[3]. Condition (1) tells us that the majority needs to offer at least $\overline{q}_M$ to the minority in order to

[3] The case $q^* > q_F$ cannot happen in equilibrium. To see this, notice that $\bar{q}_M < q_F$ and $E(U_M(\bar{q}_M, N)) > E(U_M(q_I, N)) \forall q_I > q_F$, and in case of conflict $q* = q_M < q_F$

avoid conflict, and that this policy level is increasing in the minority share π and decreasing in the cost of conflict γ.

Moreover, we define q^*, the optimal level of policy proposal of M, as

$$q^* = argmax_{q_P \in \mathbb{R}} E(U_M(q_P, S_F^*(q)))$$

We have that in case of conflict $q^* = q_M = 0$ since the policy level maximizing majority's utility is its bliss point. In terms of the q^* chosen by the majority, we can have two cases, depending on the relative position of $\overline{q}_M$ with respect to q_M.

1. Suppose $\overline{q}_M < q_M$[4], or $q_F < \sqrt{\frac{2\gamma}{\pi}}$: in this case, we have that the bliss point of the majority is able to prevent the conflict

   $$q^* = q_M$$

   This is the case in which the distance in preferences is sufficiently narrow to make it not profitable for the minority to trigger conflict. Indeed, the expected outcome from conflict would be close to q_M, and the minority would also have to bear the costs γ. The majority chooses its desired policy level, and conflict is avoided.

2. On the other hand, suppose that $\overline{q}_M > q_M$[5], which can be written as $q_F > \sqrt{\frac{2\gamma}{\pi}}$: here, we have a larger difference in preferences and proposing the bliss point of the majority does not avoid conflict. In this case, the majority must decide whether to choose $\overline{q}_M$, the minimal policy level that makes the minority indifferent between engaging in the conflict or not, or to propose q_M and trigger conflict with the minority group. Thus, the majority prefers to engage in conflict in the case that $E(U_M(q_M, C)) > E(U_M(\bar{q}_M, N))$, which is when:

   $$q_F > \frac{\sqrt{2\gamma}}{\sqrt{\pi} - \pi} \tag{2}$$

[4] This can be written as $\pi < \frac{2\gamma}{q_F^2}$. If $\gamma > (\frac{q_F}{2})^2$, this holds $\forall \pi < \frac{1}{2}$ and conflict never happens. We assume $\gamma < (\frac{q_F}{2})^2$ in order to focus on the more interesting case in which conflict is possible.

[5] This means $\pi > \frac{2\gamma}{q_F^2}$

Condition (2) can be written with respect to π, translating it into a range of values for which we have conflict: if $\pi \in [\pi_1; \pi_2]$, conflict happens and $E(q_I) = \pi q_F$[6]. Here, for the sake of exposition, we focus on the case in which $q_F > 2\sqrt{2\gamma} + 4\sqrt{\gamma}$, so that $\pi_2 > \frac{1}{2}$ and the case $\pi > \pi_2$ does not happen. We describe in Appendix A the case $q_F < 2\sqrt{2\gamma}+4\sqrt{\gamma}$ [7]. Therefore, depending on the level of π we can have either conflict or compromise between the two groups: in the first case, the expected outcome is a function of the minority group's bliss point, while in the second case we have an outcome closer to the one preferred by the minority.

Therefore, we can formalize the equilibrium strategy chosen by the majority:

$$q^* = \begin{cases} 0, & \text{if } q_F < \sqrt{\frac{2\gamma}{\pi}} \vee q_F > \frac{\sqrt{2\gamma}}{\sqrt{\pi}-\pi} \\ \overline{q}_M, & \text{if } q_F \in [\sqrt{\frac{2\gamma}{\pi}}; \frac{\sqrt{2\gamma}}{\sqrt{\pi}-\pi}] \end{cases}$$

Thus, we have that $q^* \in [0; \overline{q}_M)$. On the other hand, the equilibrium strategy decided by the minority is formalized as follows:

$$S_F^*(q^*) = \begin{cases} C, & \text{if } q_F > \frac{\sqrt{2\gamma}}{\sqrt{\pi}-\pi} \\ N, & \text{if } q_F < \frac{\sqrt{2\gamma}}{\sqrt{\pi}-\pi} \end{cases}$$

Here, we can see how conflict happens when there is a higher level of the minority group's bliss point (which is the distance in preferences) and a larger minority's share. Conversely, higher costs of conflict for the minority raise the threshold that q_F has to cross in order to trigger conflict.

3.1 Equilibrium and comparative statics

We can have three possible equilibrium outcomes in terms of the implemented policy level and of the conflict/no conflict situation. We summarize them in the following proposition:

[6] Formally, we have that $\pi_1 = \frac{q_F - 2\sqrt{2\gamma} - \sqrt{q_F^2 - 4\sqrt{2\gamma} q_F}}{2q_F}$ and $\pi_2 = \frac{q_F - 2\sqrt{2\gamma} + \sqrt{q_F^2 - 4\sqrt{2\gamma} q_F}}{2q_F}$.

[7] The assumption $q_F > 2\sqrt{2\gamma}+4\sqrt{\gamma}$ does not prevent the possibility for q_F to be $\lesseqgtr \frac{\sqrt{2\gamma}}{\sqrt{\pi}-\pi}$. This assumption is also compatible with the situations in which $q_F \lesseqgtr \sqrt{\frac{2\gamma}{\pi}}$.

Proposition 1. *According to the difference in preferences, the equilibrium in terms of the two groups' strategies and policy outcomes can be outlined as follows:*

1. *With large difference in preferences, which is when $q_F > \frac{\sqrt{2\gamma}}{\sqrt{\bar{\pi}-\pi}}$, the majority proposes its bliss point $q_M = 0$ and conflict happens. In this case, we have that $E(q_I|C) = \pi q_F$.*

2. *If the two groups have close preferences, which happens when $q_F < \sqrt{\frac{2\gamma}{\pi}}$, the majority proposes its bliss point $q_M = 0$ and there is no conflict.*

3. *In the case of intermediate distance in preferences, formally when $q_F \in [\sqrt{\frac{2\gamma}{\pi}}; \frac{\sqrt{2\gamma}}{\sqrt{\bar{\pi}-\pi}}]$, the majority chooses a policy level closer to the minority's preferences and conflict is avoided: the proposed, and then implemented, level is $q^* = q_F - \sqrt{\frac{2\gamma}{\pi}} = \overline{q}_M$.*

Briefly, the intuition is that growth in the minority's share increases the group's requests in terms of policy[8]: if these requests are too expensive for the majority, it might decide to propose its bliss point and trigger conflict. The result, in this case, is a worse policy outcome from the minority's perspective: indeed, when $\pi = \pi_1$, we have that $\bar{q}_M > E(q_I, C)$. From this, it follows that triggering the political conflict decreases the utility level of the minority, since at the turning point $\pi = \pi_1$ the minority has to bear the cost of conflict and receives a lower expected policy level:

$$E(U_F(\bar{q}_M, N)) > E(U_F(q_I, C))$$

Overall, looking at Proposition 1, we have that, given π and q_F, an increase in the cost of conflict γ reduces the range of values of q_F for which there is conflict, making it less likely. On the other hand, if we fix π and γ, an increase in the distance in preferences reduces π_1, increasing the range of values of π for which there is conflict. In addition, we can make some comparative statics of the expected policy level implemented with respect to the minority share's π as follows:

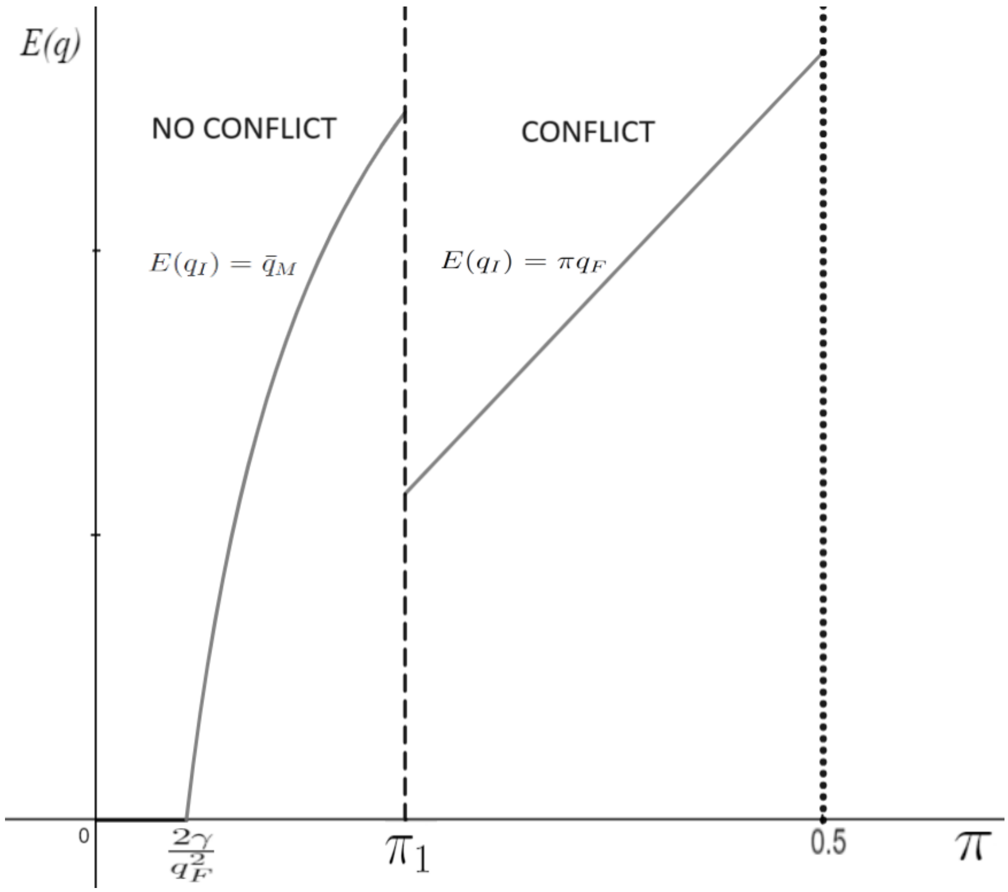
- Up to $\pi = \frac{2\gamma}{q_F^2}$, $E(q_I) = 0$
- For $\pi \in [\frac{2\gamma}{q_F^2}; \pi_1]$, we have that $E(q_I) = \overline{q}_M$, so the expected outcome is growing in π

[8] $\overline{q}_M$ moves closer to q_F

- At $\pi = \pi_1$, we have a drop in $E(q_I)$, which then starts growing again for $\pi > \pi_1$. Here, we have that $E(q_I) = \pi q_F$

Graphically, we plot the evolution of $E(q_I)$ with respect to π in Fig.1.

Fig. 1: **Evolution of $E(q_I)$ with respect to the minority share**



Looking at Fig.1 we can see how $E(q_I)$ is initially flat in π, then monotonically increasing up to $\pi = \pi_1$, where there is a drop and the conflict is triggered. For larger minority shares, $E(q_I)$ starts increasing again[9].

Now, we might wonder why the minority decides to start the conflict even if this decreases its utility, instead of moderating its policy requests. The answer lies in the dynamic structure of the game, where the majority moves first and then the answer of the minority follows.

[9] There is a global maximum of $E(q_I)$ at $\pi = 0.5$. Indeed, we have that $E(q_I)$ at $\pi = \pi_1$ is larger than $E(q_I)$ at $\pi = 0.5$ for $q_F \in [0; 4\sqrt{2\gamma}]$. This condition is never satisfied for the range of values on which we focus for the distance in preferences, which is $q_F > 2\sqrt{2\gamma} + 4\sqrt{\gamma}$.

Thus, any eventual promise of the minority is not completely credible without the existence of commitment devices. There is no reason for the majority to believe that the minority will avoid conflict, and consequently, they are not incentivized to change their policy proposals. The most interesting finding concerning the whole paper's main argument is that an increase in the share of the minority can trigger conflict between the two groups and potentially decrease the implemented policy level. This is because when the minority gains more political power, its requests might become too expensive to accommodate for the majority.

4 Setting: the 2012 gender quota

An example of a situation resembling our model's dynamic could be the introduction of a gender quota increasing the share of female members: this can be followed by increased demand for "gender-sensitive" policies[10]. These requests might become too exorbitant from the point of view of male members, who could prefer to establish their preferred policy level in terms of "gender-sensitive" policies, even if this means potentially starting a conflict with other members. This is our main result: increasing the political power of a minority has the potential to damage this group in terms of policy outcomes and utility, at least in the short term. In this section, we show how this theoretical finding can be confirmed empirically, by looking at the consequences of the introduction of a gender quota in Italy on the local expenditures for day care, possibly the most "gender-sensitive" among all categories of public expenditures.

In Italy, there are approximately 8100 municipalities, which represent the lowest sub-national level, after regions and provinces. Each municipality has a mayor assisted by a local council ("Consiglio comunale"), which owns the legislative power, and by an executive committee ("Giunta comunale") owning the executive power. The local administrators decide on the allocation of public funds over a large variety of categories since the provision of many public services is decentralized at the local level. There are three sources of financing for a

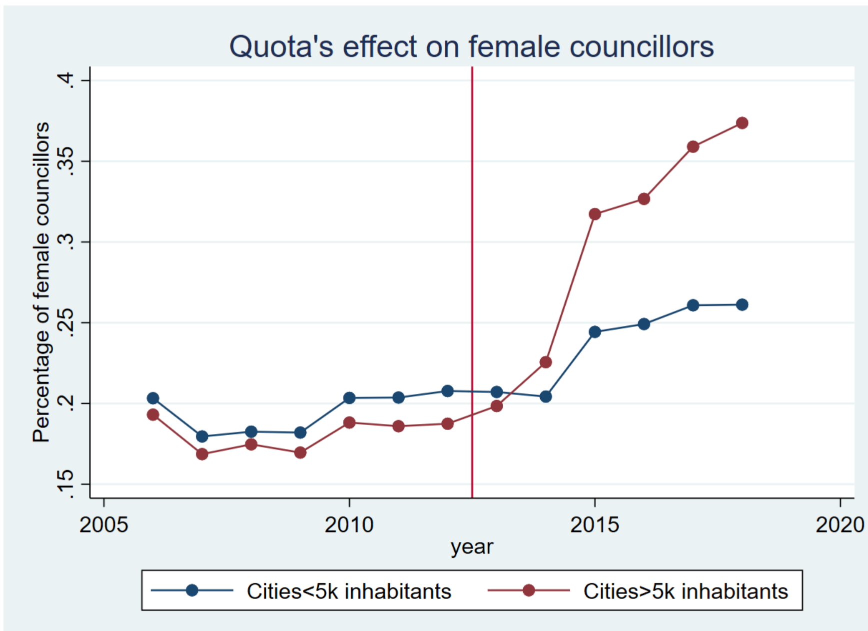
[10] We could also think of an effect similar to the one in Bonomi et al. (2021), meaning that when the gender-related topics become more salient in the political debate, this would increase the difference in preferences q_F because of polarization of the two groups' positions. This additional effect would simply reinforce our main results.

municipality: own taxes and tariffs, transfers from the central government, and revenues from fines. Elections are held every five years, and voters can express their preference for both the mayor and local councillors. We focus here on municipalities between 2000 and 10000 inhabitants, given that outside this interval there might be different electoral rules: these cities constitute a relatively homogeneous sample that allows the implementation of our identification strategy, as we explain in Section 6. For these cities, citizen over 18 years old 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. A total of 12 councillors are elected through this system.

In this paper, we focus on Law 215/2012, which introduced two major changes for local elections across Italian municipalities with more than 5000 inhabitants. First, a gender quota was imposed on candidates' lists: a maximum of two-thirds of same-sex candidates for local councillors can be included. In addition, a second important novelty was introduced regarding the voting system: citizens could cast two votes for councillors instead of one with the condition that the two candidates were of different sex. The clear aim of the law was to increase women's presence among Italian local politicians, in order to reduce the country's gender gap in politics. According to this law, non-complying parties were punished with the removal of same-sex candidates exceeding two-thirds of the total in their lists.

The consequences of this law in terms of effects on councils' composition were studied by Baltrunaite et al. (2019), who, by the means of a regression discontinuity design, found that the law increased the percentage of female councillors by an average of 18 percentage points. They claim that this effect was mainly driven by the introduction of the double vote for councillors, since lists' compositions were not particularly affected. Moreover, they investigate whether the law provoked some additional effects in terms of elected candidates' characteristics, namely age, years of education or previous occupation: none of these characteristics seemed to have changed, not even focusing on same-sex candidates. In other words, the only effect of the law was to increase the percentage of elected female councillors, and not to alter other elected councillors' characteristics[11]. In Fig.2 we can see how the percentage of female councillors changed over the years for our municipalities of focus, namely those between 2000 and 10000 inhabitants.

Fig. 2: % **Female councillors**



Note: Percentage of female councillors over the years in Italian municipalities with a population between 2000 and 10000 inhabitants. The red vertical line indicates the time when the law 215/2012 was introduced. Note that elections are staggered across municipalities, thus different municipalities might have had different election years. Data include 1434 out of 8100 Italian municipalities, observed between 2006 and 2018.

From Fig.2 we can notice how in municipalities interested by Law 215/2012 there was a sharp increase in female councillors, especially after 2013. On average, the percentage of female councillors passed from 18% in the pre-quota period to almost 38% in 2018 for the interested municipalities. It is important to point out that elections are staggered across municipalities, thus different municipalities might 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 interested by the quota held elections in that year, thus in 2013 their councillors were still the ones chosen with the old set of rules. We also need to specify that,

[11] The law's effects in terms of turnout or of turnout by gender were also negligible

even after the imposition of the new law, in almost all municipalities (93% of our sample) men detained the majority, which is required in order to pass motions on public expenditures: in other words, political power was still in the hands of male councillors. Moreover, we notice how there seemed to have been a spillover effect also for municipalities not interested by the quota, which also increased their shares of female councillors. This spillover might affect the interpretation of our empirical results in the sense that what we show in Section 6 is possibly only a lower-bound effect.

5 Data

We collected data on a total of 1422 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. 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"[12], that in 2013 was extended to all municipalities. 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, regards 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 mean age and years of education of the councillors in charge, together with the percentage of female councillors. In Tab.1 we present some summary statistics for councillors, first for the whole sample and then for the treatment and control groups.

[12] 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 to municipalities, while compliance granted a reduction in interest expenses for government loans.

Tab. 1: **Summary statistics for municipalities' councils**

Whole sample

Variable	Obs	Mean	Std.Dev.	Min	Max
Mean share of female councillors	8,291	0.26	0.14	0	0.76
Councillors' mean years of education	8,064	13.93	1.37	8	18
Councillors' mean age	8,076	51.75	4.66	25	70.25

Treatment group

Variable	Obs	Mean	Std.Dev.	Min	Max
Mean share of female councillors	3,494	0.30	0.15	0	0.76
Councillors' mean years of education	3,409	14.19	1.37	8	18
Councillors' mean age	3,412	51.66	4.66	38	69

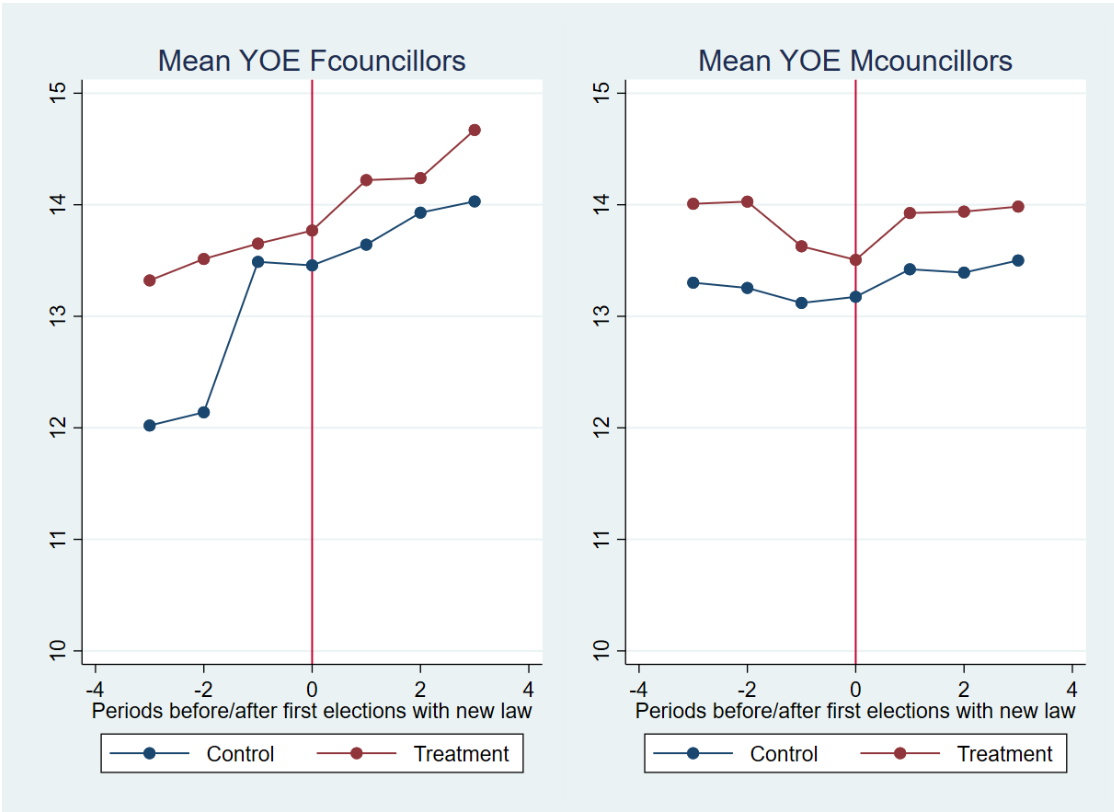
Control group

Variable	Obs	Mean	Std.Dev.	Min	Max
Mean share of female councillors	4,797	0.23	0.13	0	0.69
Councillors' mean years of education	4,655	13.74	1.34	8	18
Councillors' mean age	4,664	51.81	4.65	25	70.25

Notes: Summary statistics for sampled municipalities' councils, namely councillors' mean age, councillors' mean years of education and mean share of female councillors. The treatment group includes sampled cities with population above 5000 inhabitants, while the control group includes those below. Sampled municipalities are 1422 out of 8100 Italian municipalities, observed over the period between 2013 and 2018.

As we can observe from Tab.1, there are no striking differences between the two groups of councillors, apart from a larger female presence (due also to Law 215/2012) and a higher educational level in the treatment group. It is worth repeating that existing evidence showed that the quota did not change the educational level or average age of either male or female councillors (Baltrunaite et al. (2019)). Indeed, we can also notice this graphically: Fig.3 shows the evolution of the average educational level for male and female councillors by year, before and after the first elections under the new law. As we can see, no particular different trends appear to have been produced by Law 215/2012 for this particular variable.

Fig. 3: **% Trends in educational levels**

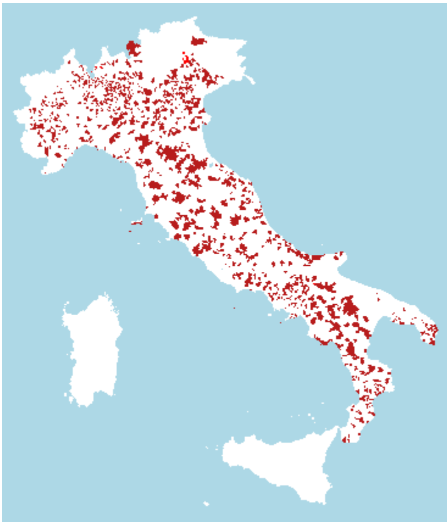


Note: Trends in average educational levels for male and female councillors. The red vertical line indicates the time when a municipality held its first elections under the law 215/2012, while the x-axis indicates the years before and after this time. The treatment group includes municipalities that were subject to the new law, while the control group the remaining ones. Data include 1434 out of 8100 Italian municipalities, observed between 2006 and 2018.

Moreover, we need to point out that most of these politicians do not have a clear ideological leaning: 95% of them belong to the so-called "liste civiche" (civic lists), which are parties not possible to locate on the right-left axis. Therefore, our sampled politicians are not influenced by any central party's guidelines, and are freer to vote on their preferred policies with respect to politicians elected at higher levels (e.g. the national Parliament or regional councils).

In Fig.4 we present the geographical distribution of sampled and treated municipalities: these two maps give us confidence that there is no geographical bias neither in our sampling or assignment to treatment procedure.

Fig. 4: **Sampled and treated municipalities**



(a) Sampled municipalities

(b) Treated municipalities

0.80.8 *Notes*: Fig.4a shows all municipalities belonging to our sample, while Fig.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 1422 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.

Together with the dataset on local politicians, we collected data on municipalities' expenditures from the Istat website. 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 a quota's effect: day care expenditures. Italy is historically conservative in its family models, and traditionally childcare is performed by mothers and grandmothers (Brilli et al. (2016)). This is due to longstanding social norms on gender roles, that privilege informal child care provided by female family members with respect to formal services (Barigozzi et al. (2019)). As a consequence, both private and public day care is less provided with respect to other European countries (Brilli et al. (2016)), with public preschool investment being one of the lowest in Europe. Public investments in day care have been growing in the last decades, but Italy still remains below the European average for the percentage of GDP devoted to child care and pre-primary schools (Barigozzi et al. (2019)). We decide to focus on this particular spending category because we have evidence from the literature that female politicians are more inclined to invest in day care (Bhalotra and Clots-Figueras (2014), Miller (2008)) with respect to male politicians. Having subsidized day care helps working mothers to cope with their job duties, increasing maternal employment and generating also possible positive effects for children's well-being (Herbst (2017), Brilli et al. (2016), Miller (2008)). Thus, we might expect that the substantial increase in female councillors had a positive effect on the portion of public funds devoted to childcare. This is why we collected data on local expenditures on childcare, together with the number of available spots and users by municipality and year. The main dependent variable of our analysis, day care expenditures per capita, is resulting from three expenditure categories added up, with the total divided by the municipality's population:

- Municipality's expenditure for directly managed day care facilities
- Municipality's funding to privately managed day care facilities[13]
- Municipality's funding for day care-related expenses. This includes all subsidies to families for day care integrated services or other related expenses, for instance the "babysitter bonus"

All these three levels of expenditures are decided by local administrators and have to be approved by the local council[14]. Here, when we speak of day care facilities, we refer to nurseries for early childhood assistance, which in Italy are meant for children aged between 0 and 3. In Tab.2 we show some summary statistics for this variable, again for the whole sample and then for treated and control municipalities.

[13] Local councils can decide to directly provide this service or outsource it to privates

[14] There might be occasional state-level contributions to daycare. For instance, national law 232 of the 11th 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.

Tab. 2: **Municipalities' day care expenditures per capita**

Whole sample

Variable	Obs	Mean	Std.Dev.	Min	Max
Day care exp. pc	8,388	7.22	12.50	0	112.95

Treatment group

Variable	Obs	Mean	Std.Dev.	Min	Max
Day care exp. pc	3,519	8.94	13.90	0	83.94

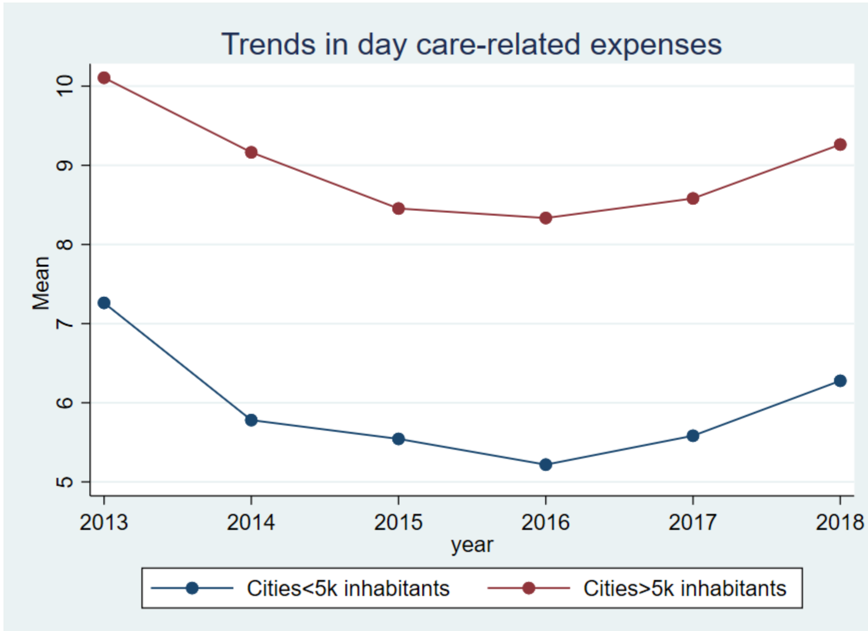
Control group

Variable	Obs	Mean	Std.Dev.	Min	Max
Day care exp. pc	4,869	5.97	11.22	0	112.95

Notes: Summary statistics for day care expenditure per capita at the municipality level. 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 a population above 5000 inhabitants, while the control group includes those below. Sampled municipalities are 1422 out of 8100 Italian municipalities, observed over the period between 2013 and 2018.

In Fig.5 we can see the evolution of this variable for both treated and control municipalities over the observed years. As we can see, trends are extremely similar, even if treated municipalities exhibit higher levels of expenditures per capita on average.

Fig. 5: **Trends in day care expenditures for treated and control muncipalities**



Note: Trends in day care-related expenditure per capita at the municipality level over the period 2013-2018. 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 population above 5000 inhabitants, while the control group includes those below. Sampled municipalities are 1422 out of 8100 Italian municipalities.

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

6 Empirical strategy and main results

The major goal of this paper's section is to empirically test our theoretical predictions, and, in particular, to understand which consequences the increase in female councillors had on local expenses on day care. One of the major implications of our model is that a large increase in the political representation of a minority could have detrimental effects in terms of policy: in our case, we identify this particular change in outcome with a decrease in day care expenditures. On the other hand, when the new law did not trigger a large increase in the political power of female politicians, this should have benefited them in terms of policies, thus we might expect an increase in daycare expenditures. For testing these implications, we make use of a Difference in Discontinuity identification strategy (Grembi et al. (2016)), which combines a Regression Discontinuity design with a Difference in Difference analysis. In our case, 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 focus, which is the municipality's day care expenditures per capita. We first present the aggregate results, and then the heterogeneous treatment effect with respect to the post-quota share of female councillors. In the next subsection, we discuss our identification strategy and its validity assumptions. After that, we present our main results and some related robustness tests.

6.1 Identification strategy and validity tests

Our model takes the following functional form:

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

With:

- $Daycarepc$ the sum of per capita day care-related expenses of municipality i at time t;
- $Treatment$ dummy for a municipality with more than 5000 inhabitants;
- $PostLaw$ dummy for expense taken by a council elected after Law 215/2012 entered into force;
- $TreatPost$ interaction between dummies $Treatment$ and $PostLaw$;
- $normPop$ normalized population, equal to population-5000;
- $Year$ year fixed effects;
- $City$ municipality fixed effects;

- $Mayor$ set of current mayor's controls (age, sex and level of education).

The treatment effect is identified by the coefficient δ, capturing the effect of the introduction of the law on treated municipalities around the threshold of 5000 inhabitants. We consider only municipalities belonging to a bandwidth $[normPop-h;normPop+h]$, with h computed by one of the methodologies suggested by Calonico et al. (2020), the standard MSE-optimal bandwidth[15]. 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 quota effect from another confounder that we have at the same 5000 inhabitants' thresholds: the mayor's wage. While for municipalities under 5000 inhabitants mayors earn 2170 euros per month, in cities above this cutoff (up to 10000 inhabitants) mayors' monthly wage increases to 2790 euros. This confounder might have effects on both selection into politics and on subsequent kinds 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 quota's effect from the wage's effect. Moreover, the fact that this strategy focuses on identifying the treatment effect at the threshold (thus, considering only municipalities belonging to a bandwidth) allows a more convincing comparison between treated and control units with respect to a simple Difference in Difference[16]. Three validity assumptions need to be satisfied for this identification strategy to be implemented (Grembi et al. (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

[15] Our results are however robust to the adoption of another, narrower, bandwidth minimizing the coverage error rate.

[16] Our major Difference-in-Discontinuity results are however robust in coefficients' signs and magnitudes to the use of a less local identification strategy such as the Difference-in-Difference

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 [17]. To check compliance with this condition, we test whether a wide set of covariates are 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[18] and demographic time-varying factors[19]. In Tables 3 and 4 we present our results.

Tab. 3: **RDD with time-invariant characteristics**

	(1)	(2)	(3)	(4)	(5)	(6)
VARIABLES	South	River	Lake	Surface	Sea distance	Altitude
RD_Estimate	-0.063	0.041	-0.069	3.034	-7.830	31.902
	(0.114)	(0.068)	(0.056)	(4.206)	(7.410)	(24.176)
Observations	3,643	3,643	3,643	3,643	3,643	3,643

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. 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 1422 out of 8100 Italian municipalities, observed over the period between 2013 and 2018.

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

[18] 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

[19] 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)

Tab. 4: **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	0.552	0.348	1.035	0.850	0.354	0.324	5213
	(1.683)	(0.299)	(0.838)	(0.677)	(0.296)	(0.302)	(0.535)
Observations	3,643	3,643	3,643	3,643	3,643	3,643	3,643

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. 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 1422 out of 8100 Italian municipalities, observed over the period between 2013 and 2018.

As we can observe, there are no jumps in municipalities' characteristics at the 5000 inhabitants' threshold, thus it seems that a set of potential outcomes is continuous in the running variable at this cutoff. 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, take as reference value the population measured by the last census, in this case 2011 [20]. And secondly, the population is measured by independent employees from Istat[21], with no presumable interest in manipulating the true value. However, in Appendix C 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

[20] Censuses in Italy have been performed every 10 years since 1861 until 2019, when they became yearly.

[21] Istat is the Italian National Institute of Statistics

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 should observe significant interaction effects between the treatment coefficient and mayors' characteristics. However, we do not observe any significant coefficients in the interaction terms, as we show in Appendix B. Therefore, we can be confident also about compliance with the third validity assumption[22].

6.2 Difference in Discontinuity results

In this section, we first show the aggregate Difference in Discontinuity results, and then we move our focus to the relevant heterogeneity analysis. In Tab.5 we present the first results, adopting the model described in section 3, with and without Mayors' controls. To preserve space, we decided to present only the coefficient capturing the treatment effect, TreatPost.

[22] 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.

Tab. 5: **Difference in Discontinuity aggregate results**

	(1)	(2)
VARIABLES	Daycarepc	Daycarepc
TreatPost	-0.715	-0.671
	(0.938)	(0.984)
Mayor's controls	NO	YES
Municipality FE	YES	YES
Year FE	YES	YES
Observations	2,690	2,629
R-squared	0.025	0.029
Number of municipality1	577	574

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

Note: 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 et al. (2020), through the minimization of the Mean Squared Error. Year and Time fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

As we can observe from Tab.5, the empirical evidence tells us that the aggregate effect of Law 215/2012 on day care-related expenses was zero. Moreover, we see a null aggregate impact also if we look at the treatment effect over time, by a dynamic Difference in Discontinuity similar to the one described by Vannutelli (2021): we show this ulterior piece of evidence in Appendix D.

However, our major theoretical implications state that increasing a minority's political representation can have different policy effects depending on the subsequent size of the minority group. In our empirical setting, we would therefore need to assess the heterogeneous law's effects on day care with respect to each council's post-quota share of female councillors. The next section is devoted to analyzing these possible heterogeneous effects.

6.3 New law's heterogeneous effects on day care

Our model concludes that an increase in the share of a minority reduces the probability of compromise with the majority, increasing the chances that the majority propose its bliss point. Therefore, in our empirical context we would expect to see a decrease in day care expenditures in places where the increase in female politicians was so large that prevented bargaining on day care between male and female councillors. Therefore, we look at heterogeneous results with respect to the post-quota share of female councillors in treated municipalities[23]. In order to do this, we proceed in the following way:

1. Generate the variable "Female post quota", indicating the percentage of female councillors elected in the first election after the new law entered into force

2. Create percentile dummies based on the distribution of this variable across treated municipalities. To be more precise, we create the dummies for being lower than the 25th, higher than the 50th, 75th or 90th percentiles

3. Interact the treatment variable $TreatPost$ with each percentile dummy and assess the heterogeneous effects of the new law on day care expenditures. We interact the 25th dummy to look for the effect on councils with low post-quota female presence. The interactions with the 50th,75th and 90th percentile should assess the effects in councils with high post-quota shares of female councillors.

We exclude from this analysis the few treated councils in which there was a female majority, since they would not comply with our model's assumptions. For our analysis to be correct, we would need gender to prevail over ideology when voting for day care, since coalitions should be forming also with members of different parties: we believe this to be true for three reasons. First, there is evidence that female politicians have stronger preferences for day care with respect to their male colleagues also in developed contexts such as Germany or Belgium (Baskaran and Hessami (2023), Slegten and Heyndels (2020)). Subsidized day care is a service that benefits disproportionally women with respect to men, and this supports the idea of a gender difference in preferences for this kind of public expenditure (Bhalotra and Clots-Figueras (2014)). In addition, we know that the 95% of politicians in our sample belong to civic lists, without a clear ideological belonging: this guarantees them more freedom in the voting process, without having to follow any national party's guidelines. Last, voting in

[23] An interesting piece of evidence to add would have been to study heterogeneous effects with respect to the gender ratio of the ruling coalition, but unfortunately, the available data does not provide this information.

the Italian local councils is secret, thus there is even more freedom for councillors in choosing the alternative that they best prefer: in other words, local politicians have the possibility to deviate from party lines when they cast their choice. In Tab.7 we show how the increase in female politicians affected day care expenditures differently on the basis of each council's post-quota share of female councillors.

Tab. 6: **Difference in Discontinuity heterogeneous results**

VARIABLES	Baseline interaction Daycarepc	25th percentile Daycarepc	median Daycarepc	75th percentile Daycarepc	90th percentile Daycarepc
TreatPost	3.936	-0.984	-0.104	-0.247	-0.615
	(2.541)	(1.059)	(1.049)	(1.022)	(1.017)
TreatPost*Fem. post quota	-12.211*				
	(6.522)				
Twenty-fifth p.		2.088**			
		(1.009)			
TreatPost*Twenty-fifth p.		1.863*			
		(1.130)			
Median			-1.484		
			(0.938)		
TreatPost*Median			-1.800*		
			(1.020)		
Seventy-fifth p.				1.518	
				(1.917)	
TreatPost*Seventy-fifth p.				-3.510	
				(2.444)	
Ninetieth p.					0.991
					(1.706)
TreatPost*Ninetieth p.					-2.447
					(2.121)
Observations	2,382	2,446	2,446	2,446	2,446
R-squared	0.035	0.032	0.034	0.037	0.032
Number of municipality1	544	559	559	559	559
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

Note: 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-quota 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 5000 inhabitants' threshold is computed following one of the procedures described by Calonico et al. (2020), through the minimization of the Mean Squared Error. Year and Time fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

Looking at Tab.6, we can immediately observe that the heterogeneous treatment effect, given by the sum between $TreatPost$ and $TeatPost * percentile$ takes different directions on the basis of the share of post-quota female politicians in cities' councils. The interaction between the treatment coefficient and the continuous measure of $Femalepostquota$ is negative, following our theoretical reasoning: in the councils with higher minority representation, the introduction of the law was followed by a decrease in day care expenditures compared to control municipalities. This is also confirmed when we look at the third, fourth and fifth columns, with the interaction between treatment and $Femalepostquota$ higher than the median, 75th and 90th percentile, all characterized by negative coefficients. The logic is that in these cities the new law increased the share of female councillors up to a level where a compromise with male councillors on day care expenditures would have been too expensive for the majority, who therefore decided to go with its respective bliss point and implementing an expenditure level lower than the pre-quota amount. On the other hand, we see in the second column of Tab.6 that in councils with low post-quota levels of female councillors (lower than the 25^{th} percentile), the increase in female politicians was followed by an increase in funding for day care compared to the control group. This is also aligned with our theoretical implications: with low minority representation, a compromise between the two groups of politicians is still possible and the larger political power of female politicians is able to move the policy outcome closer to their desired outcome. To give a sense of the magnitude of our effect, we have that in cities with low levels of female councillors the increase in expenditure was about 0,88 euros per capita, a 9.5% growth with respect to the pre-quota period and the control municipalities. On the other hand, in cities with $Femalepostquota > 50^{th}$ percentile, the size of the decrease was about 1,7 euros per capita, an 18% reduction with respect to the control group[24]. Recall that in all these municipalities men still constitute the majority of councillors, which is required in order to approve or reject motions on public expenditures: therefore, even after the quota, male politicians still detained the decisional power. We might be worried that the share of female councillors is persistent, or in other words that the cities with high/low shares of female councillors

[24] 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 a negative effect in those with high shares.

had different starting levels in this variable. In Appendix E, we add as control the pre-law share of female councillors: our main results are preserved even after the inclusion of this control. In addition, in Appendix G we show how this result can be replicated in another similar context, which is the one relative to the introduction of a gender quota in 2011 in Spanish administrative elections (Bagues and Campa (2021)). Moreover, in the next section we offer multiple pieces of evidence supporting the robustness of our results, starting from a comparison between cities with high and low shares of post-law female councillors.

7 Robustness and additional evidence

7.1 Cities and post-quota shares of female councillors

A major concern for our analysis is that what we are observing is not the heterogeneous effect with respect to the post-quota 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[25] for treated cities with high and low shares of female councillors (respectively, post-quota female councillors higher than the 75th percentile and lower than the 25th percentile), and check if the differences in means are statistically significant. In Tab.7 we can observe these comparisons.

[25] We look at variables referring to municipalities' balance sheets (e.g. Tot. expenditures), population's characteristics (e.g. Years of Education or Age Structure), geographical characteristics and local politicians' attributes.

Tab. 7: **Test for difference in means, cities with low and high levels of post-quota female councillors**

Variable	High_Share	Low_Share	Difference	p-value
Tot. Expenditures	2937784	2569516	-368268.4	.008***
Mean Y.o.E. population	8.820	8.896	.075	.300
Age Structure (M)	1.356	1.332	-.023	.608
Age Structure (F)	1.787	1.725	-.062	.349
Age Structure (Tot.)	1.564	1.521	-.043	.434
Fertility rate	39.342	38.963	-.378	.805
River	.431	.417	-.013	.824
Lake	.315	.379	.064	.262
Sea Distance	69.723	80.610	10.886	.111
Surface (km^2)	33.796	32.543	-1.253	.777
Mean altitude	227.177	231.675	4.498	.855
Population density	371.505	325.033	-46.472	.278
South	.341	.291	-.050	.383
Gender participation gap 2011	20.302	19.863	-.439	.432
Female labor force participation 2011	41.385	42.388	1.002	.256
Fertility rate	39.342	38.963	-.378	.805
Female unemployment rate 2011	12.614	12.202	-.412	.639
Female mayor	.140	.177	.037	.387
Y.o.E. mayor	15.819	15.873	.053	.874
Mean Y.o.E. councillors	14.345	14.433	.087	.582
Mean age councillors	49.599	50.620	1.0211	.027**
Mean Y.o.E. aldermen	14.434	14.475	.040	.919
Mean age aldermen	45.890	45.194	-.695	.499
Mean Y.o.E. female councillors	14.410	14.784	.374	.409
Mean age female councillors	43.063	43.360	.296	.657
Mean Y.o.E. male councillors	13.912	14.154	.242	.227
Mean age male councillors	48.009	47.778	-.230	.683
Union of municipality dummy	.319	.392	.072	.203

Note: Test for statistical differences in means of variables between group of treated municipalities with low and high shares of post-quota female councillors (post-quota female councillors<25th percentile and post-quota female councillors>75th 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 tells us that there are a couple of statistically significant differences between the two groups of cities. What we observe is that cities with high levels of post-quota 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[26]. In particular, we underline how fertility rates, which are a potential huge driver of the demand for childcare, are absolutely comparable between the two groups of cities.

7.2 Placebo test - other expenditure categories

In this section, we conduct an important placebo test, verifying the new law's effect on other expenditure categories. We know from the model that a situation of conflict is more likely in the case of divergent preferences between two groups. In this sense, we know that men and women have different preferences with respect to day care (Barigozzi et al. (2019)), but we might also think that the introduction of the gender quota raised the salience of gender topics, which can include expenditures on day care. Indeed, according to experimental evidence (Coffman (2014)), women are more likely to expose their personal ideas in environments which are not male-dominated: an increase in the share of female councillors could have encouraged them to bring their political views on the discussion table. Thus, male and female politicians should have clashed over this issue but not on other kinds of expenditures that were not gender-related: here, we verify the soundness of this hypothesis. Given that each municipality's balance sheet contains more than 100 expenditure categories, we focus here on the 10 most relevant ones in terms of share of total expenditures, with the condition

[26] I performed also the test in differences (i.e. comparing municipalities with high post-quota 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.

for them to be decided by the local council[27]:

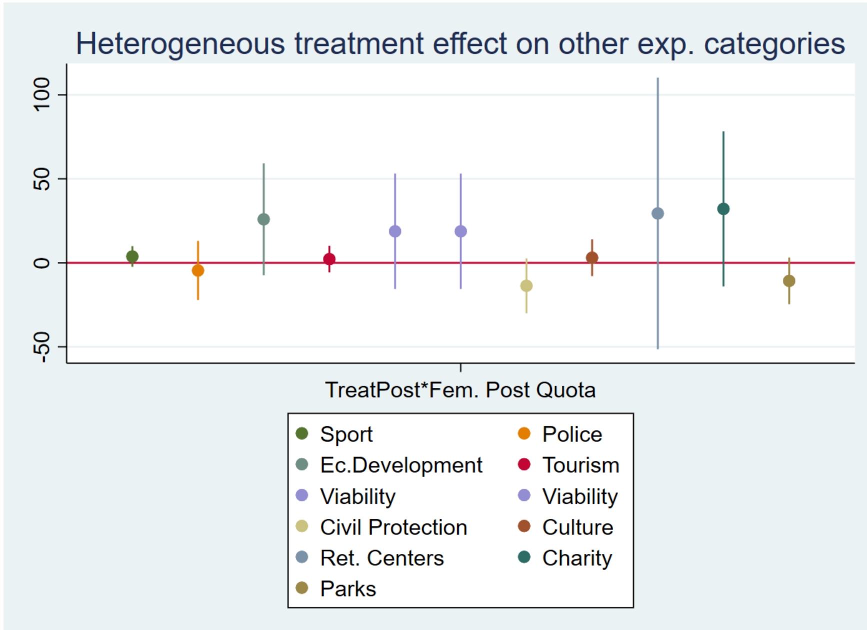
- Sport
- Police
- Economic development
- Tourism
- Public viability
- Civil protection
- Cultural events
- Retirement centers
- Charity
- Public gardens

All these expenditure categories, differently from day care, are not gender-related: therefore, there should not be a difference in preferences between male and female councillors[28]. To preserve space, in Fig.6 we exclusively show the interaction coefficients between the treatment dummy and the Female post-quota variable. Each coefficient displayed refers to the interaction between Treatpost and Female post-quota when we perform our Difference in Discontinuity regression with the outcomes indicated in the legend.

[27] We focus on current, and not capital, expenditures per capita since they are more likely to be affected by sudden changes in the political class composition. However, results with capital expenditures are not different in terms of magnitude/significance level of coefficients.

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

Fig. 6: **Law's effect on other expenditure categories**



Note: Difference in Discontinuity heterogeneous effects with respect to the city council's post quota shares of female councillors. Each coefficient corresponds to the interaction between Treatpost and Female Post Quota, when we perform our Difference in Discontinuity regression with the expenditure categories indicated in the legend as the main outcomes. Confidence intervals are at the 95% level.

From Fig.6, we can see how there seem to be no significant effects from the interaction between Treatpost and Female Post Quota on any of these expenditure categories: we exclusively observe our main result on day care expenditures, given its gender-related connotation. Indeed, in line with our model, the exogenous shock brought by the gender quota increased the divergence over gender-related expenditures between male and female politicians, raising the chances of conflict. On the other hand, conflict was not generated on other kinds of expenditures, given that the two groups of politicians do not have different preferences over them. In addition, this placebo is also useful to reduce the concern 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 addition to this robustness, in Appendix F we perform two additional placebo tests with artificial cutoffs and year of the law's implementation. Results from these two more tests confirm that what we observe in our main regressions are not simply spurious results.

7.3 2SLS panel regression

We perform a further robustness analysis by changing the identification strategy and recurring to an instrumental variable approach. Indeed, we might think that the Female post quota variable is endogenous to some unobserved municipality's characteristics: the shares of elected female politicians can indeed depend on a large variety of factors intrinsic to the electorate's and city's characteristics (Hessami and da Fonseca (2020)). Thus, given the possible endogeneity of this variable, in this section we instrument the percentage of female councillors with Treatpost: this way, we are able to 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 councillors on funding for day care[29][30]. As we explain in Section 4, the law 215/2012 had a positive effect on the share of female councillors. This positive effect is confirmed in Tab.8,

[29] This effect is "Local" in the sense that it interests only those treated municipalities that had their shares of female councillors increased by the new law

[30] For this identification strategy to be valid, we need the instrument to influence the outcome only through its effect on the 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 expenditures' composition, and it did not change how the local administrators could spend public money, plus it did not affect their revenues. Moreover, it's hard to think that the new law generated any effect on the demand for daycare, since we saw how other variables such as the number of spots did not change. Thus, we argue that being subject to the gender quota did not alter directly the municipalities' expenditures for daycare, but it affected daycare only through its effect on female councillors.

which displays our first stage in the 2SLS setting.

Tab. 8: **First stage regression**

	(1)
VARIABLES	Female councillors
TreatPost	0.099***
	(0.007)
Municipality FE	YES
Year FE	YES
Observations	7,896
Number of municipalities	1,442
R-squared	0.347
First Stage F test	120.0

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

Note: First stage regression in the 2SLS setting. TreatPost corresponds to the interaction between the dummy identifying the treatment group and the dummy for the post-treatment period. The dependent variable is the share of female councilors in the municipalities. All 1442 municipalities belonging to our sample are included in the regression. Year and Time fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

As we can see, the introduction of the law increased on average the female presence in city councils across treated cities, which was indeed the main goal of Law215/2012. In Tab.9 we show the second stage results, together with OLS base results.

Tab. 9: **OLS and second stage regression**

VARIABLES	OLS Daycarepc	2SLS Daycarepc	2SLS Daycarepc
Female councillors	-0.105	-5.555***	-5.096*
	(0.734)	(1.928)	(2.806)
treatment	-2.180***		-2.057***
	(0.726)		(0.718)
Postquota	-0.215		0.346
	(0.204)		(0.372)
Observations	7,896	7,896	7,896
Number of municipalities	1,422	1,422	1,422
Muni FE	yes	yes	yes
Year FE	yes	yes	yes

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

Note: Second stage regression in the 2SLS setting, together with OLS base results. Female councillors is instrumented with Treatpost, the interaction between the dummy identifying the treatment group and the dummy for the post-treatment period. The dependent variable is the per capita expenditures on day care measured at the municipality level. All 1442 municipalities belonging to our sample are included in the regression. Year and Time fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

As we can see, the post-law increase in female councillors decreased on average day care expenditures across treated municipalities[31]. On the other hand, the correlation between polarization and day care expenditures is almost null and non-significant in the OLS results: this is not particularly surprising, since this correlation takes into account also control municipalities and the period before the introduction of the quota. The magnitude of the last column's effect is larger with respect to our Diff-in-Disc results: the decrease in 2SLS results is of roughly 5 euros per capita, with respect to the 1,8 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, Tab.11 presents a LATE effect, which regards 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-

[31] 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.

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-quota variable: it might be that these biases are shrinking 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.

7.4 Effect on councils' dissolution

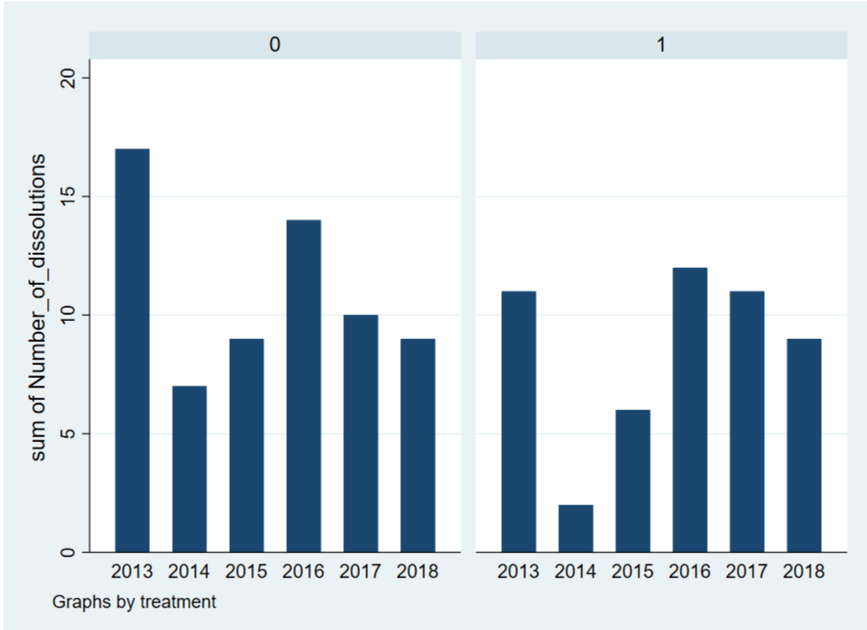
In Italy, municipal councils can be dissolved for different reasons:

- Persistent law violations and threats to public policy (e.g. mafia infiltrations)
- Mayor's forfeiture because of resignation, promotion to higher office or death
- Resignation of more than half of local councillors
- Passage of a nonconfidence motion[32]
- Prolonged financial distress

In case the council is dissolved, the municipality's local administrators are replaced with external commissioners, and local public finances are placed under supervised management until new elections are held. Some of these cases can signal a heavy divergence of opinions among local councillors: in this subsection, we focus on councils dissolved because of the resignation of more than half of the local councillors. This possibility happens mostly for persistent disagreement about the implementation of the political agenda. In Fig.7 we see trends in dissolutions for this specific reason over time, for treated and control municipalities belonging to our sample.

[32] The nonconfidence motion must be presented by at least 2/5 councillors and has to be approved by the majority of councillors in order to pass.

Fig. 7: **Number of dissolutions because of councillors' resignation**



Note: Number of local councils dissolved because of councillors' resignation by year, for treated (label 1) and control (label 0) municipalities. The whole sample of 1422 municipalities is included.

As we see, control municipalities get dissolved more frequently for this reason with respect to treated municipalities: however, yearly trends between the two groups of municipalities are similar[33]. In relative terms, the resignation of the majority of councillors is a rare event: the percentage of our sampled municipalities that experiences this kind of council dissolution is 3.8%. Nevertheless, these specific dissolutions can signal a conflict between local councillors and, in line with our theoretical reasoning, we might expect that this happened more frequently when the new law generated a larger presence of female councillors. Indeed, in our model, a larger minority presence is decreasing the probability of compromise between the two groups. In our empirical setting, this finding can translate into increased chances of conflict among councillors after the introduction of the gender quota. To assess whether Law 215/2012 raised the frequency of councils' dissolutions, we create a "Dissolution dummy" taking value 1 for the year when the local council was dissolved for the resignation of the majority of local councillors. The "Dissolution dummy" takes value 1 also for the previous

[33] Note that so far we excluded the year of dissolution for municipalities dissolved for any reason from our main analysis, since, as we pointed out, public finances are placed under a special regime in these cases.

years of the specific dissolved councils, since it identifies the future conflicting councils. This variable is the outcome of Tab.10, which shows new Difference in Discontinuity regressions with heterogeneous results aimed at verifying whether dissolutions were more frequent after the Law 215/2012 introduction.

Tab. 10: **Heterogeneous Law's effect on councils' dissolutions**

VARIABLES	(1) Dissolution	(2) Dissolution	(3) Dissolution	(4) Dissolution	(5) Dissolution
TreatPost	-0.098	0.010	-0.018	-0.002	0.000
	(0.062)	(0.026)	(0.028)	(0.027)	(0.027)
TreatPost*Twentyfifth p.		-0.081			
		(0.068)			
TreatPost*Fem. post quota	0.265*				
	(0.140)				
Twenty-fifth p.		-0.442			
		(1.388)			
TreatPost*Twenty-fifth p.		1.002			
		(1.738)			
Median			0.465		
			(0.353)		
TreatPost*Median			0.055***		
			(0.020)		
Seventy-fifth p.				0.465	
				(0.353)	
TreatPost*Seventy-fifth p.				0.028**	
				(0.013)	
Ninetieth p.					0.465
					(0.353)
TreatPost*Ninetieth p.					0.033**
					(0.014)
Observations	2,483	2,547	2,547	2,547	2,547
R-squared	0.037	0.059	0.062	0.053	0.053
Number of municipalities	515	528	528	528	528

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

Note: 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-quota 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 et al. (2020), through the minimization of the Mean Squared Error. Year and Time fixed effects are included in each regression. Data were constructed from the Ministry of Interior database. Robust standard errors in parentheses are clustered at the municipality level.

From Tab.10 we see that in councils where the Law's introduction generated a larger increase in female councillors, this led to a raised probability of council's dissolution compared to the control group: columns 1,3,4 and 5 confirm our theoretical prediction. If we interpret the dissolution as a signal of conflict, we can say that having a more balanced composition of male and female councillors increased the probability of political disagreement[34]. On the other hand, in councils with low levels of post-quota female politicians, this probability was not affected (column 2). To get an idea of the magnitude of the effect, for levels of post-quota female councillors higher than the median, the effect of the law was an increase of 3.7 percentage points in the probability of dissolution. This evidence hints at the possibility that the increase in female politicians leads to further conflict in some municipalities, not only by changing public expenditures but also ending up dissolving the local council. Again, this is a reduced form effect and we cannot exclude possible other effects leading to councils' dissolution. However, this other piece of evidence, together with the other presented ones, supports our major claim that the Law increased the probability of political conflict in some cities[35].

8 Conclusion

In the last decades, Western societies have been acting in order to increase the political power and representation of some social groups that were unable to influence the process of political decision-making. These measures, such as the extension of suffrage to Afro-Americans and the introduction of gender quotas, have been proven effective in benefiting these groups, in terms of granting them more social rights and access to welfare programs (Miller (2008), Fujiwara (2015), Chattopadhyay and Duflo (2004)). However, in this paper we highlight how

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

[35] We might wonder why these dissolutions did not generate ulterior inefficiencies on other kinds of expenditures apart from day care. On this matter, we cannot exclude that the political conflict extended to other kinds of expenditures, but given the large number of categories (over 100), it is possible that different conflicting councils clashed over different categories. In other words, it is hard to believe that all of these councils conflicted on the same expenditure category, apart from day care, given its gender-related connotation. Therefore, we do not observe a clear effect on sport or tourism, for instance, because it is not obvious that each conflicting council would raise the funding for these specific categories after the reduction in day care funding.

in the short run the increase of political representation of an under-represented social group might potentially generate unintended consequences, which could both benefit and damage the group. Indeed, increasing the representation of a minority can reduce the possibility of compromising with the majority, as the claims of the minority group might also be stronger and more distant from the majority's preferences. As we show in our theoretical model, the growth in the political minority group's share decreases the probability of reaching a mutual agreement on a policy for which the two groups have distant preferences. The growth of the minority's share increases its probability of winning the conflict and its demands in terms of policy outcomes: as a consequence, compromise becomes too costly for the majority which might decide to propose its bliss point as the policy level. This situation moves the final decision on the policy level to a more extreme outcome which is less preferred by the minority with respect to the outcome that would have been obtained with a compromise. The probability of compromise between groups is negatively related to the minority group's share and to the difference in the two groups' bliss points concerning the policy level. On the other hand, our model predicts that having greater costs of conflict raises the probability of avoiding a conflict. We test some of these theoretical implications by exploiting the introduction of a gender quota in Italy in 2012, which substantially raised the share of female politicians at the local level. We assess the consequences of this shock in the political class composition on the level of investment for day care, a service for which men and women typically have different preferences (Barigozzi et al. (2019), Bhalotra and Clots-Figueras (2014)), and find that there were heterogeneous consequences with respect to the post-law share of female councillors. By the means of a Difference-in-Discontinuity strategy (Grembi et al. (2016)), we show that where the quota substantially raised the share of female politicians in local councils, there was a cut in expenditure for day care with respect to the cities not subject to the quota. On the other hand, with low shares of female councillors, the level of funding for this particular expenditure category was increased with respect to the control group. The empirical findings are in line with our theoretical model since the minority group's share influences the probability of compromise between the two groups of politicians. A series of robustness tests support the soundness of our results: we prove that there are no substantial differences between cities with high and low post-law shares of female councillors, and three different placebo tests reinforce our claims of causality. Moreover, through an instrumental variable analysis, we show that the negative effect on day care expenditures was indeed driven by the quota-induced increase in female councillors. Lastly, we show more evidence supporting our idea that an increase in the minority group size can fuel political disagreement in the short run. Indeed, the councils with larger post-law shares of female councillors were more likely to be dissolved in the years after elections with respect to councils with low scores in this variable.

In conclusion, the short-run consequences of enhancing a minority's political representation can be more complex than expected. If we needed to suggest a policy implication, it would be to cautiously assess potential heterogeneity in preferences between politicians: giving more political power to a group facing another one with distant preferences could be detrimental in the short run. However, increasing the political power of minorities, as we said previously, grants consistent long-run benefits, thus we are not advocating a limit to their enfranchisement. More simply, we are suggesting to assess also possible short-run distortions that might be caused by sudden shocks in the composition of the political class.

References

Abhay P Aneja and Carlos F Avenancio-Leon. The effect of political power on labor market inequality: Evidence from the 1965 voting rights act. *Unpublished working paper*, 2019.

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

Audinga Baltrunaite, Alessandra Casarico, Paola Profeta, and Giulia Savio. Let the voters choose women. *Journal of Public Economics*, 180:104085, 2019.

Francesca Barigozzi, Helmuth Cremer, and Chiara Monfardini. The gender gap in informal child care: theory and some evidence from italy. 2019.

Thushyanthan Baskaran and Zohal Hessami. Women in political bodies as policymakers. 2023.

Lori Beaman, Esther Duflo, Rohini Pande, and Petia Topalova. Female leadership raises aspirations and educational attainment for girls: A policy experiment in india. *science*, 335(6068):582–586, 2012.

Timothy Besley and Anne Case. Does electoral accountability affect economic policy choices? evidence from gubernatorial term limits. *The Quarterly Journal of Economics*, 110(3): 769–798, 1995.

Timothy Besley and Stephen Coate. An economic model of representative democracy. *The quarterly journal of economics*, 112(1):85–114, 1997.

Sonia Bhalotra and Irma Clots-Figueras. Health and the political agency of women. *American Economic Journal: Economic Policy*, 6(2):164–97, 2014.

Giampaolo Bonomi, Nicola Gennaioli, and Guido Tabellini. Identity, beliefs, and political conflict. *The Quarterly Journal of Economics*, 136(4):2371–2411, 2021.

Ylenia Brilli, Daniela Del Boca, and Chiara D Pronzato. Does child care availability play a role in maternal employment and children's development? evidence from italy. *Review of Economics of the Household*, 14(1):27–51, 2016.

Sebastian Calonico, Matias D Cattaneo, and Max H Farrell. Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs. *The Econometrics Journal*, 23(2):192–210, 2020.

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

Irma Clots-Figueras. Are female leaders good for education? evidence from india. *American Economic Journal: Applied Economics*, 4(1):212–44, 2012.

Katherine Baldiga Coffman. Evidence on self-stereotyping and the contribution of ideas. *The Quarterly Journal of Economics*, 129(4):1625–1660, 2014.

Anthony Downs et al. An economic theory of democracy. 1957.

Joan Esteban and Debraj Ray. Conflict and distribution. *Journal of Economic Theory*, 87 (2):379–415, 1999.

Joan-Maria Esteban and Debraj Ray. On the measurement of polarization. *Econometrica: Journal of the Econometric Society*, pages 819–851, 1994.

Fernando Ferreira and Joseph Gyourko. Does gender matter for political leadership? the case of us mayors. *Journal of Public Economics*, 112:24–39, 2014.

Thomas Fujiwara. Voting technology, political responsiveness, and infant health: Evidence from brazil. *Econometrica*, 83(2):423–464, 2015.

Stefano Gagliarducci and Tommaso Nannicini. Do better paid politicians perform better? disentangling incentives from selection. *Journal of the European Economic Association*, 11(2):369–398, 2013.

Stefano Gagliarducci and M Daniele Paserman. Gender interactions within hierarchies: evidence from the political arena. *The Review of Economic Studies*, 79(3):1021–1052, 2012.

Andrew Gelman and Guido Imbens. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business & Economic Statistics*, 37(3):447–456, 2019.

Benny Geys and Rune J Sørensen. The impact of women above the political glass ceiling: Evidence from a norwegian executive gender quota reform. *Electoral Studies*, 60:102050, 2019.

Veronica Grembi, Tommaso Nannicini, and Ugo Troiano. Do fiscal rules matter? *American Economic Journal: Applied Economics*, pages 1–30, 2016.

Gene M Grossman and Elhanan Helpman. Identity politics and trade policy. *The Review of Economic Studies*, 88(3):1101–1126, 2021.

Chris M Herbst. Universal child care, maternal employment, and children's long-run outcomes: Evidence from the us lanham act of 1940. *Journal of Labor Economics*, 35(2): 519–564, 2017.

Zohal Hessami and Thushyanthan Baskaran. Competitively elected women as policy makers. 2019.

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

Grant Miller. Women's suffrage, political responsiveness, and child survival in american history. *The Quarterly Journal of Economics*, 123(3):1287–1327, 2008.

Caroline Slegten and Bruno Heyndels. Within-party sex gaps in expenditure preferences among flemish local politicians. *Politics & Gender*, 16(3):768–791, 2020.

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

Appendices

A Case $q_F < 2\sqrt{2\gamma} + 4\sqrt{\gamma}$

In the case $q_F < 2\sqrt{2\gamma} + 4\sqrt{\gamma}$, we have that $\pi_2 < \frac{1}{2}$, and an increase in the minority share can switch a situation of conflict to a compromise. In this case, the equilibrium depends on π: we can have four ranges of the minority share with different groups' strategies and policy outcomes.

1. For low minority shares, which is when $\pi < \frac{2\gamma}{q_F^2}$, the majority proposes its bliss point q_M and there is no conflict.

2. When $\pi \in [\frac{2\gamma}{q_F^2}; \pi_1]$, the majority chooses a policy level closer to the minority's preferences and conflict is avoided: the implemented level is $q^* = q_F - \sqrt{\frac{2\gamma}{\pi}} = \overline{q}_M$.

3. When $\pi \in [\pi_1; \pi_2]$, the majority proposes its bliss point q_M and conflict happens. In this case, the expected policy outcome is a weighted average of the two bliss points: $E(q_I, C) = \pi q_F + (1-\pi) q_M = \pi q_F$.

4. For large minority shares, which is when $\pi > \pi_2$, for the majority it is more profitable to compromise with the minority and offer $q^* = q_F - \sqrt{\frac{2\gamma}{\pi}} = \overline{q}_M$.

The interpretation of this last point is that in this case conflict has a high chance of being won by the minority, and the majority prefers to avoid this risk: by offering $\overline{q}_M$, conflict is avoided. Here, an increase in the minority share can change its optimal strategy from conflict to no conflict, since the group sees its requests satisfied. Thus, we have a situation in which an increase in the size of the minority group can be beneficial for it, given that at $\pi = \pi_2$, $E(q_I))$ increases.

B Difference in Discontinuity: third validity assumption

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 et al. (2016)). In this subsection we discuss the third assumption, namely the absence of relevant interaction effects between the treatment and the other confounder present at the 5000 inhabitants' threshold, which is the increase in the mayor's salary. 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. Here, our reasoning is that the higher salary might influence the selection into politics and bring different mayors into charge who could interact diversely with the new councillors elected thanks to the gender quota. Thus, if our treatment effect is different because of this confounder, we should observe significant interaction effects between the treatment coefficient and mayors' characteristics. The mayor's characteristics we consider are the ones for which we have data on the Ministry of Interior's website: mayor's sex, education and age. In Tab.11 we check whether any of this interaction effects is relevant or present.

Tab. 11: **Interactions treatment coefficient and mayor's characteristics**

VARIABLES	(1) Daycarepc	(2) Daycarepc	(3) Daycarepc
TreatPost	-1.423	0.112	2.364
	(1.878)	(0.837)	(2.148)
Mayor_fem	0.224		
	(0.464)		
TreatPost*Mayor_fem	-2.150		
	(1.410)		
Mayor_age		-0.011	
		(0.020)	
TreatPost*Mayor_age		0.026	
		(0.032)	
Mayor_Ed			-0.066
			(0.072)
TreatPost*Mayor_Ed			-0.156
			(0.125)
Municipality FE	YES	YES	YES
	YES	YES	YES
Year FE	YES	YES	YES
	YES	YES	YES
Observations	3,345	3,409	3,354
R-squared	0.032	0.033	0.033
Number of municipalities	704	707	704

Robust standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

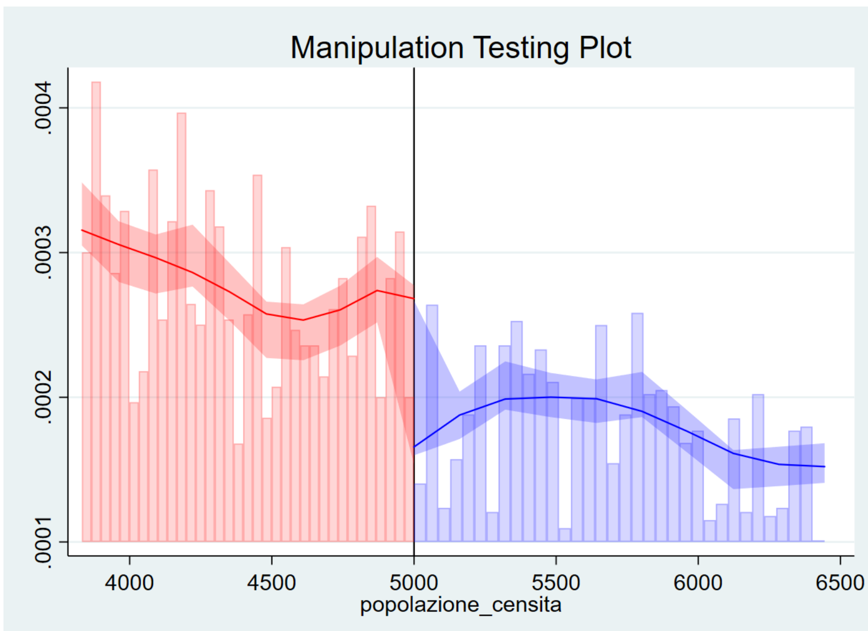
Note: Difference in discontinuity regressions. TreatPost corresponds to the treatment coefficient, the interaction between the dummy identifying the treatment group and the dummy for the post-treatment period. TreatPost is interacted with each one of the current mayor's characteristics, namely sex, age 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 et al. (2020), through the minimization of the Mean Squared Error. Year and Time fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

As we can observe from Tab.11, there are no significant interaction effects between our treatment coefficient and each of the three mayor's characteristics we have data on, namely sex, age and education. We can conclude that the third validity assumption for our Difference in Discontinuity strategy is satisfied.

C McCrary test

We want to test whether there could have been manipulation in the running variable at the threshold of 5000 inhabitants. Fig.8 shows the McCrary test, which analyzes whether there is a significant difference in observational densities at the cutoff.

Fig. 8: **% McCrary test for manipulation of running variable**



Note: 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 1422 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.

D Dynamic Difference in Discontinuity

The dynamic Difference-in-Discontinuity Vannutelli (2021) we exploit here interacts the dummy identifying the treatment group with each of the periods observed, 1/2/3/4 years before the first elections with the new law and 1/2/3/4 years after. This way, we are able both to observe the treatment effect in a dynamic way and to perform a placebo test, checking whether the pre-treatment periods present null treatment effects[36]. Also in this case, we focus on the reduced sample of municipalities belonging to the bandwidth computed before. Tab.12 presents the point estimates, while Fig.9 shows the Dynamic Difference in Discontinuity results graphically, to allow an easier understanding of this other result.

[36] In other words, we are empirically testing a parallel trends assumption

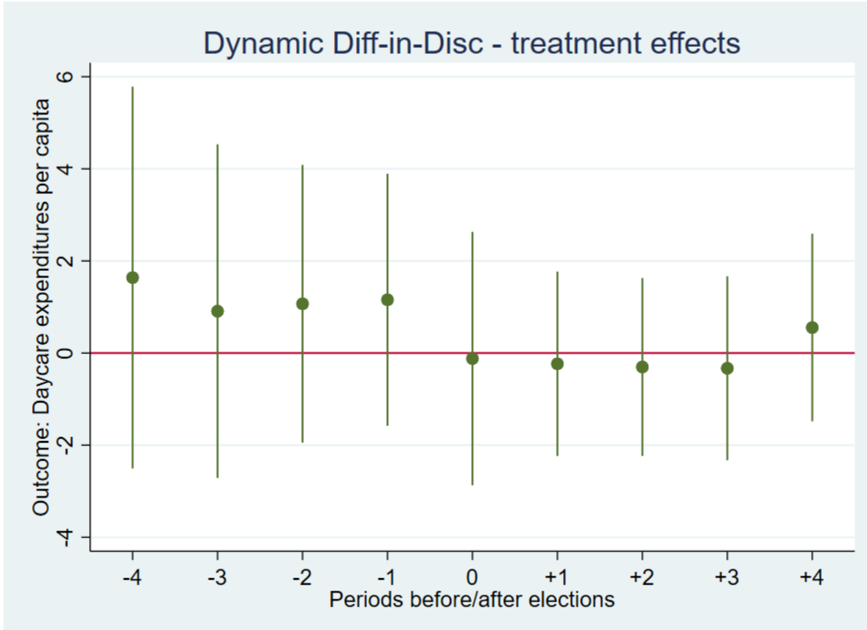
Tab. 12: **Dynamic Difference in Discontinuity results**

	(1)	(2)
VARIABLES	Daycarepc	Daycarepc
Treated*4 years before	1.551	1.639
	(1.980)	(2.110)
Treated*3 years before	0.873	0.909
	(1.758)	(1.844)
Treated*2 years before	0.996	1.070
	(1.425)	(1.535)
Treated*1 year before	1.137	1.157
	(1.288)	(1.393)
Treated*Election year	-0.022	-0.120
	(1.303)	(1.400)
Treated*1 year after	-0.208	-0.233
	(0.949)	(1.020)
Treated*2 years after	-0.289	-0.303
	(0.918)	(0.983)
Treated*3 years after	-0.304	-0.331
	(0.960)	(1.017)
Treated*4 years after	00.516	0.554
	(0.984)	(1.037)
Mayor's controls	NO	YES
Municipality FE	YES	YES
Year FE	YES	YES
Observations	2,641	2,582
R-squared	0.022	0.025
Number of municipalities	576	573

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

Note: Dynamic Difference in Discontinuity results. The dynamic treatment effect was 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 et al. (2020), through the minimization of the Mean Squared Error. Year and Time fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

Fig. 9: **Dynamic Difference in Discontinuity results**



Note: 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 et al. (2020), through the minimization of the Mean Squared Error. Year and Time fixed effects are included in each regression.

Looking at both Tab.12 and Fig.9, we can conclude that the aggregate treatment effect has not been significantly different than zero, even if we look at it in a dynamic sense. The results we see from the Dynamic Difference in Discontinuity also exclude differences between treated and control municipalities in political budget cycles: even graphically, we do not observe recurring paths in the periods before or after the elections. This absence of differences in spending cycles contributes to supporting the idea of comparability between treated and control municipalities.

E Different level of pre-quota female councillors

A potential concern we need to discuss comes from the fact that the heterogeneity that we found might depend on pre-quota levels of female councillors, which in turn might be driven by other municipalities' characteristics also linked to day care services. Even if this particular heterogeneity should be captured by municipality's fixed effects, we perform an additional robustness and introduce the pre-quota level of female councillors as a control in our Difference in Discontinuity regressions, to exclude this potential confounder for our analysis. Tab.13 presents this additional evidence.

Tab. 13: **Difference in Discontinuity heterogeneous results**

VARIABLES	Baseline interaction Daycarepc	25th percentile Daycarepc	median Daycarepc	75th percentile Daycarepc	90th percentile Daycarepc
TreatPost	4.014	-0.831	0.075	-0.116	-0.482
	(2.559)	(1.068)	(1.065)	(1.040)	(1.033)
TreatPost*Fem. post quota	-12.242*				
	(6.549)				
Twenty-fifth p.		5.244***			
		(1.273)			
TreatPost*Twenty-fifth p.		1.875*			
		(1.132)			
Median			-5.190***		
			(1.350)		
TreatPost*Median			-1.798*		
			(1.033)		
Seventy-fifth p.				0.469	
				(2.698)	
TreatPost*Seventy-fifth p.				-3.584	
				(2.576)	
Ninetieth p.					-0.271
					(2.503)
TreatPost*Ninetieth p.					-2.426
					(2.264)
Fem. pre-quota		17.977**	23.362***	9.252	10.798
		(8.985)	(8.114)	(10.984)	(10.831)
Observations	2,382	2,446	2,446	2,446	2,446
R-squared	0.036	0.034	0.037	0.039	0.034
Number of municipality1	544	559	559	559	559
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

Note: 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-quota 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-quota 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 et al. (2020), through the minimization of the Mean Squared Error. Year and Time fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

What we see from Tab.13 is that our main coefficients are not modified in magnitude or significance level. Therefore, the inclusion of the share of pre-quota 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.

F Placebo tests - artificial cutoffs/year of implementation

In this section, we conduct two additional placebo analyses to support the soundness of our result and corroborate our thesis that Law 215/2012 had these heterogeneous effects with respect to city councils' post quota shares of female councillors. For the first placebo analysis we create two subsamples with only treated or control municipalities and artificially move the threshold of enforcement for the law to either 4000 or 7000 inhabitants. Then, we verify the absence of significant effects on day care-related expenditures. In addition, we perform another placebo test by moving the law's implementation year to 2015 and checking the resulting treatment effects[37]. The first placebo test's results are shown in Tab.14 and 15: in the first Table, the sample is composed of only control municipalities and the fictional cutoff is at 4000 inhabitants, while in the second Table, we have only treated municipalities with a 7000 inhabitants cutoff.

[37] Since it might be argued that these fictional cutoff and year are chosen arbitrarily, we performed additional placebos and observed null heterogeneous effects also with other "fake" cutoffs/years, such as 6000 inhabitants or 2014-2016.

Tab. 14: **Placebo test with artificial population cutoff - control municipalities**

VARIABLES	Baseline interaction Daycarepc	25th percentile Daycarepc	median Daycarepc	75th percentile Daycarepc	90th percentile Daycarepc
TreatPost	0.880	1.592	0.948	1.404	1.075
	(1.874)	(1.045)	(1.304)	(1.150)	(1.122)
TreatPost*Fem. post quota	-1.102				
	(4.595)				
Twenty-fifth p.		5.570***			
		(0.326)			
TreatPost*Twenty-fifth p.		1.425			
		(1.688)			
Median			-3.304*		
			(1.987)		
TreatPost*Median			-0.474		
			(0.836)		
Seventy-fifth p.				-3.313*	
				(1.985)	
TreatPost*Seventy-fifth p.				0.173	
				(0.972)	
Ninetieth p.					-3.317*
					(1.984)
TreatPost*Ninetieth p.					3.111**
					(1.487)
Observations	1,594	1,655	1,655	1,655	1,655
R-squared	0.020	0.022	0.020	0.020	0.024
Number of municipalities	392	419	419	419	419
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

Note: 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-quota 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 et al. (2020), through the minimization of the Mean Squared Error. Year and Time fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

Tab. 15: **Placebo test with artificial population cutoff - treatment municipalities**

VARIABLES	Baseline interaction Daycarepc	25th percentile Daycarepc	median Daycarepc	75th percentile Daycarepc	90th percentile Daycarepc
TreatPost	-3.127	0.004	-1.421	-0.800	-0.438
	(2.104)	(1.210)	(0.985)	(1.169)	(1.145)
TreatPost*Fem. post quota	7.332				
	(5.795)				
TreatPost*Twenty-fifth p.		-2.168			
		(1.339)			
Median			-3.840***		
			(1.454)		
TreatPost*Median			1.746		
			(1.096)		
Seventy-fifth p.				-1.651	
				(1.537)	
TreatPost*Seventy-fifth p.				1.178	
				(1.220)	
Ninetieth p.					-0.816
					(0.845)
TreatPost*Ninetieth p.					-0.277
					(1.496)
Observations	968	993	993	993	993
R-squared	0.037	0.037	0.038	0.035	0.033
Number of municipalities	243	249	249	249	249
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

Note: 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 7000 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-quota 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 7000 inhabitants' threshold is computed following one of the procedures described by Calonico et al. (2020), through the minimization of the Mean Squared Error. The dummy for the twenty-fifth percentile has been dropped due to multicollinearity. Year and Time fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

Results from Tables 14 and 15 present no evidence of significant heterogeneous effect with respect to these "fake" population thresholds, supporting our claim that the results we observe are due to the new legislation introduced.

In the second placebo test, instead of an artificial population threshold, we move the implementation of the law to year 2015, thus considering as "post treatment" period the years 2016, 2017 and 2018. Tab.16 shows this second placebo's results.

Tab. 16: **Placebo test with artificial year of law's introduction**

VARIABLES	Baseline interaction Daycarepc	25th percentile Daycarepc	median Daycarepc	75th percentile Daycarepc	90th percentile Daycarepc
TreatPost	0.734	0.013	0.300	0.557	0.451
	(3.454)	(1.329)	(1.339)	(1.259)	(1.270)
TreatPost*Fem. post quota	-7.892				
	(8.598)				
Twenty-fifth p.		-0.442			
		(1.388)			
TreatPost*Twenty-fifth p.		1.002			
		(1.738)			
Median			1.150		
			(1.178)		
TreatPost*Median			-0.452		
			(0.971)		
Seventy-fifth p.				1.300	
				(1.075)	
TreatPost*Seventy-fifth p.				-1.316	
				(0.984)	
Ninetieth p.					1.279
					(1.071)
TreatPost*Ninetieth p.					-1.187
					(1.025)
Observations	926	2,441	2,441	2,441	2,441
R-squared	0.065	0.031	0.031	0.033	0.032
Number of municipalities	515	528	528	528	528
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

Note: 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-quota 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 et al. (2020), through the minimization of the Mean Squared Error. Year and Time fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

Also this second placebo presents non-significant coefficients for the interaction terms and the treatment variable.

G Gender quotas in 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 it was extended in 2011 for municipalities with at least 3000 inhabitants. The quota imposed a minimum of 40% individual from each sex to be present in candidates' lists. A recent paper (Bagues and Campa (2021)) studied the effect of this quota on a wide variety of outcomes, finding mainly an increase in the percentage of female councillors by 4pp, but no effects on politicians' quality and 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 interesting 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 by 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[38], 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 performed on the Spanish population[39]. We start by replicating our Difference-in-Discontinuity results from Tab.6, and we present them in Tab.17.

[38] This because at 5k inhabitants there is another confounder which is a higher level of transfers received from the central government.

[39] 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.

Tab. 17: **Difference in Discontinuity - Spanish municipalities**

VARIABLES	Baseline interaction F. Expenditures	25th percentile F. Expenditures	median F. Expenditures	75th percentile F. Expenditures	90th 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)				
Twenty-fifth p.		-11.741*			
		(6.603)			
TreatPost*Twenty-fifth p.		-6.788			
		(13.301)			
Median			3.779		
			(6.900)		
TreatPost*Median			-6.473		
			(15.528)		
Seventy-fifth p.				-3.670	
				(8.741)	
TreatPost*Seventy-fifth p.				-66.693***	
				(12.569)	
Ninetieth p.					-3.841
					(8.793)
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

Note: 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 et al. (2020), through the minimization of the Mean Squared Error. Year and Time fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

As we can see from Tab.17, 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 we can see from the column related to the 25th percentile. We can replicate another of our results, the 2SLS regressions in Tab.9. In brief, we instrument the percentage of female councillors with the variable $TreatPost$, and in this way we are able to assess the effect of the quota-related increase in female councillors on female expenditures. In Tab.18

we present these additional results.

Tab. 18: **OLS and second stage regression**

VARIABLES	OLS F. Expenditures	2SLS F. Expenditures	2SLS F. Expenditures
Sh. female councillors	23.491	-734.717***	-510.063***
	(16.166)	(216.847)	(185.687)
treatment	-42.622***		-43.939***
	(8.343)		(9.290)
Observations	32,767	32,767	32,767
Number of codigoine	4,365	4,365	4,365
Muni FE	yes	yes	yes
Year FE	yes	yes	yes

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

Note: Second stage regression in the 2SLS setting, together with OLS base results. Female councillors is instrumented with Treatpost, the interaction between the dummy identifying the treatment group and the dummy for the post-treatment period. The dependent variable is the per capita female expenditures measured at the municipality level. The treatment group includes municipalities with more than 3000 inhabitants, while municipalities with more than 5000 inhabitants have been excluded from the sample as they have been treated for more than one electoral cycle. Year and Time fixed effects are included in each regression. Robust standard errors in parentheses are clustered at the municipality level.

We can see from Tab.18 that the quota-induced increase in female councillors had a negative effect on average on treated municipalities with respect to control ones. The effect is sizeable, as in the most complete specification one percentage point increase in female councillors lead to a reduction of 500 euros per capita in female expenditures in treated municipalities with respect to control ones[40]. Overall, evidence from Tab.17 and Tab.18 tells us that the main result of our paper is not exclusive to the Italian context, but there are also other examples in developed countries where a sudden increase in the share of female politicians provoked a reduction in the most "gender-sensitive" categories of expenditures. We argue that this is always due to the potential conflict triggered by this sudden change in proportion between the majority and the minority relative sizes: in this case, the consequence

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

can be a policy outcome that is more adverse to the minority's preferences.