Fiscal Rules and the selection of politicians: theory and evidence from Italy*

Matteo Gamalerio† Federico Trombetta‡

April 20, 2024

Abstract

Fiscal rules, or constraints on the policymaking discretion of elected officials, are widely used to regulate fiscal policies. Using data on Italian municipalities, we employ a difference-in-discontinuity design to provide evidence of the negative effect of fiscal rules on mayoral candidates' education. Municipalities in which fiscal rules meaningfully restrict the action space of politicians drive the effect. These results are consistent with a formal model of fiscal rules and political selection. We highlight that reducing discretion may affect the composition of the pool of players: it may alleviate pork-barrel spending but also negatively affect the education of politicians.

Keywords: fiscal rules, selection of politicians, deficit, difference-in-discontinuity.

JEL Classification: D72, H62, H70, H72.

*We thank the editor, Erzo Luttmer, and three anonymous referees. Furthermore, we thank Mohit Agrawal, Matteo Alpino, Pablo Beker, Dan Bernhardt, Christopher Berry, Jeremy Boccanfuso, Massimo Bordignon, Fernanda Brollo, Estelle Cantillon, Davide Cipullo, Tommaso Colussi, Stefano DellaVigna, Mirko Draca, James Fenske, Claudio Ferraz, Dana Foarta, Dirk Foremny, Gonzalo Gaete, Juan Carlos Gozzi, Edoardo Grillo, Giorgio Gulino, Stuti Khemani, Eliana La Ferrara, Marco Le Moglie, Nicola Limodio, Alessandro Lizzeri, Ben Lockwood, Daniel Montolio, Helios Herrera, Niall Huges, Mario Jametti, Francesco Nava, Raphael Parchet, Luigi Pascali, Petra Persson, Torsten Persson, Amedeo Piolatto, Giacomo Ponzetto, Francesco Porcelli, Marta Reynal-Querol, Leonzio Rizzo, Edoardo Slerca, Jim Snyder, Pilar Sorribas-Navarro, Albert Sole Olle, Tilko A. Swalve, Silvana Tenreyro, Andrea Tesei, Davide Ticchi, Daniel Treisman, Judit Vall Castello, Stephane Wolton, and all participants at the 2017 EEA-ESEM Congress in Lisbon, the 2016 Marco Fanno Alumni meeting in London, the 2017 RES Congress in Bristol, the 2017 RES Junior Symposium in Bristol, the 2017 Warwick Ph.D. conference, the IEB seminar, the JRC, the University of Lugano, the 2019 MPSA congress, the 2019 APSA congress, the 2019 SIEP conference, the 2019 ZEW Public Finance conference, APSA Formal Theory Virtual Workshop, Università Cattolica, the XV GRASS workshop, the 5th Bruneck-Brunico workshop on Political Economy, the 2022 EPCS conference, and the 2022 Petraila Workshop for their insightful comments. Special thanks to Mario Luca for assisting with the collection of the corruption data. Any remaining mistakes are ours. We did not receive any funding from agencies in the public, commercial, or not-for-profit sectors.

†Institut d'Economia de Barcelona (IEB), University of Barcelona, m.gamalerio@ub.edu.

‡Department of International Economics, Institutions and Development (DISEIS), Faculty of Political and Social Sciences, Università Cattolica del Sacro Cuore, Milano, federico.trombetta@unicatt.it

This paper shows that reducing policymakers' discretion with fiscal rules has the important side effect of negatively impacting the education of candidates willing to run for political office.

Fiscal rules are constraints on policies, and are widely adopted around the world[1] to reduce the incentives of national and local governments to accumulate debt and run deficits. The literature on fiscal rules includes theoretical (Battaglini and Coate, 2008; Halac and Yared, 2014; Azzimonti, Battaglini, and Coate, 2016; Halac and Yared, 2018; Halac and Yared, 2019) and empirical (Grembi et al., 2016; Daniele and Giommoni, 2020) contributions, both of which describe fiscal rules as a reduction in policymaking discretion that involves a trade-off between commitment and flexibility. On the one hand, fiscal rules provide commitment, limiting the incentives toward excessive spending. On the other hand, they impose a cost in terms of reduced flexibility and discretion, confirmed by anecdotal evidence,[2] since they limit governments' ability to respond to shocks.[3] Existing work has mainly addressed the effect of such rules on fiscal stability and economic outcomes (Alesina and Perotti, 1996; Wyplosz, 2012; Grembi et al., 2016).

Here, we provide evidence of a "general equilibrium" effect of fiscal rules. Specifically, rather than focusing on the consequences for fiscal stability, we show that the application of fiscal rules, with the associated reduction in poli-

[1]As Grembi et al. (2016) report, many countries have adopted rules to constrain local governments' fiscal policies in recent years, among them Argentina, Austria, Brazil, Canada, China, Colombia, Czech Republic, Denmark, Italy, Mexico, Poland, Spain, Sweden, and Turkey. One of the most famous examples is the European Stability Pact, which was introduced in 1997 by the European Union and applied to member countries.

[2]For two examples of Italian mayors complaining about the reduction in discretion implied by fiscal rules see Anci (2013a and 2013b).

[3]It is important to point out that this reduction in flexibility can happen, with fiscal rules, even when they are optimal from an ex-ante point of view. As they introduce a constraint, the reduced flexibility may be costly from an ex-post perspective in some states of the world.

cymaking discretion and, potentially, in the value of holding office, can change the type of candidates choosing to run for office, proxied by their education. This is relevant on two different levels. First, the existing literature on rules versus discretion, from Kydland and Prescott (1977) onward, typically studies its trade-offs keeping fixed the pool of players. Our paper shows that it may not be the case: allowing for different levels of rules or discretion affects the type of people willing to enter the game. Second, the education of politicians correlates with measures of policy outcomes (Besley et al., 2011), administrative competence (Carreri, 2020), and public goods production (Martinez-Bravo, 2017), which may be welfare-improving for the voters.[4]

Our main contribution is empirical. We use data on Italian municipalities from 1993 to 2012 to estimate the effect of fiscal rules on the composition of the political class, focusing on mayoral candidates' education. Italy provides an interesting context for study in that, in 1999, the government introduced fiscal rules to limit incentives for accumulating debt and running deficits. These rules initially applied to all municipalities and were introduced under the so-called "Domestic Stability Pact" (DSP). In 2001, the central government removed the rules for all towns with less than 5,000 inhabitants. This relaxation remained in place until 2013 when the cutoff changed from 5,000 to 1,000 inhabitants.

This institutional framework would be ideal for a Regression Discontinuity Design if fiscal rules were the only policy change at the 5,000-inhabitant threshold. However, at the same cutoff, there is also a sharp increase in the

[4]For example, as argued by Besley and Reynal-Querol (2011, p. 563), "given the large amount of evidence of the importance of education in private and public spheres, it would be surprising if there were no relationship between the leaders' education and the quality of policy making". It is important to point out, however, that university education may also measure things that are different from human capital (e.g., socioeconomic status, ideology) and that the evidence on the relationship between the education of politicians and policy outcomes is mixed (see Carnes and Lupu (2016) and the discussion in section I.).

wages paid to the mayor and the municipal aldermen ("assessori"), based on a policy introduced by the Italian government in the 1960s (Gagliarducci and Nannicini, 2013; Grembi et al., 2016; Boffa et al., 2022). This policy thus represents a confounding factor. Indeed, Gagliarducci and Nannicini (2013), using data on Italian municipalities between 1993 and 2001, demonstrate that higher wages attract more educated individuals into politics. Hence, we exploit the 2001 removal of fiscal rules for municipalities below 5,000 inhabitants to estimate a Difference-in-Discontinuity (*Diff-in-Disc*) model, making it possible to measure the effect of fiscal rules on political selection separately from that of the wage increase (Grembi et al., 2016), under the assumption that the effect of the wage increase is constant over time. We find that fiscal rules induce a 9 percentage point reduction in the share of mayoral candidates with a university degree. Furthermore, we find that fiscal rules bring about a similar reduction in the probability of electing a mayor with a post-secondary education.[5]

To explain this empirical evidence and shed light on underlying mechanisms, we propose a simple model of fiscal rules and political selection. Building on the existing theoretical literature on political selection (see Dal Bo and Finan, 2018, for a recent review), our model explicitly incorporates fiscal rules and their well-known flexibility-commitment trade-off (e.g., Halac and Yared, 2014) in a set up where politician types are multidimensional (differing in their education and pro-deficit bias). Fiscal rules avoid biased politicians'

[5]As described in section IV., in the empirical analysis, we consider in the treatment group the municipalities above 5000 inhabitants, which are those that kept fiscal rules after 2001. Having structured the empirical analysis in this way, in the paper, we describe the results as being the effect of introducing fiscal rules, which negatively affect the education of politicians. An alternative interpretation is that removing fiscal rules (as happened to municipalities below 5000 inhabitants in 2001) leads to a positive effect on the education of politicians.

improper use of public money, but they also constrain the fiscal policy choice, impacting the value of holding office for different types of prospective mayors. Since higher-educated politicians are better equipped to understand the state of the economy correctly and more likely to choose the correct policy without fiscal rules, these rules have a greater negative impact on the value they derive from holding office (both in absolute terms and relative to the payoffs from office without fiscal rules). Higher-educated politicians are also unambiguously discouraged from running, and this can increase the running incentive for lower-educated candidates. As a result, fiscal rules are predicted to have an overall negative effect on the education level of the pool of candidates.

Therefore, the model illustrates that an important channel through which fiscal rules affect political selection is heterogeneity in the ability to understand the state of the economy, proxied by education, combined with some degree of policy motivation by politicians and the fact that fiscal rules constitute a restriction to what can be achieved. We can incorporate two important elements of this mechanism into our analysis. First, the mechanism requires fiscal rules to be a meaningful restriction to what can be done once in office. Hence, we should observe a smaller or even zero effect if the action space is already *de facto* restricted, as in the case of municipalities with a more rigid budget. As the literature explains (Grembi et al., 2016; Pavese and Rubolino, 2022; Vannutelli, 2022), municipalities with a large share of rigid expenditures, such as personnel costs and debt repayment, have less policy discretion as these expenditures cannot be adjusted in the short run and reduce the capacity to reallocate resources. Second, the mechanism requires higher educated politicians to be better, on average, at understanding the state of the economy, which may translate into choosing a more appropriate countercyclical fiscal policy. Consequently, when fiscal rules are not in place, there must be a difference between these politicians and less educated ones in terms of fiscal policies. This difference should disappear when fiscal rules are in place.

Both those elements are consistent with the empirical analysis. Specifically, we build a measure for the degree of pre-existing budget rigidity for the municipalities in our analysis. The heterogeneity analysis developed with this measure shows that municipalities with lower pre-treatment rigidity drive the negative effect of fiscal rules on politicians' education. Conversely, the negative effect disappears in the sample of municipalities subject to greater rigidity. Furthermore, a regression discontinuity design based on mixed electoral competitions between mayors with and without a university degree shows that, without fiscal rules, higher educated politicians are more likely to choose a counter-cyclical fiscal policy, which requires the ability to understand the state of the economy. The same analysis shows that higher educated politicians are also more likely to increase investment expenditures and provide more public services without increasing the deficit, *vis-à-vis* less educated politicians. All these differences disappear in places with fiscal rules.

Finally, we show that alternative stories, like a reduction in real terms of the difference in wages paid to mayors or different out-of-politics options for individuals with different levels of education, do not explain the results. Besides, we show the potential effect of fiscal rules on other politicians' characteristics that could correlate with education, such as political experience and ideology,[6] does not explain the results. Indeed, fiscal rules do not affect these variables. In addition, voters do not seem to change their behavior (in terms

[6] In Italy, the difference in vote share for left-wing parties between the top 10% and the bottom 90% of educated voters is around 5 percentage points, while it is 0 for right-wing parties (Gethin et al., 2022). The left vs. right difference in this measure is positive but smaller than many comparable countries (Germany, France, Uk).

of preferences for highly educated candidates) depending on fiscal rules. Consequently, the effect we observe may not be attributed mainly to changes in voters' demand for highly educated politicians induced by fiscal rules. Finally, we show that the empirical results are unlikely to be due to the effect of fiscal rules on corruption estimated in the literature (Daniele and Giommoni, 2020).

I. Related literature

Our study relates to three strands of literature. The first consists of studies on the selection of politicians (Besley, 2005; Braendle, 2016; Dal Bo and Finan, 2018; Krcmaric, 2021; Gulzar, 2021; Carnes and Lupu, 2023), to which we contribute from both a theoretical and empirical point of view. On the theory side,[7] the timing of the model and our assumptions about the information structure are similar to Dal Bo and Finan (2018); the model of candidates' outside option echoes Besley (2004); we assume that high ability politicians have an advantage in performing office-related duties as in Caselli and Morelli (2004); and that there can be a scarcity of high quality politicians, as in Galasso and Nannicini (2011). However, none of these models consider the theoretical implications of fiscal rules on political selection, modeling them as a flexibility-commitment trade-off consistent with Halac and Yared (2014). Like us, Le Borgne and Lockwood (2002) and Izzo (2020) explore the effect of economic dimensions on political selection, though their focus is on political budget cycles and economic crises, respectively.

The empirical literature on political selection has analyzed many different

[7]See, among others, Le Borgne and Lockwood, 2002; Besley, 2004; Caselli and Morelli, 2004; Messner and Polborn, 2004; Mattozzi and Merlo, 2008; Galasso and Nannicini, 2011; Mattozzi and Merlo, 2015; Galasso and Nannicini, 2017; Dal Bo and Finan, 2018; Izzo, 2020.

institutions that potentially affect the selection of politicians.[8] However, to the best of our knowledge, the role of fiscal rules has not been addressed. We contribute to this body of work by showing how a reduction in policymaking discretion and the value of holding office due to the application of fiscal rules can negatively affect the quality of the political class. Our results parallel those of Gagliarducci and Nannicini (2013) who, using data from Italian municipalities around the 5,000-inhabitant threshold from 1993 to 2001, show how higher wages paid to politicians can attract more competent individuals into politics. Our empirical analysis suggests that a reduction in policymaking discretion due to fiscal rules can offset the positive effect of higher pay for local politicians. These findings imply that, while higher pay for politicians may help to attract skilled individuals, their decision to enter politics depends on several different factors.

Our paper also intersects with work analyzing the effect of fiscal rules on fiscal stability and more general economic and political outcomes. The results of these studies are mixed, with some (Alesina and Perotti, 1996, and Wyplosz, 2012) noting that fiscal rules may not work for reasons of commitment. Grembi et al. (2016) offer a recent contribution along these lines. They use data from Italian municipalities to show that fiscal rules can effectively reduce the deficit run by local governments. Specifically, they show that municipalities where fiscal rules are relaxed increase their deficit by 20 euros per capita compared

[8]The institutions and determinants studied include the wage paid to politicians (Besley, 2004; Ferraz and Finan, 2011; Gagliarducci and Nannicini, 2013; Kotakorpi and Poutvaara, 2011; Dal Bo et al., 2013; Fisman et al., 2015; Braendle, 2015), the role of outside earnings (Gagliarducci et al., 2010; Fedele and Naticchioni, 2013; Grossman and Hanlon, 2013), the role of monitoring institutions (Grossman and Hanlon, 2013; Artiles et al., 2020), the level of fiscal autonomy (Brollo et al., 2013; Peralta and Pereira dos Santos, 2018; Bordignon, Gamalerio, and Turati, 2020), political parties (Cervellati, Gulino, and Roberti, 2021), electoral rules (Beath et al., 2015; De Benedetto, 2018; Gulino, 2020), gender quotas (Baltrunaite et al., 2014), voter turnout (Lo Prete and Revelli, 2021), disclosure laws (Fisman, Schulz, and Vig, 2019), and criminal organizations (Daniele and Geys, 2015).

to those where fiscal rules are maintained. They also show that treated municipalities reduce taxes and that the effect is driven by municipalities where the mayor can be re-elected, with more parties in the legislative assembly and an older population. Daniele and Giommoni (2020) document that fiscal rules reduce corruption.[9] Carreri and Martinez (2021) also examine fiscal rules' political outcomes. However, they focus on a different type of fiscal rule (the so-called "golden rule") in a different context (Colombia) and with different outcomes (support for the party of the incumbent mayor and protests against the municipal government). We contribute to these studies by investigating an unexplored consequence of fiscal rules on the selection of politicians.

Third, and more broadly, our study speaks to the literature on incentives and selection into the public sector[10] and beyond,[11] with a focus on intrinsic incentives. Deserrano (2018) observes that higher wages may deter prosocial candidates from applying for a government job; Ashraf et al. (2019) find similar results in terms of career perspectives. Bartling et al. (2012) experimentally document the complementarity between the possibility of screening employees' past performance and the amount of on-the-job discretion allowed by employers. The quasi-experimental setting of our paper means that we can explore the direct effect of a reduction in agents' discretion on the quality of "applicants" for public office, holding financial returns fixed. We show that more discretionary power attracts more highly educated candidates, at least for this type of "executive" job. In this respect, the relevance of intrinsic in-

[9]Other papers in the literature have exploited the Italian context to study the consequences of fiscal rules on outcomes such as firms' dynamics (Coviello et al., 2021), the political budget cycle (Bonfatti and Forni, 2019), and distributional policies (Alpino et al., 2020). A study by Revelli (2016) examines the impact of tax limits on electoral turnout and local election outcomes. Finally, Vannutelli (2022) shows that fiscal rules are even more effective when enforcement involves independent auditors.

[10]See Finan et al. (2015) for a review focused on field experiments.

[11]See Oyer and Schaefer (2010) for a review.

centives highlighted herein is consistent with the findings of Gulzar and Khan (2023), who observe that increasing the salience of prosocial incentives motivates prosocial people to run for office and behave in ways more aligned with citizens' preferences.

We use the level of education of mayoral candidates and elected mayors as our main outcome variable. Various studies justify this focus on education, showing that the latter positively affects socio-economic outcomes such as wages (Card, 1997) and measures of citizenship (Dee, 2004). University-educated people are also more likely to engage in politics and work on community issues (Milligan et al., 2004). Of particular relevance to our paper, previous work documents that electing more educated political leaders can positively affect economic growth (Besley et al., 2011), the production of public goods (Martinez-Bravo, 2017), and fiscal sustainability (Meriläinen, 2021). Education is extensively used as an indicator of competence in the political selection literature (e.g., Besley and Reynal-Querol, 2011; Gagliarducci and Nannicini, 2013; Galasso and Nannicini, 2011), and it positively correlates with administrative competence (Carreri, 2020). Daniele and Giommoni (2020) observe that fiscal rules reduce corruption to a greater extent in municipalities with more educated mayors. Additionally, Mitra (2020) shows that educated mayors increase public investment in education without worsening the municipality's financial situation. Notably, Boffa et al. (2022) observe that higher education levels in local politicians correspond with decreased support for anti-establishment populist parties.

However, it is important to point out that the available evidence on the relationship between politicians' formal education and the outcomes they produce once in power is not unambiguous. In particular, Carnes and Lupu (2016) show that the main result of Besley et al. (2011) is weak and does not extend to other outcomes beyond GDP growth. Furthermore, they use RDD analysis to show that college education does not correlate with more legislative productivity or less corruption. They argue that "schooling is an imperfect measure of competence, and general competence alone may not be enough to make someone a good politician" (p. 47). However, in the absence of better available measures of human capital, we argue that looking at changes in the education level of candidates allows us to proxy changes in the composition of the pool of politicians that may be relevant for how policymaking is implemented and ultimately, for voters' welfare. More broadly, we see education as a useful proxy for the type of candidates choosing to run for office, and we see that the level of discretion affects this pool in important ways.

Finally, we look specifically at elected mayors and mayoral candidates for two reasons. First, as described in Section III.A., Italian mayors are powerful at the municipal level. Second, the seminal paper by Besley (2005) suggests that the scope of authority enjoyed by elected politicians affects the selection of directly elected chief executives such as presidents, governors, and mayors to a greater extent than politicians in positions with less direct power. Thus, we can expect a reduction in policymaking discretion to significantly affect politicians in powerful positions, like mayors, rather than politicians in less prominent positions, like municipal councilors.

II. Theoretical framework

II.A. Model set up

There is a large number n of municipalities. Each has a representative voter V and two potential candidates affiliated with political parties. As in Dal Bo and

Finan (2018), we use the general term "politicians" to indicate party members who may be chosen to run for office. "Candidates" are those that accept to run and "perspective" or "potential" candidates are those picked by parties but observed before their decision.

II.A.I Politician types and payoffs

Politicians' education level is denoted by $\Gamma \in \{H, L\}$ and is observable at the election stage. Higher education implies a better understanding of the state of the world: both high and low education politicians receive an informative signal about the state of the economy, but one is more precise than the other.[12]

When in office, politicians derive an office rent $E > 0$ capturing the direct motivation of the office (e.g., salary) and a policy-related utility, weighted by $k > 0$. Specifically, some politicians are biased in favor of spending, meaning that they receive a payoff of 1 when they choose deficit spending. We denote their bias $b \in \{0, 1\}$, with the common prior $Pr(b = 1) = \tau \in (0, 1)$. We assume that the bias is not correlated with education,[13] implying that discretion may be valuable for both high and low education politicians, when they are biased toward public spending. Moreover, politicians learn their bias once in office. This assumption simplifies the game, since the entry decision might otherwise become a signalling game. Several political attitudes are contained in this idea of bias toward deficit spending: present-biased politicians (Halac and Yared (2014), Piguillem and Riboni (2015)); dishonest politicians using deficit-financed funds for private purposes; politicians able to manipulate vot-

[12]The results are qualitatively unchanged if we assume instead that signal precision depends on an underlying unknown level of true ability, positively correlated with the education level.

[13]Appendix E.4.2 shows under what conditions our result holds when education and spending bias are correlated.

ers' views about their competence (Murtinu et al. 2021); pork-barrel spending or political budget cycles.[14] Clearly, some of those incentives are related with office-motivation and re-election concerns, and they may interact with the level of education of politicians in ways that are not fully captured here. A more complete model should take them into account explicitly. In this one we choose to privilege simplicity capturing all those issues under the idea of (exogenous) pro deficit bias. The remaining politicians are unbiased, meaning that they want to choose the correct policy with no pre-existing preference for or against deficit spending. Finally, note that τ can also be interpreted as the probability that, in any given municipality, the incentive structure is such that running a budget deficit is rewarded, *vis-à-vis* choosing the correct policy (for example, because political budget cycles are very effective).

II.A.II State of the economy and policies

The economy is summarized by a binary state of the world $\theta \in \{0, 1\}$ where $\theta = 0$ implies that the budget should be balanced and $\theta = 1$ implies that there should be deficit spending.[15] Assume $Pr(\theta = 1) = p \in (0, 1)$.

There are 2 possible actions: $x \in \{0, 1\}$, where $x = 0$ denotes a balanced budget and $x = 1$ is deficit spending. The voter's payoff is $u^V = \begin{cases} 1, & \text{if } x = \theta \\ 0, & \text{otherwise} \end{cases}$.

Politicians receive a signal s with realizations $\{0, 1\}$ such that $Pr(s = \theta|\theta, \Gamma) = \phi^\Gamma$, where $1 > \phi^H > \phi^L > max[p, 1-p]$. This implies that the higher educated politician has a better understanding of the state of the world than

[14] See Yared (2019) for a recent taxonomy of political economic reasons behind pro-deficit biases in democracies.

[15] For example, in response to a negative economic shock.

the less educated one, and offers a way of capturing the advantage of more competent politicians once in office, consistent with Caselli and Morelli (2004).

II.A.III Running decision

In each municipality, there are two political parties. One member from each party is selected to become a perspective candidate. For simplicity, we assume that every party is independently able to find a high education perspective candidate with probability 0.5, and this probability is known to the players.[16] Simultaneously, and without observing the education level of the opponent, the selected perspective candidates choose whether to accept and run, as in Brollo et al. (2014) and Dal Bo and Finan (2018). If neither of the two party members run, a default mayor is in place.[17]

If the chosen politician decides not to run, she keeps her salary w^i, private information of potential candidate i. We assume that w^i is drawn from a uniform distribution on $[0, W^\Gamma]$, where $W^H > W^L$ are assumed to be sufficiently large to ensure an interior solution (i.e., $W^L > E + k$). This means that high education politicians have, on average, higher salaries in the private sector. Finally, we set the payoff of running and losing to zero.

Formally, we define $d^i \in \{r, nr\}$ as the decision to run or not by politician i, whose party has selected her to run. γ^i_Γ is her probability of winning the election (in equilibrium, this is endogenous as it depends on the decision of the opponent and the latter's education level). As a consequence, if i is highly

[16] It is known that parties may select their candidates for reasons other than their education level or ability (Mattozzi and Merlo, 2015), or they may have a limited supply of high education candidates (Galasso and Nannicini, 2011).

[17] This particular timing is identical to Dal Bo and Finan (2018), with the exception of the default mayor (in their case, parties draw candidates again if nobody accepts to run). However, the results are qualitatively unchanged if we assume that first potential politicians self-select into joining one of the parties and then parties randomly select one of them.

educated and runs for office, we have the following expected payoff:

$$u^i(H, d^i = r) = \gamma_H^i \left(E + k\mathbb{E}_{b,\theta,s} u_H^P\right) \tag{1}$$

The term in parenthesis is the payoff from office in case of victory: politician i receives the office rent E and the expectation of the policy-related utility u_H^P with respect to her bias, the state of the world and the signal realization. The subscript H captures the fact that high education politicians have a more precise signal of the true state of the economy. u^P is equal to u^V when $x = \theta$ if $b = 0$ and to 1 when $x = 1$ if $b = 1$, and 0 otherwise. $k > 0$ measures the relative importance of the policy motivation. On the other hand, if politician i chooses not to run once selected, she receives

$$u^i(H, d^i = nr) = w^i \tag{2}$$

In the case of a potential candidate with low education we have

$$u^i(L, d^i = r) = \gamma_L^i \left(E + k\mathbb{E}_{b,\theta,s} u_L^P\right) \tag{3}$$

$$u^i(L, d^i = nr) = w^i \tag{4}$$

The difference between (1) and (3) lies in how informative the signal is in each case. In equilibrium, all unbiased politicians follow it, but H politicians are more likely to choose the correct policy.

II.A.IV Modelling fiscal rules

We model fiscal rules as a restriction on the action space of the incumbent in order to capture in the simplest possible way the flexibility-commitment trade off proposed by Halac and Yared (2014), among others. In particular, we assume that when fiscal rules are introduced, the action space is reduced to $x = 0$, i.e. politicians cannot run budget deficits. This implies that some flexibility is lost (i.e., an incumbent cannot choose $x = 1$ when the state of the economy requires it), but also that a biased politician cannot choose $x = 1$. This implies that both H and L politicians are damaged by fiscal rules when they are biased toward spending, because vote buying, pork-barrel policies and political budget cycles are no longer possible. As such, voters are indifferent in their electoral preferences for high and low education politicians when fiscal rules are in place. The way we break this indifference is not crucial for our result. However, in order to capture the fact that education can be correlated with in-office performance on other issues separate from the budget policy, we assume that voters, if otherwise indifferent, choose the high education candidate.[18]

II.A.V Timing and solution concept

The game is one shot. The timing is as follows:

1. One politician per party is selected to run for office in each municipality. They simultaneously decide whether to run or not, without knowing the education level of the opponent.

2. Candidates' education level is revealed to voters. If there are two candidates, voters vote sincerely. If there is only one candidate, she wins directly. If there are no candidates, a default mayor is in place.

[18] The result is qualitatively unchanged if, for example, V mixes with equal probability when indifferent among candidates of different type.

3. The winning politician privately learns b. Then she observes s and chooses x.

4. u^V is realized, payoffs are paid and the game ends.

Our solution concept is a symmetric Perfect Bayesian Nash Equilibrium (PBNE). As a tie-breaking rule, we assume that indifferent politicians choose to run and that the voter mixes with equal probability when indifferent between two candidates of the same education level.

II.A.VI Probability of highly educated candidate

We are interested in the effects of fiscal rules on ex ante selection into politics. One way to measure this is to look at the probability of a candidate being highly educated, i.e.,

$$\hat{\lambda} := Pr(\Gamma^i = H | d^i = r) = \frac{Pr(d^i = r | \Gamma^i = H)0.5}{Pr(d^i = r)} = \frac{1}{1 + \frac{p_L}{p_H}} \tag{5}$$

where we define p_H (p_L) as the probability that a randomly picked higher (less) educated member of a party chooses to run. This is also equivalent to the expected share of high education candidates in any given municipality, conditional on having at least one candidate running.

Obviously, both p_H and p_L are determined in equilibrium and are influenced by the presence or absence of fiscal rules.

II.B. Analysis

In this section, we present a shorter and informal examination of the game, leaving the formal analysis and all the relevant proofs for Appendix E.1.

First, note that without fiscal rules the voter anticipates that each type of politician will choose her own individually optimal strategy, following the signal if unbiased, and she thus prefers to elect the high education candidate, if available. Candidates anticipate this and compare their payoffs from running or not running. In other words, they compare the respective expected utilities as described by equations (1) and (2) for the H education type, and equations (3) and (4) for the L type, taking into account the equilibrium behavior. This implies that, in the PBNE without fiscal rules, the expected policy related payoffs for politicians choosing to run is:

$$\mathbb{E}u_{\Gamma}^{P} = ((1-\tau)\phi^{\Gamma} + \tau) \tag{6}$$

In either case, the running decision is captured by a threshold in w^i, such that only individuals with a private sector salary below the threshold decide to enter politics. Those thresholds vary by education level and depend on the conjectured threshold strategies of different types of opponents. In a symmetric equilibrium, those conjectures are correct and politicians of the same type follow the same strategy. Thus, equilibrium thresholds are the solution of the following system:

$$\bar{w}_H = \left(1 - \frac{1}{2}\frac{1}{2}\frac{\bar{w}_H}{W^H}\right)\left(E + ((1-\tau)\phi^H + \tau)k\right)$$
$$\bar{w}_L = \left(1 - \frac{1}{2}\frac{1}{2}\frac{\bar{w}_L}{W^L} - \frac{1}{2}\frac{\bar{w}_H}{W^H}\right)\left(E + ((1-\tau)\phi^L + \tau)k\right)$$

The interpretation of the thresholds is straightforward. All politicians win when running unopposed. H politicians know that, if they run, they are sure to win against a low education opponent and they win with probability 0.5 against an high education opponent. On the opposite side, L politicians know that they will lose against an H opponent and win with probability 0.5 against a L opponent. The rest is their expected payoff from being in office. As a consequence, in equilibrium, $p_H = \frac{\bar{w}_H}{W^H}$ and $p_L = \frac{\bar{w}_L}{W^L}$. The thresholds can be derived in a closed form as follows:

$$\bar{w}_H = \frac{4W^H\left(E + ((1-\tau)\phi^H + \tau)k\right)}{4W^H + (E + ((1-\tau)\phi^H + \tau)k)} \tag{7}$$

$$\bar{w}_L = \frac{4W^H - \left(E + ((1-\tau)\phi^H + \tau)k\right)}{4W^H + (E + ((1-\tau)\phi^H + \tau)k)} \frac{4W^L\left(E + ((1-\tau)\phi^L + \tau)k\right)}{4W^L + (E + ((1-\tau)\phi^L + \tau)k)} \tag{8}$$

When fiscal rules are in place, all politicians in office are constrained to choose $x = 0$, therefore (6) becomes $(1-\tau)(1-p)$. As a consequence, the relevant thresholds are:

$$\bar{w}_H^{FR} = \frac{4W^H\left(E + (1-\tau)(1-p)k\right)}{4W^H + (E + (1-\tau)(1-p)k)} \tag{9}$$

$$\bar{w}_L^{FR} = \frac{4W^H - (E + (1-\tau)(1-p)k)}{4W^H + (E + (1-\tau)(1-p)k)} \frac{4W^L\left(E + (1-\tau)(1-p)k\right)}{4W^L + (E + (1-\tau)(1-p)k)} \tag{10}$$

We are now in a position to compare the probability that a candidate is highly educated with and without fiscal rules. Proposition 1 summarizes our findings:

Proposition 1. *The probability that a candidate is highly educated is higher without fiscal rules.*

Proposition 1 (and its proof) provides several insights. First, note that, because of the uniform distribution of w, we can focus on the comparison of the ratios between $\bar{w}_H$ and $\bar{w}_L$, with or without fiscal rules.

Second, fiscal rules have two distinct effects, both of which lead to a decrease in the probability that a candidate is an H type. The first effect is a reduction in the expected policy payoff from office for unbiased politicians, which shuts down the advantage H candidates have in matching the state of the world. Note that without fiscal rules, unbiased politicians can match the state of the world with probability $\phi^H > \phi^L$. With fiscal rules, this happens with probability $1-p$ irrespective of the education level. Hence, the reduction is relatively larger for H politicians, and the ratio between expected payoffs from being in office for H over L types is higher without fiscal rules, implying that the overall probability that a candidate is H should be higher without fiscal rules.[19] The second, related, effect originates from strategic considerations related with the probability of victory. H politicians are less likely to run with fiscal rules in place. This, in turn, increases the incentive of L politicians to run, as they can win by running unopposed or against an opponent of the same type.

Third, the model highlights that both effects are simultaneously at play,[20] and the importance of the first effect, which results from the different ability and policy-related incentives. As shown in Appendix E.4, if we assume purely office-motivated politicians the effect of fiscal rules disappears.

[19] To capture the first effect of fiscal rules on $\hat{\lambda}$ one should focus on the ratio between the payoffs from holding office for high over low education politicians. When fiscal rules cause a reduction in this ratio (meaning that their cost, as a share of the office payoff without fiscal rules, is higher for H candidates), they are reducing $\hat{\lambda}$ (see Online Appendix E.2). In our model, the fact that fiscal rules lead to the same expected payoffs from being in office for both H and L politicians implies that the effect of the reduction in expected payoffs from office for L types, also caused by fiscal rules and that would increase $\hat{\lambda}$ in isolation, is always dominated. If payoffs had other education-specific components, the effect would be there when the ratio of payoffs from office is higher without fiscal rules than with fiscal rules.

[20] If we shut down considerations related with changes in the probability of victory, the effect of fiscal rules would still be there when $\phi^H > \phi^L$ (see Corollary E3). If we assume $\phi^H = \phi^L$, the negative effect of fiscal rules survives, combined with the assumption that the voter chooses the H type even when two candidates of the same education level give the same expected payoffs.

II.B.I Additional implications

Along with providing a rationale for our main result, the model suggests an important mechanism for the observed effect. This relies on fiscal rules meaningfully restricting the set of available policies, combined with a differential probability to understand the correct state of the economy between higher and less educated politicians once they take office. Both of these implications are consistent with the empirical analysis.

Rigid municipalities Consider the case of a municipality where the discretion in policymaking is limited by the high share of rigid, pre-existing expenditures in her budget (i.e., personnel and debt repayment, similar to Grembi et al., 2016). As they cannot be easily adjusted in the short run, it is more difficult for the incumbent mayor to react to a negative shock. The role of rigid expenditures in reducing policymaking discretion has been pointed out by Grembi et al. (2016) and Pavese and Rubolino (2022). Anecdotal evidence is also in Openpolis (2016).[21] In such cases, the action space is effectively constrained to $x = 0$ irrespective of fiscal rules, meaning that fiscal rules do not meaningfully restrict the action space of politicians in office. Hence, if the channel implied by the model does drive politicians' behavior, we should observe no effect of fiscal rules on the composition of the candidate pool in rigid municipalities. Where discretion is already severely limited, there is no further differential impact on the expected value of holding office.[22]

Education level and policy choice The model's conclusions are partly driven by higher educated politicians being more able to match the state of

[21]See section IV.C.II for more details on our measure of spending rigidity.

[22]A similar logic holds even if rigidity implies that municipalities are constrained to choose $x = 1$. Appendix E.3 discusses this case formally.

the world, which they are (sometimes) willing to do. However, fiscal rules may restrict them from doing so. In the realm of the model, this implies that, without fiscal rules, higher educated politicians are more likely to choose a counter-cyclical fiscal policy, on average. More generally, they should be more likely to be efficient administrators when unconstrained by fiscal rules. However, this difference should disappear once fiscal rules are in place.

III. Institutional Setting

III.A. Italian municipalities

There are around 8000 municipalities in Italy. Municipalities are responsible for municipal police, infrastructure, transport, welfare, housing, garbage collection, and water supply. They manage 10% of total public expenditures, with an expenditure per capita approximately equal to 1856 euros (years 2015-2017, data from Aida Pa, Bureau Van Dijk), and around 20% of their revenues coming from local taxes. The rest of the revenues are discretionary transfers from higher levels of government, like provinces, regions, and the central state. The most important local taxes include property tax, introduced in 1993 by Legislative Decree 504/1992, and a surcharge on residents' income tax, introduced in 1999. Since 1993 (see Law 81 of 1993), mayors of Italian municipalities are directly elected by voters. In municipalities below 15,000 inhabitants, mayors are elected using a single round plurality rule, while a run-off system is employed above this threshold. Mayors serve for a period of five years and, since 1993, for a maximum of two consecutive terms. They play a powerful role in municipal governments, as they can choose and dismiss the aldermen that form part of the municipal government. Furthermore, if the municipal council decides to dismiss the mayor, new elections must be held.

III.B. The "Domestic Stability Pact" (DSP)

Fiscal rules for Italian municipal governments were introduced in 1999, following the European Stability and Growth Pact (SGP), signed in 1997 by various European countries.[23] In Italy, these rules were called the "Domestic Stability Pact" (DSP) (In Italian, Patto Interno di Stabilita'). The DSP was introduced through Law 448 of 23 December 1998, Article 28. The goal of the DSP was to reduce the incentives for local governments to accumulate debt and run deficits. Table 1 describes how the target imposed by the DSP changed between 1999 and 2015. We see that the target has not been constant over time, though, with the exception of 2005-2006, local governments did have to balance their budget each year.[24] The initial penalties introduced by the central government for not complying with the rules consisted of a 5% cut

[23]De Biase and Dougherty (2022) use a dataset of 27 European Union members to show that budget balance rules are the most common type of fiscal rules among local governments of European countries, while expenditure rules (such as the one considered in Carreri and Martinez, 2021) are more common amongst central governments. Looking at OECD countries, budget balance rules are the most common among state governments. Among local governments, they were the most common type of fiscal rule in 2006 and 2011, becoming second slightly behind borrowing constraints in 2018. De Biase and Dougherty (2022) also coded a "fiscal rules strength index" (FRI) based on European Commission's data. The index, between 0 and 10, captures the strength of five institutional features of fiscal rules: statutory/legal base of the rule, nature of the body in charge of rule monitoring and the correction mechanism, nature of the body in charge of monitoring correction of deviations from the rule, correction mechanism in case of deviation from the rule and resilience to shocks or events outside the control of the government. While the median strength for budget balance rules among local governments is always below 5, local fiscal rules in Italy are "among the highest FRI scores for any fiscal rule for local governments" (De Biase and Dougherty, 2022, p. 31), moving between 6.92 and 7.69.

[24]The definition of balanced budget used in the target for most years has been based on the so-called fiscal gap (Grembi et al., 2016), or the municipal deficit net of transfers and debt service. The limits on the target have varied: in some years, municipalities were asked to apply a cap to the growth of the target; in other years, they were asked to cut the target. These limits have always been calculated with reference to past values of the target in specific reference years.

in grants transferred by the national government, a cut in reimbursement and non-absenteeism bonuses for municipal employees, and a ban on new municipal hires. Municipalities complying with the rules were rewarded with decreased interest expenses on loans received from the central government. In 2008, as described by Coviello et al. (2021), harsher penalties for not complying with the rules were instituted, including an increased cut to central government grants and an automatic 30% cut to the salaries of mayors and municipal councilors.

As can be seen in Table 1, fiscal rules applied to all municipalities for the first two years (1999-2000). In 2001, the central government then removed the fiscal rules for all the municipalities below 5,000 inhabitants in order to lift onerous constraints on places that were disadvantaged by economies of scale. Specifically, the law (Legislative Decree no. 267, article 156) identifies the municipalities subject to fiscal rules based on the number of residents as measured by the Italian Statistical Office (Istat) at the end of the second-most recent year (e.g., for the year 2002, the population figure from December 31, 2000 was used as a reference). In 2002, the Regions with Special Statute (i.e., Sardinia, Sicily, Valle d'Aosta, Trentino-Alto Adige, Friuli-Venezia Giulia) were allowed to establish their own fiscal rules; we accordingly exclude municipalities in these regions from the analysis. The 5,000-inhabitant threshold remained in place until 2013.[25] In that year, the threshold was reduced from 5,000 to 1,000 inhabitants for 2013-2015 (Daniele and Giommoni, 2020). Finally, the DSP

[25] At the end of 2004, Budget Law 311/2004 extended the application of fiscal rules to municipalities between 3,000 and 5,000 inhabitants for the year 2005. However, Decree-Law 44/2005 and Law 88/2005 canceled this provision, such that the threshold remained at 5,000 inhabitants. Budget Law 266/2005 confirmed the 5,000-inhabitant threshold for 2006, and Budget Law 296/2006 confirmed it for 2007. The threshold then remained unchanged until Budget Law 183/2011 extended the application of fiscal rules to all municipalities with more than 1,000 inhabitants starting from the year 2013.

Table 1: Fiscal rules in Italy: the Domestic Stability Pact (DSP)

Year	Target	Reference Year	Covered municipalities
1999	Budget Balance	1997	All
2000	Budget Balance	1998	All
2001	Budget Balance	1999	> 5000
2002	Budget Balance	2000	> 5000
	Current Expenditures	2000	
2003	Budget Balance	2001	> 5000
2004	Budget Balance	2003	> 5000
2005	Total Expenditures	2002-2004	> 5000
2006	Current Expenditures	2004	> 5000
	Capital Expenditures	2004	
2007	Budget Balance	2003-2005	> 5000
2008	Budget Balance	2003-2005	> 5000
2009	Budget Balance	2007	> 5000
2010	Budget Balance	2007	> 5000
2011	Budget Balance	2006-2008	> 5000
2012	Budget Balance	2006-2008	> 5000
2013	Budget Balance	2007-2009	> 1000
2014	Budget Balance	2009-2011	> 1000
2015	Budget Balance	2010-2012	> 1000

Notes. Domestic Stability Pact: fiscal rules enacted by the Italian central government, which apply to the covered municipalities on an annual basis. Columns: "Year" = year in which the rules were applied; "Target" = target decided by the central government for a specific year. The limits on the target decided by the central government are imposed with respect to specific past reference years, which are reported in the column "Reference years"; "Covered municipalities" = this indicates the municipalities that must apply the fiscal rules based on their resident population measured at the end of the second-most recent year. Legislative sources: annual national budget law (Legge Finanziaria) from 1999 to 2015. Other sources: Grembi et al. (2016); Chiades and Mengotto (2013). As described by Grembi et al. (2016), the main definition of budget balance used during the years has been the so-called fiscal gap, which is defined as municipal deficit net of transfers and debt service.

was abolished in 2016 and replaced by a new set of balanced budget rules for all municipalities.

Table 2 reports the legislative population thresholds applicable to municipalities with less than 15,000 inhabitants. We observe that the wages paid to the mayor and aldermen in the municipal government change at the 5,000-

inhabitant threshold. This wage increase is a policy that dates back to the 1960s (Gagliarducci and Nannicini, 2013), which has remained constant in real terms until today. The mayor's pay is based on the population recorded in the last available population Census.

Table 2: Legislative population thresholds in Italy: Municipalities below 15,000

Population	Wage Mayor	Wage Aldermen	Size Government	Size Council
< 1000	1,291	15 %	4	12
1000-3000	1,446	20 %	4	12
3000-5000	2,169	20 %	4	16
5000-10,000	2,789	50 %	4	16
10,000-15000	3,099	55 %	6	20

Notes. Legislative population thresholds that apply to Italian municipalities with less than 15,000 inhabitants. Columns: Population = municipal population as measured by the last Census; Wage Mayor = the wage paid to the mayor, expressed in Euros per month at 2000 prices; Wage Aldermen = wage paid to the aldermen as a percentage of the wage of the mayor; Size Government = maximum number of aldermen that can be appointed in the municipal government; Size Council = number of seats in the municipal council. All wage thresholds date back to 1960, except the 1,000 and 10,000 thresholds, which were introduced in 2000. Sources: Gagliarducci and Nannicini (2013); Grembi et al. (2016).

IV. Empirical Evidence

IV.A. Empirical Strategy: Difference-in-discontinuity

To study the impact of fiscal rules on political selection, we use the variation over time in the application of fiscal rules around the 5,000-inhabitant threshold. A standard regression discontinuity design (RDD) focusing on electoral terms from 2001 to 2012 would be invalidated by differences in mayoral pay across this threshold. To circumvent this issue, we exploit the 2001 removal of fiscal rules for municipalities below the threshold and adopt a Difference-in-Discontinuity (*Diff-in-Disc*) method, as suggested by Grembi et al. (2016).

The *Diff-in-Disc* approach is a strategy (Lalive, 2008; Campa, 2011; Leonardi and Pica, 2013; Casas-Arce and Saiz, 2015; Grembi et al., 2016) that combines the *pre/post treatment* variation typical of a Difference-in-Differences design with the variation *around a threshold* that characterizes an RDD approach. In the Italian context, adopting this strategy allows to estimate the effect of fiscal rules on the selection of politicians while controlling for differences in mayoral wages, which are constant over time in real terms.

We estimate the following empirical model, using data at the municipality and electoral year level:

$$\begin{aligned} Y_{it} = \rho_0 + \phi_0(> 5000_{it}) * (Post_t) + \beta_0(> 5000_{it}) + \pi_0(Post_t) + \\ R_{it} * [\rho_1 + \phi_1(> 5000_{it}) * (Post_t) + \beta_1(> 5000_{it}) + \pi_1(Post_t)] + \eta_{it} \end{aligned} \tag{11}$$

where Y_{it} is the level of education of politicians. The variable $R_{it} = P_{it-1} - 5000$ is the normalized population, which measures the distance of municipality i from the 5,000-inhabitant threshold at time t. P_{it-1} denotes the population used for the application of fiscal rules. The dummy variable $(> 5000_{it})$ is equal to 1 if municipality i is above the 5,000-inhabitant threshold, while the dummy variable $(Post_t)$ is equal to 1 for elections starting from 2001. As the selection of (new) politicians can only happen during election years, we estimate the model using data at the election year level, unlike Grembi et al. (2016), who use yearly data. The treatment variable is the interaction term between $(> 5000_{it})$ and $(Post_t)$. The coefficient of interest is ϕ_0, which captures the effect of fiscal rules on the selection of politicians, comparing municipalities that continue to apply fiscal rules and those that became exempt starting in 2001. The electoral years from 2001 on comprise the treatment period as this is the year when fiscal rules started to apply differently across the 5,000-

inhabitant threshold. Therefore, we can expect differential behavior in terms of political selection due to fiscal rules to emerge across the cutoff only from 2001.

We estimate model (11) with a local linear regression (Gelman and Imbens, 2018), using the subsample of observations that lie within the interval $R_{it} \in [-h, +h]$ around the threshold, where the optimal bandwidth h is calculated using the MSE-optimal bandwidth following Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). We cluster standard errors at the local labor market area level.[26] In the Appendix section D.1, we provide additional details on how in practice we estimate model (11).

This identification strategy requires three main assumptions, which we test in the analysis below. First, there must not be manipulative sorting of the running variable R_{it} around the 5,000-inhabitant threshold before and after 2001, such that municipalities must not be able to self-select themselves and decide which side of the cutoff they wish to stay on. We test this assumption, and we do not find any discontinuity in the density of R_{it} at the 5,000-inhabitant threshold (see Figure A1 in the appendix).[27] Second, other potential outcomes and municipal characteristics must be balanced around the threshold before and after 2001. We test this assumption by running model (11) using municipal characteristics as dependent variables. We find that municipal and geographical characteristics are balanced around the threshold before and after 2001 (see Table A2 in the appendix). Third, municipalities just below and

[26]Local labor market areas are geographical units composed of neighboring municipalities that share the same local labor market and exhibit common socio-economic and population characteristics. In our dataset, we were able to identify 517 local labor areas using information from the 2001 Census. The results are robust to clustering the standard errors at different levels (e.g., municipal or provincial levels).

[27]We apply this test to R_{it} measured in the years before and after 2001. We also test the continuity of the difference between the density of the average population before and after 2001.

just above the 5,000-inhabitant threshold must be on parallel trends before the 2001 relaxation, as is typical in a difference-in-differences analysis. We test this assumption in Section IV.C.I.

Finally, an additional requirement is necessary for employing the Diff-in-Disc model in the Italian context. As described in section III.B., the populations of reference for applying fiscal rules and deciding the mayor's wage are not the same, even though the correlation between the two is 0.97. Therefore, there may be cases of municipalities above 5000 inhabitants according to one population but not the other. Hence, the use of *Diff-in-Disc* is justified if the indicator for the higher wage jumps discontinuously at the 5000-inhabitant threshold calculated using the population that decides the application of fiscal rules. In addition, this discontinuous jump in the wage indicator should not change before and after the 2001 fiscal rules relaxation. We run model (11) using the high wage dummy variable as the dependent variable, and we find that these requirements are confirmed in our data (see Table A3).

IV.B. Data

We use data from Italian municipalities with less than 15,000 inhabitants for the period of 1993-2012. There are various reasons behind the selection of this sample. First, municipalities with less than 15,000 inhabitants use a single-ballot majoritarian electoral system, while those above this threshold employ a run-off system (Gamalerio, Morelli, and Negri, 2021). To keep electoral institutions constant, we thus exclude the latter group. Second, in 1993, following a corruption scandal called *Mani Pulite* (Clean Hands), new electoral municipal laws and a municipal property tax were introduced (Bordignon, Gamalerio, and Turati, 2020). Third, the different application of the Domestic Stability

Pact (DSP) across the 5000-inhabitant threshold remained in place until 2012. Finally, we exclude municipalities in the Special Statute Regions (i.e., Sardinia, Sicily, Valle d'Aosta, Trentino-Alto Adige, Friuli-Venezia Giulia) since they have distinct political and fiscal institutions, and a different set of fiscal rules in place since 2002.

The data set contains information on the characteristics of elected municipal politicians and mayoral candidates for the years 1993-2012.[28] The main observable characteristics, available from the Italian Home Office, are gender, age, years of past political experience at all levels of politics, political orientation (i.e., left, right or independent), past professional background, and education. Information on municipality characteristics comes from Istat and includes the share of the population with a university degree, the share of the active population (i.e., the population between 15 and 64 years old), the population of seniors (i.e., above 65 years old), income per capita, the number of firms and non-profit associations per capita, the area of the municipality in square kilometers, and population density. We use all of these variables for the balance tests. They were measured in 2001. We collected the data on municipal budget outcomes from the Aida PA database, an online archive managed by the Bureau Van Dijk. The collected data contains information on the fiscal items of the budgets of all Italian municipalities, covering the years 2000-2012. Finally, data on average income and income growth rate at the municipal level is provided by the Italian Ministry of Economics and Finance for the years 2000-2012.

The final sample consists of 26,005 electoral terms and 6,170 municipalities. Table A1 in the Appendix reports the summary statistics of this sample,

[28]The data set used in this paper was initially collected by Gamalerio (2020) and represents an extension of that initial data set.

distinguishing between municipalities below and above the cutoff. Table 3 reports the summary statistics of the municipalities included in the main CCT optimal bandwidth (i.e., the one used in the analysis for mayoral candidates) used in the *Diff-in-Disc* analysis below.

Table 3: Descriptive statistics:
Municipalities below 5000 vs. Municipalities above 5000
Sample within CCT optimal bandwidth

	(1) Below 5000	(2) obs	(3) Above 5000	(4) obs	(5) p-value
Politicians characteristics					
Female mayors	0.088	656	0.096	358	0.495
Age mayors	47.816	656	48.074	358	0.508
High skills job mayors	0.292	656	0.304	358	0.557
Graduate mayors	0.470	656	0.493	358	0.289
Political experience mayors	8.064	656	8.190	358	0.605
Female mayoral candidates	0.111	656	0.110	358	0.898
Age mayoral candidates	47.826	656	48.112	358	0.296
High skills job mayoral candidates	0.274	656	0.303	358	0.036
Graduate mayoral candidates	0.446	656	0.491	358	0.004
Municipal characteristics					
South	0.213	656	0.324	358	0.000
Centre	0.154	656	0.123	358	0.177
North-West	0.409	656	0.307	358	0.001
North-East	0.224	656	0.246	358	0.434
Population density	316.953	656	374.594	358	0.022
Area	31.756	656	39.330	358	0.002
No profit associations	0.004	656	0.004	358	0.278
Firms per capita	0.073	656	0.074	358	0.541
Income per capita	9800	656	9444	358	0.045
% elderly	0.183	656	0.183	358	0.879
% 15-64 years old	0.674	656	0.673	358	0.510
% graduate	0.046	656	0.048	358	0.107

Notes. Municipalities between 3886 and 6114 inhabitants (i.e., municipalities within the main CCT optimal bandwidth used in the analysis below). Electoral terms between 1993 and 2012. *Below 5000* = 1 for municipalities below 5,000 inhabitants. *Above 5000* = 1 for municipalities above 5,000 inhabitants. Columns (1) and (3) report the mean values for the two samples; *obs* is the number of observations; *p-value* is the p-value of the difference between the means of the two samples.

IV.C. Results

IV.C.I The effect of fiscal rules on the selection of politicians

The main results from the *Diff-in-Disc* analysis developed through model (11) appear in Table 4. This analysis utilizes data spanning from 1993 to 2012 and leverages the relaxation of fiscal rules in 2001. Column 1 presents estimates without controlling for covariates, whereas column 2 includes controls for election year fixed effects. In column 3, we further incorporate region fixed effects.

Two main findings emerge. First, the positive coefficients in front of the dummy variable ($> 5000_{it}$) indicate that in the years prior to 2001 (i.e., when fiscal rules applied equally across the cutoff), mayoral candidates in municipalities just above 5,000 inhabitants more frequently had a university education. This result is consistent with the fact that mayoral pay is higher in municipalities with more than 5,000 inhabitants, enabling them to attract more skilled candidates (Gagliarducci and Nannicini, 2013). Second, the negative coefficient in front of the interaction term between ($> 5000_{it}$) and ($Post_t$) suggests that the application of fiscal rules from 2001 onward in municipalities above 5,000 inhabitants offsets the positive selection effect induced by the higher wage paid. The results indicate that fiscal rules induced a reduction of around 9 percentage points in the share of mayoral candidates with a university degree. They also indicate that fiscal rules led to a similar reduction in the probability of electing a higher educated mayor.[29]

[29] As a robustness test, we check how the estimated coefficients change with the bandwidth used. As is typical in an RDD/*Diff-in-Disc* setup, moving toward smaller bandwidths produces larger (more negative) coefficients (i.e., lower bias) and larger confidence intervals (i.e., more inefficiency). For more detail see Figure A2 in the appendix. In addition, we demonstrate that the results are not due to random chances (DellaVigna and La Ferrara, 2010). Specifically, we run a series of *Diff-in-Disc* local linear regressions at 5,542 fictional thresholds for mayoral candidates and 4298 for mayors. We select the fictional thresholds

Table 4: The effect of fiscal rules on the education of politicians

	(1)	(2)	(3)
Control Function	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT
Election Year FE	No	Yes	Yes
Region FE	No	No	Yes
Panel A: mayoral candidates with university degree			
(> 5000)	0.087	0.086	0.060
	(0.033)	(0.033)	(0.031)
(Post)	0.069		
	(0.029)		
(Post)*(> 5000)	-0.091	-0.099	-0.101
	(0.043)	(0.043)	(0.041)
Observations	3,433	3,433	3,433
Bandwidth	1114	1114	1114
Mean outcome	0.469	0.469	0.469
Panel B: mayors with university degree			
(> 5000)	0.055	0.054	0.029
	(0.048)	(0.048)	(0.048)
(Post)	0.057		
	(0.047)		
(Post)*(> 5000)	-0.107	-0.114	-0.109
	(0.060)	(0.060)	(0.059)
Observations	4,383	4,383	4,383
Bandwidth	1425	1425	1425
Mean outcome	0.486	0.486	0.486

Notes. Diff-in-disc estimates of the impact of fiscal rules on the education level of politicians. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting after 2001. The outcome variable is the share of mayoral candidates with a university degree in Panel A, and a dummy variable equal to 1 for mayors with a university degree in Panel B. The bandwidth is calculated using the MSE-optimal bandwidth h selector following Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

After describing the baseline effect of fiscal rules on political selection, we test one of the three assumptions of the *Diff-in-Disc* methodology, which is that the municipalities just below and just above the 5,000-inhabitant thresh-

such that the bandwidths around the placebo cutoff do not contain the 5000 inhabitants threshold or any other population threshold at which other policies change (i.e., 1000, 3000, and 10,000 inhabitants thresholds). We run the regressions in these placebo tests using the same optimal bandwidths as in the main specification in Table 4 and controlling for election year and region fixed effects. We observe no distinct pattern in the estimated coefficients, with most confidence intervals including the zero line. Furthermore, the c.d.f. of the coefficients indicates that the majority are bigger compared to our coefficients in Table 4. Additionally, the c.d.f. of the t-statistics associated with these coefficients reveal that most values fall within the interval (-2,2), suggesting that it is rare to find statistically significant results at these fictional thresholds (see Figure A3 in the appendix).

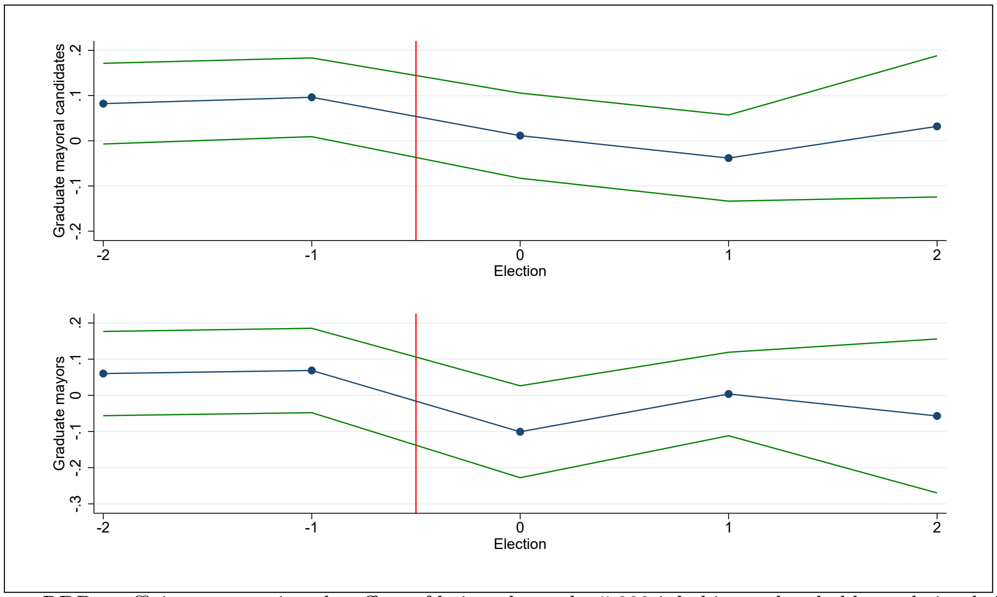
old must have been on parallel trends before the 2001 reform (see Section IV.A. for results on the other two assumptions). We provide evidence on the parallel trends assumption in Figure 1. We run a series of cross-sectional RDD regressions comparing the level of education of politicians across municipalities just below and above the 5,000-inhabitant threshold, where we group observations from different municipalities in a single regression depending on the distance of the electoral year from the 2001 relaxation of fiscal rules. For example, as reported on the x-axis of the two graphs in Figure 1, a value of 0 indicates the first elections ran immediately after the 2001 relaxation. The value of -1 refers to the elections run immediately before the 2001 reform, while -2 indicates elections run just before the elections at time -1. We run the cross-sectional RDD regressions using the same CCT optimal bandwidths used for the analysis in Table 4 (i.e., a bandwidth equal to 1114 for mayoral candidates and a bandwidth equal to 1425 for mayors).[30]

As we can observe from the top graph in Figure 1, in the two elections immediately before the 2001 removal of fiscal rules (i.e., -2 and -1, during which time fiscal rules applied in the same way across the threshold), municipalities above the threshold attracted more educated mayoral candidates. This evidence is consistent with the results in Panel A of Table 4 and the analysis of Gagliarducci and Nannicini (2013). Crucially for the parallel trends assumption, the RDD coefficients are stable between elections -2 and -1, signaling that municipalities below and above the threshold followed the same trends during the pre-2001 period. Consistent with the evidence of a negative effect of fiscal rules on political selection, Figure 1 clearly shows that the magnitude of the coefficients after 2001 (i.e., elections 0, 1, and 2, for which fiscal rules applied differently across the threshold) is reduced substantially compared to the pre-

[30]The estimates that define the points in Figure 1 are reported in Table A4.

2001 coefficients. The bottom graph in Figure 1 provides similar evidence for elected mayors.[31]

Figure 1: Cross-sectional RDD coefficients over time



Notes. RDD coefficients capturing the effect of being above the 5,000-inhabitant thresholds vs. being below it. On the x-axis, which goes from -2 to 2, we report the elections before and after the 2001 removal of fiscal rules, where 0 indicates the elections immediately after the relaxation of fiscal rules. We run the cross-section RDD regressions using the same CCT optimal bandwidths reported in Table 4 (i.e., a bandwidth equal to 1114 for mayoral candidates and a bandwdiths equal to 1425 for mayors). The blue lines connect the estimated coefficients, while the green lines represent the 95 percent confidence interval.

[31] We refrained from conducting the main analysis using a simple difference-in-difference model on the entire original sample, as the parallel trends assumption seems not to be met, particularly in the regressions for mayoral candidates (see Table A5 in the appendix). Furthermore, we employ model (11) to examine the potential interactive effects on politicians' education arising from the interaction between the 1999 introduction of fiscal rules and the different mayoral wages paid across the cutoff. This analysis helps demonstrate that municipalities on both sides of the threshold did not respond differently to the fiscal rules' introduction in 1999 (Grembi et al., 2016). The results of this analysis, detailed in Table A6 in the appendix, indicate that this appears to be the case.

IV.C.II Evidence on municipal budget rigidity

We conduct the analysis again, differentiating between municipalities with high pre-treatment budget rigidity in their spending and those with low levels of rigidity. We use municipalities' balance sheets, classifying as rigid the expenditures that leave little room for adjustments, such as personnel costs and debt repayment. These expenditures significantly constrain fiscal flexibility, as municipal governments face challenges in changing employee wages or firing staff, and must repay debts incurred by previous administrations (Persson and Svensson, 1989; Alesina and Tabellini, 1990; Pettersson Lidbom, 2001; Alt and Lassen, 2006; Eslava, 2010). Using data from municipalities' balance sheets in the year 2000 (i.e., the year before the 2001 fiscal rules relaxation), we estimate the rigidity of a municipality's balance sheet as the ratio between personnel plus debt repayment expenditures and total current revenues. In our data, this ratio takes an average value of 0.38 with a standard deviation of 0.14.

In columns 1 and 3 of Table 5, we run the *Diff-in-Disc* analysis using only the subsample of municipalities with a level of rigidity below the median. In columns 2 and 4, we keep only the municipalities with a level of rigidity above the median. We find that municipalities with a low level of rigidity drive the negative effect of fiscal rules on the education of politicians. Furthermore, there are no statistically significant differences between municipalities just above and below the threshold in the subsample of municipalities with a high level of rigidity.[32] This evidence further confirms that higher educated politicians are less likely to enter politics if they cannot enjoy a high level of policymaking

[32] In Table A7, we repeat the heterogeneity analysis by distinguishing between personnel expenditures as a share of current revenue and debt repayment expenditures as a share of current revenues. While the results go in the same direction, personnel expenditures have a more relevant role in guiding the heterogeneity analysis results.

discretion.[33]

Table 5: Heterogeneity analysis based on municipal budget rigidity

	(1)	(2)	(3)	(4)
Dependent variable	*Mayoral candidates with university degree*		*Mayors with university degree*	
Control Function	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT
Sample	*Rigidity<median*	*Rigidity>median*	*Rigidity<median*	*Rigidity>median*
Panel A: without controlling for election year and region FE				
(Post)*(> 5000)	-0.201	0.070	-0.169	0.007
	(0.047)	(0.071)	(0.075)	(0.092)
Observations	2,247	1,510	2,425	2,578
Bandwidth	1279	1140	1386	1862
Mean outcome	0.424	0.524	0.451	0.527
Panel B: controlling for election year and region FE				
(Post)*(> 5000)	-0.189	0.030	-0.163	-0.006
	(0.046)	(0.068)	(0.076)	(0.088)
Observations	2,247	1,510	2,425	2,578
Bandwidth	1279	1140	1386	1862
Mean outcome	0.424	0.524	0.451	0.527

Notes. Diff-in-disc estimates of the impact of fiscal rules on the education of politicians. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Rigidity is defined as the level of personnel and debt expenditures as a fraction of total current revenues. Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The outcome variable is the share of mayoral candidates with a university degree in columns 1-2, and a dummy variable equal to 1 for mayors with a university degree in columns 3-4. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

IV.C.III Graphical visualization of the *Diff-in-Disc* estimates

Figure 2 shows the discontinuity of politicians' education across the population threshold of 5,000 inhabitants, both before 2001, when the only difference at this threshold was given by politicians' wages, and from 2001 onward, when the application of fiscal rules at the threshold is also different. We present a scatter plot that shows the dependent variables, averaged over bins of 100 inhabitants. This plot includes data within a range equal to twice the optimal bandwidth and is accompanied by fitted regression lines.

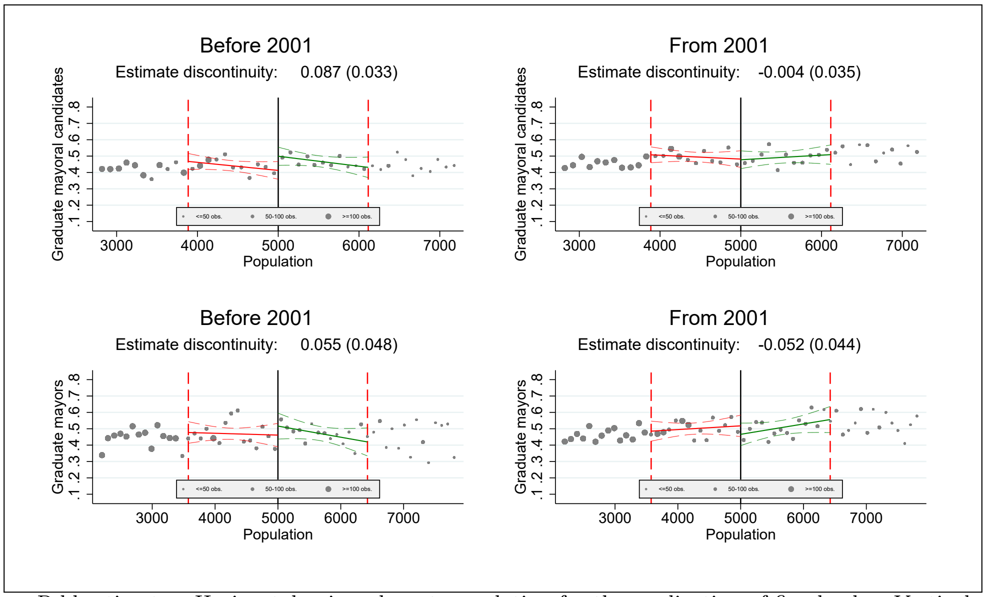
[33]We repeat the same graphical analysis of Figure 1 by separating municipalities depending on the level of budget rigidity. The evidence confirms that the sample of municipalities with low rigidity drive our results (see Figures A4-A5).

As shown in the graph for mayoral candidates running before 2001 (top left), there is a discontinuity across the threshold for the electoral years before 2001. This shows that higher wages paid above the threshold increased the share of more highly educated candidates running by about 9 percentage points. Conversely, as shown in the graph for candidates since 2001 (top right), the discontinuity around the cutoff disappears in the years 2001-2012, when fiscal rules no longer applied to municipalities below the cutoff. The bottom graphs provide similar evidence for elected mayors. Before 2001, municipalities above the threshold had a higher probability of electing graduate mayors, even though, the estimated discontinuity is not statistically different from zero. Since 2001, the estimated discontinuity becomes negative, even though not statistically significant.

To identify the *Diff-in-Disc* effect of fiscal rules on politicians' education (i.e., the main focus of this paper) in Figure 2, we need to calculate the difference between the estimated discontinuities in the graphs on the right of the Figure (i.e., -0.004 for mayoral candidates and -0.052 for elected mayors) and the estimated discontinuities in the graphs on the left (i.e., 0.087 for mayoral candidates and 0.055 for mayors). These two differences are equal to -0.091 for mayoral candidates and -0.107 for mayors and coincide with the *Diff-in-Disc* estimated coefficients of the interaction term between ($> 5000_{it}$) and ($Post_t$) reported in column 1 of Table 4.

Second, we provide direct graphical evidence of the discontinuities behind the *Diff-in-Disc* estimates in Figure 3. Specifically, we plot the change in politicians' education between the years after and before the 2001 fiscal rules relaxation, across the population threshold, through a scatter plot of the dependent variables averaged over bins of 100 inhabitants for all the observations in the data that lie in a bandwidth equal to double the optimal bandwidth. We

Figure 2: RDD graphical evidence



Notes. Rdd estimates. Horizontal axis: relevant population for the application of fiscal rules. Vertical axis: share of mayoral candidates (top graphs) and mayors (bottom graphs) with a university degree. Scatter points are averaged over bins of 100 inhabitants. The central line represents a linear regression of the outcome variable in the population, fitted separately on each side of the threshold. We run the linear RDD regressions using the same CCT optimal bandwidths used for the analysis in Table 4 (i.e., a bandwidth equal to 1114 for mayoral candidates and a bandwidth equal to 1425 for mayors). The other two dashed lines represent 95 percent confidence intervals. The vertical dashed lines indicates the limit of the optimal bandwidth used in the regressions. Number of observations in each graph: 1) top left graph: 3,454 observations in total, 1605 within the optimal bandwidth; 2) top right graph: 3886 observations in total, 1828 within the optimal bandwidth; 3) bottom left graph: 4803 observations in total, 2046 within the optimal bandwidth; 4) bottom right graph: 5331 observations in total, 2337 within the optimal bandwidth.

also add to the graphs the regression lines and 95 percent confidence intervals estimated by the *Diff-in-Disc* model (11).[34]

The top graphs of Figure 3 refer to the mayoral candidates, and the bottom graphs to the elected mayors. The graphs on the left refer to the effects estimated with the entire sample. The discontinuities are statistically significant and coincide in size with the estimated coefficients of the interaction term between ($> 5000_{it}$) and ($Post_t$) reported in column 1 of Table 2. The graphs on the right refer to the subsample of municipalities with low pre-treatment budget rigidities, which is the sample driving our results. In these graphs, the discontinuities are even larger and coincide with the estimated coefficients reported in columns 1 and 3 of Panel A of Table 5.[35] Overall, Figure 3 shows that following the different implementation of fiscal rules in 2001, the education level of mayoral candidates and mayors in municipalities just above the 5,000-inhabitant cutoff grew less compared to those in municipalities just below the threshold.

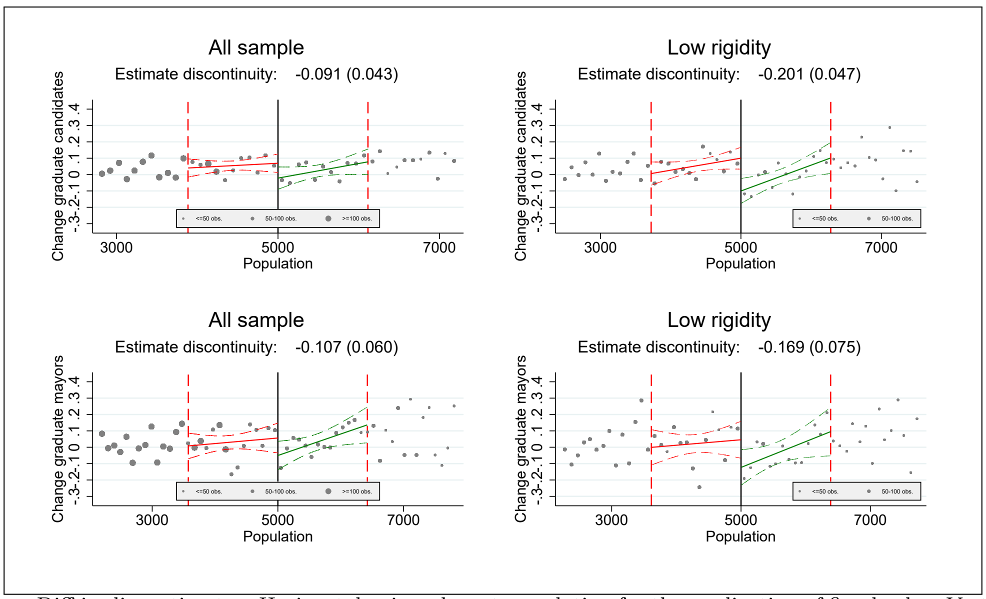
IV.C.IV Results on education level and policy choices

We explore whether highly educated politicians, as suggested by our theoretical model, are more likely to implement counter-cyclical fiscal policies than their less educated counterparts, indicating greater administrative competence. We use a Regression Discontinuity Design (RDD) based on close mixed electoral competitions between graduate and non-graduate mayors. Indeed, in mixed races decided by a narrow margin, election outcomes are likely determined

[34] We obtain the scatter points in Figure 3 by calculating the difference between the scatter points in the right and left graphs of Figure 2.

[35] We report in Figure A6 the graphical visualization of the *Diff-in-Disc* estimates obtained with the subsample of municipalities with high pre-treatment budget rigidities. Consistent with the estimated coefficients reported in columns 2 and 4 of Panel A of Table 5, we do not detect any discontinuity in Figure A6.

Figure 3: Diff-in-Disc graphical evidence (Entire and low-rigidity samples)



Notes. Diff-in-disc estimates. Horizontal axis: relevant population for the application of fiscal rules. Vertical axis: the change over time in the share of mayoral candidates (top graphs) and mayors (bottom graphs) with a university degree. Scatter points are averaged over bins of 100 inhabitants. The central line represents a linear regression of the outcome variable in the population, fitted separately on each side of the threshold. The other two dashed lines represent 95 percent confidence intervals. The vertical dashed lines indicates the limit of the optimal bandwidth used in the regressions. Number of observations in each graph: 1) top left graph: 7340 observations in total, 3433 within the optimal bandwidth; 2) top right graph: 4806 observations in total, 2247 within the optimal bandwidth; 3) bottom left graph: 10134 observations in total, 4383 within the optimal bandwidth; 4) bottom right graph: 5293 observations in total, 2425 within the optimal bandwidth.

by random factors and not by systematic municipal characteristics that could also affect policy outcomes. We estimate the following model using data at the municipality and electoral year level:

$$Y_{it} = \rho_0 + \rho_1 MV_{it} + \beta_0 Graduate_{it} + \beta_1 Graduate_{it} \cdot MV_{it} + \eta_{it} \tag{12}$$

where the dependent variable Y_{it} is the probability that a mayor chooses a counter-cyclical fiscal policy over the electoral mandate, seen as a proxy of her ability to match the state of the economy.[36] The treatment is the dummy variable $Graduate_{it}$, which is equal to 1 for mayors with a university degree and 0 otherwise. The assignment to treatment is uniquely determined by the margin of victory MV_{it}, which is the difference between the vote share of the graduate candidate minus the votes share of the non-graduate one.

We run model (12) on the sub-sample of municipalities in the interval $MV_{it} \in [-h, +h]$, where the optimal bandwidth h is calculated following the MSE-optimal bandwidth of Calonico, Cattaneo, and Titiunik (2014), and Calonico, Cattaneo, and Farrell (2018). We provide additional details on how in practice we estimate model (12) in the Appendix section D.2. In the Appendix B.1, we show the robustness of the two main assumptions required for

[36]To build this dependent variable, we create two dummy variables. The first is equal to 1 for mayors that, in a specific year, run a deficit above the median, calculated across the entire sample of mixed electoral competitions, and 0 otherwise. The second dummy variable is equal to 1 for years characterized by a growth in municipal income below the median of the entire sample of mixed electoral competitions and 0 otherwise. Hence, the dummy variable for a mayor implementing a counter-cyclical fiscal policy in a specific year is equal to 1 when these two dummy variables are equal and 0 otherwise. We then collapse the data at the municipality and electoral term level, resulting in a dependent variable capturing the probability of a mayor's fiscal policy matching the state of the economy throughout the term. Furthermore, to deal with the potential endogeneity of municipal income growth during a mayor's term, Table B3 shows that we get similar results if we use an alternative version of this dependent variable built using the prediction of income growth based on pre-determined municipal characteristics.

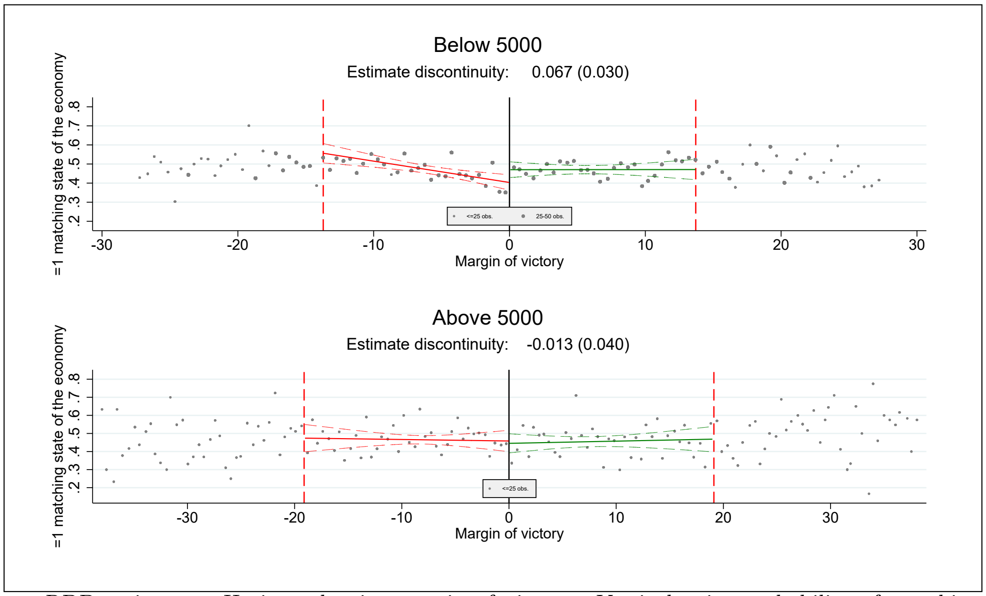
this identification to work correctly, which are that there must be no sorting around the threshold $MV_{it} = 0$ and that observable municipal characteristics should vary smoothly at the threshold $MV_{it} = 0$.

We implement this analysis for the electoral terms between 2001-2012. We report RDD results in Figure 4 and Table 6. Panel A of Table 6 reports the results for municipalities not constrained by fiscal rules in the period 2001-2012, and Panel B the results for municipalities affected by fiscal rules. In column 1, we run model (12) without control variables. In column 2, we control for other personal characteristics of the mayors such as age, gender, political experience, professional background, and political orientation. In column 3, we add regional and election year fixed effects. The results show that, where fiscal rules do not apply, higher educated mayors have a higher probability of matching the state of the economy compared to less educated ones. Conversely, we do not find statistically significant differences between graduate and non-graduate mayors in municipalities affected by fiscal rules.[37]

Finally, we observe qualitatively similar results when using alternative policy outcomes, such as investment levels, public services provision, and fiscal sustainability measures, to capture the competence of mayors. For more detail on this analysis see Appendix B.2 and Table B5.

[37]The p-values presented in the last column of Table 6 indicate that the coefficients in Panels A and B, in the specifications incorporating region and electoral year fixed effects, are statistically different at conventional levels of significance. The different results for municipalities below and above the 5,000-inhabitant threshold may be due to the different wages paid to the mayors. To rule out this possibility, in Table B4, we extend our dataset to the period 2013-2015 to show that differences in matching the state of the economy disappear when fiscal rules applied equally across the threshold.

Figure 4: Graduate mayors and matching the state of the economy



Notes. RDD estimates. Horizontal axis: margin of victory. Vertical axis: probability of matching the state of the economy over the electoral mandate. Scatter points are averaged over bins of 0.5 percentage points. The central line represents a linear regression of the outcome variable in the margin of victory, fitted separately on each side of the threshold. The other two dashed lines represent 95 percent confidence intervals. The vertical dashed lines indicates the limit of the optimal bandwidth used in the regressions. Number of observations in each graph: 1) top graph: 3459 observations in total, 2098 within the optimal bandwidth; 2) bottom graph: 1,587 observations in total, 1085 within the optimal bandwidth.

Table 6: Graduate mayors and matching the state of the economy

	(1)	(2)	(3)
Control Function	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT
Year of election FE	No	No	Yes
Region FE	No	No	Yes
Mayoral covariates	No	Yes	Yes
Panel A: municipalities below 5000			
Graduate Mayor	0.067	0.079	0.070
	(0.030)	(0.030)	(0.028)
Observations	2098	2032	2038
Bandwidth	13.71	13.19	13.21
Panel B: municipalities above 5000			
Graduate Mayor	-0.013	-0.007	-0.014
	(0.040)	(0.041)	(0.039)
Observations	1085	1077	1010
Bandwidth	19.11	18.83	17.28
P-Value difference Panel A vs. B	0.110	0.091	0.082

Notes. Municipalities below 15,000 inhabitants. Electoral terms between 2001 and 2012. Dependent variable: probability of matching the state of the economy over the electoral mandate. Treatment variable equal to 1 when mayor has a university degree, 0 otherwise. Estimation by RDD using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Mayoral covariates in columns 2-3: 1) age of the mayor; 2) political experience: years of past political experience of the mayor at any level of politics; 3) high skills job = 1 if mayor worked in a high skills occupation in the past; 4) female = 1 if mayor is a woman; 5) left = 1 for a center-left mayor. Robust standard errors clustered at the local labor area level are in parentheses. The last row presents the p-value for the test comparing whether the coefficients in panel A are the same to that in panel B.

IV.C.V Alternative stories

Different alternative narratives could account for this paper's results. The first alternative explanation is that our findings are attributable to a real-term reduction in the wage differences for mayors across the threshold. Although we cannot directly discount this hypothesis, the evidence in Figure 1 does not seem to align with it. Specifically, if wage differences were the driving factor, a more gradual change in the education level of mayoral candidates and mayors over time would be expected, rather than the sharp decline observed just after 2001.

Second, the application of fiscal rules may require the selection of more politically experienced politicians, who may be less educated. To rule out this explanation, we test the effect of fiscal rules on other personal characteristics of local politicians, such as past professional background, age, gender, and past political experience. While for characteristics potentially correlated with education, the estimated coefficient goes in the expected direction (i.e., a decline in the share of politicians from skilled occupations), we do not find any effect of fiscal rules on years of past political experience (see Table C1). In addition, we check whether fiscal rules negatively affected municipal councilors' education level. As described in section I., our expectation about the effect of fiscal rules was that these should affect politicians in powerful positions, like mayors, rather than politicians in less prominent positions, like municipal councilors. In line with this expectation, we do not find that fiscal rules affect the characteristics of municipal councilors (see Table C2).

Third, we demonstrate that the labor market's different career options outside of politics for individuals with varying education levels do not explain the results of this paper (see Table C3). Fourth, we find that fiscal rules did not impact the ideological orientation of mayoral candidates and mayors, another dimension potentially correlated with education (see Table C4). Fifth, we provide evidence suggesting that the effect we observe may not be solely attributed to changes in voter demand for highly educated politicians induced by fiscal rules (see Tables C5-C6). Finally, in Tables C7-C8, we rule out the possibility that our results are driven by the effect of fiscal rules on corruption, as estimated in the literature (Daniele and Giommoni, 2020). For further details on how we rule out these alternative explanations, see Appendix C.

V. Concluding remarks

This paper investigates the effect of a reduction in policymaking discretion on the selection of politicians. We theoretically and empirically document that such a reduction, induced by the application of fiscal rules, negatively affects politicians' level of education. Beyond the political economy literature, this finding aligns with recent studies on discretion and public procurement, showing that, while political discretion can create opportunities for misuse of public money, it can also enable greater efficiency (Decarolis et al., 2021) and better procurement outcomes (Coviello et al., 2022). Obviously, our results do not imply that fiscal rules are generally inefficient or welfare-decreasing. This paper only documents an additional and previously unseen effect on political selection.

Our results highlight three considerations for future research. First, they underline the possibility that electing fewer educated politicians may lead to worse policies. The evidence we provide is not conclusive, and it calls for further research on the policy implications of reducing the share of well-educated elected politicians. Second, though we analyze the ex-ante politicians' education, we do not address the latter's representativeness relative to the electorate. Indeed, the fact that the election of fewer educated and skilled politicians may lead to worse policies is but one facet of the story (Dal Bo et al., 2020; Carreri, 2020). The election of fewer educated individuals may also translate into a better representation of marginalized groups in terms of labor market performance and socioeconomic background. The existing literature says little about whether a better political representation of marginalized groups may lead to more targeted policies toward them, an important issue that might be further explored. Third, this paper provides evidence using data from one country and strategies such as *Diff-in-Disc* and RDD. Such an approach enables avoiding the limitations of cross-country analyses. Moreover, the methodologies employed have strong internal validity and provide casual estimates. However, there is a potential cost in terms of external validity, calling for further research using data from other countries to understand whether our findings also apply in other contexts.

Finally, there are two policy implications related to our results. First, our findings suggest that the negative selection effect of fiscal rules should be taken into account in the design of such rules. Although local, the effect is sizeable and significant, indicating that the composition of the candidate pool is endogenous to the rules themselves. Second, we show that reduced discretion brought about by fiscal rules compensates for the positive selection implied by higher wages paid just above the 5,000-inhabitant threshold. This suggests that one possible way to compensate for the negative selection effect of fiscal rules would be to combine them with higher pay for politicians.

References

Alesina, A., and R. Perotti (1996). "Fiscal Discipline and the Budget Process." *American Economic Review Papers and Proceedings*, 86: 401–407.

Alesina, A. and Tabellini, G. (1990). "Voting on the budget deficit." *American Economic Review*, 80(1): 37-49.

Alpino, M., Asatryan, Z., Blesse, S., and N. Wehrhöfe (2020). "Austerity and Distributional Policy." CESifo Working Paper No. 8644.

Alt, J. and Lassen, D. (2006). "Fiscal transparency, political parties, and debt in OECD countries." ıEuropean Economic Review, 50: 1430-1439.

Anci (2013a). *"Patto stabilità – Sindaco Ravenna si ribella: Abbiamo 14 milioni da spendere, venga rivisto subito."*. Available at: `https://www.anci.it/patto-stabilita-sindaco-ravenna-si-ribella-abbiamo-14-milioni-da-spendere-venga-rivisto-subito/` [Accessed 09/02/2023].

Anci (2013b). *"Patto stabilità – Avanzo di 500 mila euro bloccato, sindaco del padovano non ci sta: Non alzo le tasse, mi dimetto."*. Available at: `https://www.anci.it/patto-stabilita-avanzo-di-500-mila-euro-bloccato-sindaco-del-padovano-non-ci-sta-non-alzo-le-tasse-mi-dimetto/` [Accessed 09/02/2023].

Artiles, M., Kleine-Rueschkamp, L., and G. León-Ciliotta (2020). "Accountability, Political Capture and Selection into Politics: Evidence from Peruvian Municipalities." *Review of Economics and Statistics*, forthcoming.

Ashraf, N., Bandiera, O., Davenport, E., and S. S. Lee (2019). "Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services." *American Economic Review*,110(5): 1355-1394.

Azzimonti, M., Battaglini, M., and S. Coate (2016). "The costs and benefits of balanced budget rules: Lessons from a political economy model of fiscal policy." *Journal of Public Economics*, 136: 45-61.

Bandiera, O., Prat, A., and T. Valletti (2009). "Active and Passive Waste in Government Spending: Evidence from a Policy Experiment." *American Economic Review*, 99 (4): 1278-1308.

Baltrunaite, A., Bello, P., Casarico, A. and P. Profeta (2014) "Gender quotas and the quality of politicians." *Journal of Public Economics*, 118(9–10): 62–74.

Bartling, B., Fehr, E. and K. M. Schmidt (2012). "Screening, Competition, and Job Design: Economic Origins of Good Jobs." *American Economic Review*, 102(2): 834-864.

Battaglini, M., and S. Coate (2008). "A dynamic theory of public spending, taxation and debt." *American Economic Review*, 98(1): 201–236.

Beath, A., Christia, F., Egorov, G. and R. Enikolopov (2014). "Electoral rules and the quality of politicians: theory and evidence from a field experiment in Afghanistan." NBER Working Papers No. 20082. National Bureau of Economic Research.

Besley, T. (2004). "Paying politicians: theory and evidence." *Journal of the European Economic Association*, 2: 193-215.

Besley, T. (2005). "Political Selection." *Journal of Economic Perspectives*, 19(3): 43-60.

Besley, T., Montalvo, J. G., and M. Reynal-Querol (2011). "Do Educated Leaders Matter?" *The Economic Journal*, 121(554): 205-227.

Besley, T., and M. Reynal-Querol (2011). "Do Democracies Select More Educated Leaders?" *American Political Science Review*, 105(3): 552-566.

Boffa, F., Mollisi, V. and G. Ponzetto (2022). "Do Incompetent Politicians Breed Populist Votes? Evidence from Italian Municipalities." *working paper*.

Bonfatti, A., and Forni, L. (2019). "Fiscal rules to tame the political budget cycle: Evidence from Italian municipalities," *European Journal of Political Economy*, 2019, 60, 101800.

Bordignon, M., Gamalerio, M., and G. Turati (2020). "Manager or professional politician? Local fiscal autonomy and the skills of elected officials." *Regional Science and Urban Economics*, 83.

Braendle, T. (2015) "Does remuneration affect the discipline and the selection of politicians? Evidence from pay harmonization in the European Parliament." *Public Choice* 162(1): 1–24.

Braendle, T. (2016) "Do institutions affect citizens' selection into politics?" *Journal of Economic Surveys*, 30(2): 205-227.

Brollo, F., Nannicini, T., Perotti, R., and G. Tabellini (2013). "The Political Resource Curse". *American Economic Review*, 103: 1759-1796.

Calonico, S., Cattaneo, M., and R. Titiunik (2014). "Robust Nonparametric Bias Corrected Inference in the Regression Discontinuity Design." *Econometrica*, 82(6): 2295-2326.

Calonico, S., Cattaneo, M., and M. H. Farrell (2018). "On the effect of bias estimaton on coverage accuracy in non parametric inference" *Journal of the American Statistical Association*, 113(522): 767-779.

Campa, P. (2011): "Gender Quotas, Female Politicians and Public Expenditures: Quasi Experimental Evidence", mimeo, Stockholm University.

Card, D., (1999). “The causal effect of education on earnings.” In Orley Ashenfelter and David Card, editors, Handbook of Labor Economics Volume 3A. Amsterdam: Elsevier.

Carnes, N., and N. Lupu (2016). “What Good Is a College Degree? Education and Leader Quality Reconsidered.” *Journal of Politics*, 78(1): 35-49.

Carnes, N., and N. Lupu (2023). “The Economic Backgrounds of Politicians.” *Annual Review of Political Science*, 26, 11.1-11.18.

Carreri, M. (2020). “Can Good Politicians Compensate for Bad Institutions? Evidence from an original survey of Italian mayors.” *Journal of Politics*, forthcoming.

Carreri, M. and L. R. Martinez (2021). “Economic and Political Effects of Fiscal Rules: Evidence from a Natural Experiment in Colombia”. Working paper.

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

Caselli, F., and M. Morelli (2004) “Bad politicians.” *Journal of Public Economics* 88(3-4): 759–782.

Cervellati, M., Gulino, G., and Roberti, P. (2021). “Misvoting in the Ballot: Random Power to Parties and its Effect on Politicians and Policies.” Cepr DP14906 (v. 2).

Chiades, P., and V. Mengotto (2013). “Il calo degli investimenti nei Comuni tra Patto di stabilità interno e carenza di risorse,” Bank of Italy working paper.

Coviello, D., Guglielmo, A., Lotti, C., and G. Spagnolo (2022). “Procurement with Manipulation”. *CEPR working paper DP17063.*

Coviello, D., Marino, I., Nannicini, T., and N. Persico (2021). “Demand Shocks and Firm Investment: Micro-evidence from fiscal retrenchment in Italy.” *The Economic Journal*, accepted.

Dal Bo, E. and F. Finan (2018). “Progress and Perspectives in the Study of Political Selection.” *Annual Review of Economics*, 10, 541–575.

Dal Bo, E., Finan, F. and M. A. Rossi (2013). “Strengthening state capabilities: the role of financial incentives in the call to public service.” *Quarterly Journal of Economics* 128(3): 1169–1218.

Dal Bo, E., Finan, F., Folke, O., Persson, T., and J. Rickne (2017). “Who Becomes A Politician?” *Quarterly Journal of Economics*, 132(4): 1877–1914.

Dal Bo, E., Finan, F., Folke, O., Rickne, J., and T. Persson (2020) “Economic and Social Outsiders but Political Insiders: Sweden’s Radical Right.” Working paper.

Daniele, G., and B. Geys (2015). “Organised crime, institutions and political quality: evidence from Italian municipalities”, *The Economic Journal*, 125(586): F233-F255.

Daniele G., and T. Giommoni (2020). “Corruption under austerity.” Working paper.

De Benedetto, M. A. (2018). “Politicians Quality and Electoral System. Evidence from a Quasi-experimental Design for Italian Cities.” Working paper.

De Biase, P., and S. Dougherty (2022). “The past and future of subnational fiscal rules: analysis of fiscal rules over time”, *OECD Working Papers on Fiscal Federalism*, n. 41.

Debrun, X., Moulin, L., Turrini, A., Ayuso-i-Casals, J., and M. S. Kumar (2008). “Tied to the Mast? National Fiscal Rules in the European Union.” *Economic Policy*, 54: 297–362.

Decarolis, F., Fisman, R., Pinotti, P., and S. Vannutelli (2021). “Rules, Discretion, nad Corruption in Procurement: Evidence from Italian Government Contracting”. *NBER working paper 28209.*

Dee, T. S. (2004). “Are there civic returns to education?” *Journal of Public Economics*, 88(9–10): 1697-1720.

DellaVigna, St., and La Ferrara, E. (2010). “Detecting Illegal Arms Trade.” *American Economic Journal: Economic Policy*, 2 (November 2010): 26–57.

Deserrano, E. (2018). “Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda.” *American Economic Journal: Applied Economics*, 11(1): 277-317.

Eslava, M. (2010). “The political economy of fiscal deficits: a survey.” *Journal of Economic Surveys*, 25(4): 645-673.

Fedele, A., and P. Naticchioni (2013). “Moonlighting politicians: motivation matters!” IZA Discussion Paper, No. 7500: Bonn.

Ferraz, F. and C. Finan (2011). “Motivating Politicians: The Impacts of Monetary Incentives on Quality and Performance.” Working paper.

Finan, F., Olken, B. A., and R. Pande (2015). “The Personnel Economics of the State.”. *NBER working paper 21825.*

Fisman, R., Harmon, N. A., Kamenica, E. and I. Munk (2015). “Labor Supply of Politicians”. *Journal of the European Economic Association*, 13(5): 871-905.

Fisman, R., Schulz, F., and V. Vig (2019). “Financial disclosure and political selection: Evidence from India.” Working paper.

Gagliarducci, S., and T. Nannicini (2013). “Do Better Paid Politicians Perform Better? Disentangling Incentives from Selection.” *Journal of the European Economic Association*, 11: 369–398.

Gagliarducci, S., Nannicini, T., and P. Naticchioni (2010). “Moonlighting politicians.” *Journal of Public Economics*, 94(9–10): 688–699.

Galasso, V., and T. Nannicini (2011). “Competing on good politicians.” *American Political Science Review*, 105(1): 79–99.

Galasso, V., and T. Nannicini (2017). “Political Selection under Alternative Electoral Rules.” *Public Choice*, 171: 257–281.

Gamalerio, M., (2020). “Do national political parties matter? Evidence from Italian municipalities.” *European Journal of Political Economy*, Volume 63, June 2020.

Gamalerio, M., Morelli, M., and M. Negri, (2021). “The Political Economy of Open Borders. Theory and Evidence on the role of Electoral Rules.” Working paper.

Gelman, A., and G. Imbens (2018). “Why high-order polynomials should not be used in regression discontinuity designs.” *Journal of Business and Economic Statistics*, 37(3): 447-456.

Gethin, A., Martinez-Toledano, C., and T. Piketty (2022). “Brahim left versus merchant right: changing political cleavages in 21 western democracies, 1948-2020.” *Quarterly Journal of Economics*, 137(1): 1-48.

Grembi, V., Nannicini, T., and U. Troiano (2016). “Do fiscal rules matter?” *American Economic Journal: Applied Economics*, 8: 1-30.

Grossman, G., and W. W. Hanlon (2013). “Do better monitoring institutions increase leadership quality in community organizations? Evidence from Uganda.” *American Journal of Political Science*, 58(3): 669–686.

Gulino, G. (2020). “Electoral Systems, Selection, and Re-election: Evidence from Italian Municipalities.” *The Journal of Law, Economics, and Organization*, Volume 37, Issue 3, November 2021.

Gulzar, S., (2021). “Who Enters Politics and Why?” *Annual Review of Political Science*, 24(1), 253–75.

Gulzar, S., and M. Y. Khan (2023). “Good Politicians”: Experimental Evidence on Motivations for Political Candidacy and Government Performance.” *Review of Economic Studies*, accepted (pending data review).

Halac, M., and P. Yared (2014). "Fiscal rules and discretion under persistent shocks." *Econometrica*, 82(5): 1557-1614.

Halac, M., and P. Yared (2018). "Fiscal Rules and Discretion in a World Economy." *American Economic Review*, 108(8): 2305-2334.

Halac, M., and P. Yared (2019). "Fiscal Rules and Discretion under Limited Enforcement." Working paper.

Hallerberg, M., and J. Von Hagen (1999). "Electoral Institutions, Cabinet Negotiations, and Budget Deficits in the EU." in J. Poterba and J. Von Hagen (eds.), Fiscal Institutions and Fiscal Performance, Chicago, IL: University of Chicago Press.

Izzo, F. (2020). "Who runs when?." Working paper.

Kotakorpi, K., and P. Poutvaara (2011). "Pay for Politicians and Candidate Selection: An Empirical Analysis." *Journal of Public Economics*, 95(7-8): 877-885.

Lo Prete, A., and F. Revelli (2021). "Voter Turnout and City Performance: Evidence from Italian Municipalities." *Journal of Law, Economics and Organization*, 37(1): 168-197.

Lalive, R. (2008). "How Do Extended Benefits Affect Unemployment Duration: A Regression Discontinuity Approach," *Journal of Econometrics*, 142(2): 785–806.

Le Borgne, E. and B. Lockwood (2002). "Candidate Entry, Screening, and the Political Budget Cycle." IMF working paper WP/02/48.

Leonardi, M., and G. Pica (2013). "Who Pays for It? The Heterogeneous Wage Effects of Employment Protection Legislation." *Economic Journal*, 123(573): 1236-1278.

Krcmaric, D., Nelson, S. C., Roberts, A. (2020). "Studying Leaders and Elites: The Personal Biography Approach." *Annual Review of Political Science*, 23, 133-151.

Martinez-Bravo, M. (2017). "The Local Political Economy Effects of School Construction in Indonesia" *American Economic Journal: Applied Economics*, 9(2): 256–289.

Mattozzi, A., and A. Merlo (2008). "Political careers or career politicians?" *Journal of Public Economics*, 92: 597–608.

Mattozzi, A., and A. Merlo (2015). "Mediocracy" *Journal of Public Economics*, 130: 32–44.

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

Meriläinen, J. (2021). "Political Selection and Economic Policy." *Economic Journal*, conditionally accepted.

Messner, M., and M. K. Polborn (2004). "Paying Politicians." *Journal of Public Economics*, 88: 2423–2445.

Milligan, K., Moretti, E., and P. Oreopoulos (2004). "Does education improve citizenship? Evidence from the United States and the United Kingdom." *Journal of Public Economics*, 88: 1667–1695.

Mitra, A. (2020). "Should you want an educated mayor? Evidence from close elections in Italy." Working paper.

Murtinu, S., Piccirilli, G., and A. Sacchi (2021). "Rational inattention and politics: how parties use fiscal policies to manipulate voters." *Public Choice*, forthcoming.

Openpolis (2016). *"Rigidità della spesa, quando il sindaco ha le "mani legate""*. Available at: `https://blog.openpolis.it/2016/10/12/rigidita-della-spesa-quando-il-sindaco-ha-le-mani-legate` [Accessed 05/01/2023].

Oyer, P., and J. Schaefer (2010). "Personnel Economics: Hiring and Incentives." NBER working paper 15977.

Pavese, C., and E. Rubolino (2022). "Austerity Harmed Student Achievement." Working paper.

Peralta, S., and J. Pereira dos Santos (2018). "Who seeks reelection: local fscal restraints and political selection." *Public Choice*, 184: 105–134.

Persson, T., and L. Svensson (1989). "Why a Stubborn conservative would run a deficit." *Quarterly Journal of Economics*, 104(2): 325-345.

Pettersson-Lidbom, P. (2001). "An Empirical Investigation of the Strategic Use of Debt." *Journal of Political Economy*, 109(3): 570-583.

Piguillem, F. and A. Riboni (2015). "Spending biased legislators: discipline through disagreement." *Quarterly Journal of Economics*, 130(2): 901–949.

Poterba, J. (1994). "State Responses to Fiscal Crises: The Effects of Budgetary Institutions and Politics." *Journal of Political Economy*, 102: 799–821.

Poterba, J. (1996). "Budget Institutions and Fiscal Policy in the U.S. States." *American Economic Review Papers and Proceedings*, 86: 395–400.

Repetto, L. (2016). "Balance sheet disclosure and the budget cycle of Italian municipalities." Working paper.

Revelli, F. (2016). "Tax limits and local democracy." *Public Choice*, 166: 53–68.

Vannutelli, S. (2022). "From Lapdogs to Watchdogs: Random Auditor Assignment and Municipal Fiscal Performance in Italy." Working paper.

Wyplosz, C. (2012). "Fiscal Rules: Theoretical Issues and Historical Experiences." NBER Working Paper 17884.

Yared, P. (2019). "Rising Government Debt: Causes and Solutions for a Decades-Old trend." *Journal of Economic Perspectives*, 33(2): 115-140.

ONLINE APPENDIX

Fiscal Rules and the selection of politicians: theory and evidence from Italy

Matteo Gamalerio

Institut d'Economia de Barcelona (IEB), University of Barcelona

&

Federico Trombetta

Department of International Economics, Institutions and Development (DISEIS), Faculty of Political and Social Sciences, Università Cattolica del Sacro Cuore, Milano

A Appendix: Additional tables and figures

Table A1: Descriptive statistics:
Municipalities below 5000 vs. Municipalities above 5000

	(1) Below 5000	(2) obs	(3) Above 5000	(4) obs	(5) p-value
Politicians characteristics					
Female mayors	0.088	4836	0.094	1334	0.230
Age mayors	48.236	4836	47.786	1334	0.023
High skills job mayors	0.228	4836	0.310	1334	0.000
Graduate mayors	0.374	4836	0.516	1334	0.000
Political experience mayors	8.329	4836	8.226	1334	0.381
Female mayoral candidates	0.105	4836	0.110	1334	0.310
Age mayoral candidates	48.110	4836	48.076	1334	0.814
High skills job mayoral candidates	0.213	4836	0.307	1334	0.000
Graduate mayoral candidates	0.356	4836	0.500	1334	0.000
Municipal characteristics					
South	0.253	4836	0.289	1334	0.008
Centre	0.136	4836	0.166	1334	0.006
North-West	0.504	4836	0.307	1334	0.000
North-East	0.107	4836	0.239	1334	0.000
Population density	146.931	4836	496.301	1334	0.000
Area	25.328	4836	43.145	1334	0.000
No profit associations	0.005	4836	0.004	1334	0.000
Firms per capita	0.067	4836	0.076	1334	0.000
Income per capita	8907	4836	9795	1334	0.000
% elderly	0.229	4836	0.177	1334	0.000
% 15-64 years old	0.643	4836	0.677	1334	0.000
% graduate	0.043	4836	0.052	1334	0.000

Notes. Municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. *Below 5000* = 1 for municipalities below 5,000 inhabitants. *Above 5000* = 1 for municipalities above 5,000 inhabitants. Columns (1) and (3) report the mean values for the two samples; *obs* is the number of observations; *p-value* is the p-value of the difference between the means of the two samples.

Table A2: Balance test on municipal covariates

Diff-in-Disc

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Characteristics municipal population							
Dependent variables	% university degree	% 15-64	% 65+	(log) income per capita	# firms	no-profit ass	area	population density
(Post)*(> 5000)	-0.000	-0.007	-0.002	-0.022	0.000	-0.000	0.159	-7.468
	(0.002)	(0.004)	(0.006)	(0.043)	(0.003)	(0.000)	(5.195)	(51.177)
Observations	2,950	2,644	2,509	4,172	2,810	1,716	3,536	3,465
Bandwidth	954.3	859.8	818.2	1346	909.8	568.3	1151	1131
	Panel B: Geographical characteristics municipalities, deficit and re-election/re-run status							
Dependent variables	NE	NW	CEN	SOU	Second term mayor	Re-elected		
(Post)*(> 5000)	-0.065	0.014	0.057	0.032	0.089	-0.031		
	(0.050)	(0.061)	(0.040)	(0.064)	(0.068)	(0.070)		
Observations	3,865	3,772	5,105	3,182	2,706	2,819		
Bandwidth	1246	1220	1644	1034	879.8	1418		

Notes. Diff-in-disc estimates of the impact of fiscal rules on municipal covariates. Municipalities between 0-15,000. Electoral years between 1993 and 2012. Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5,000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the municipality level are in parentheses.

Table A3: Jump of higher wage indicator at the threshold

	(1)	(2)	(3)
Control Function	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT
Election Year FE	No	Yes	Yes
Region FE	No	No	Yes
Dependent variable = 1 if mayor paid higher wage			
(> 5000)	0.335	0.329	0.316
	(0.071)	(0.070)	(0.071)
(Post)*(> 5000)	0.041	0.035	0.027
	(0.099)	(0.093)	(0.092)
Observations	1,418	1,418	1,418
Bandwidth	466.3	466.3	466.3

Notes. Diff-in-disc estimates of the impact of fiscal rules on higher wage indicator. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting after 2001. The outcome variable is a dummy variable equal to 1 if mayor receive higher wage, according to the Census population. The bandwidth is calculated using the MSE-optimal bandwidth h selector following Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

Table A4: Cross-sectional RDD coefficients over time

	(1)	(2)	(3)	(4)	(5)
Control Function	Linear	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT	CCT
Election Year FE	No	No	No	No	No
Region FE	No	No	No	No	No
Election	-2	-1	0	1	2
Panel A: Mayoral candidates with university degree					
($>$ 5000)	0.082	0.096	0.011	-0.038	0.032
	(0.046)	(0.044)	(0.048)	(0.049)	(0.080)
Observations	775	802	783	774	267
Bandwidth	1114	1114	1114	1114	1114
Panel B: Mayors with university degree					
($>$ 5000)	0.060	0.069	-0.101	0.004	-0.057
	(0.059)	(0.060)	(0.065)	(0.059)	(0.109)
Observations	984	1,022	1,001	1,005	327
Bandwidth	1425	1425	1425	1425	1425

Notes. RDD coefficients capturing the effect of being above the 5,000-inhabitant thresholds vs. being below it. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the Table: 1) ($>$ 5000) = 1 for municipalities with more than 5000 inhabitants. The outcome variable is the share of mayoral candidates with a university degree in Panel A, and a dummy variable equal to 1 for mayors with a university degree in Panel B. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Specifically, we run the cross-section RDD regressions using the optimal CCT bandwidths reported in Table 4. Robust standard errors clustered at the local labor area level are in parentheses.

Table A5: The effect of fiscal rules on the education of politicians
Difference-in-differences estimates

	(1)	(2)	(3)	(4)	(5)	(6)
Election Year FE	No	No	Yes	No	No	Yes
Municipal FE	No	No	Yes	No	No	Yes
Dependent	Share mayoral candidates with university degree			= 1 for Mayors with university degree		
(> 5000)	0.145	0.154	0.028	0.135	0.140	-0.017
	(0.012)	(0.013)	(0.021)	(0.016)	(0.017)	(0.038)
(Post)	0.037	0.035		0.023	0.020	
	(0.005)	(0.006)		(0.007)	(0.008)	
(Post)*(> 5000)	-0.010	-0.019	-0.004	0.007	0.001	0.023
	(0.010)	(0.011)	(0.011)	(0.016)	(0.019)	(0.020)
Pre		-0.004			-0.006	
		(0.005)			(0.006)	
(Pre)*(> 5000)		-0.018	-0.011		-0.011	0.003
		(0.011)	(0.011)		(0.015)	(0.015)
Observations	26,005	26,005	26,005	26,005	26,005	26,005

Notes. Difference-in-differences estimates of the impact of fiscal rules on the education of politicians. Municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5,000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001; 3) (Pre) = 1 for election immediately before 2001 fiscal rules removal. The outcome variable is the share of mayoral candidates with a university degree in columns 1-3, and a dummy variable equal to 1 for mayors with a university degree in columns 4-6. Robust standard errors clustered at the local labor area level are in parentheses.

Table A6: Introduction of fiscal rules

	(1)	(2)
Dependent Variables	Share mayoral candidates with university degree	= 1 for Mayors with university degree
Control Function	Linear	Linear
Bandwidth	CCT	CCT
Election Year FE	No	No
Region FE	No	No
(>= 1999)*(> 5000)	-0.042	-0.043
	(0.061)	(0.077)
Observations	1,926	2,210
Bandwidth	1335	1534
Mean outcome	0.441	0.464

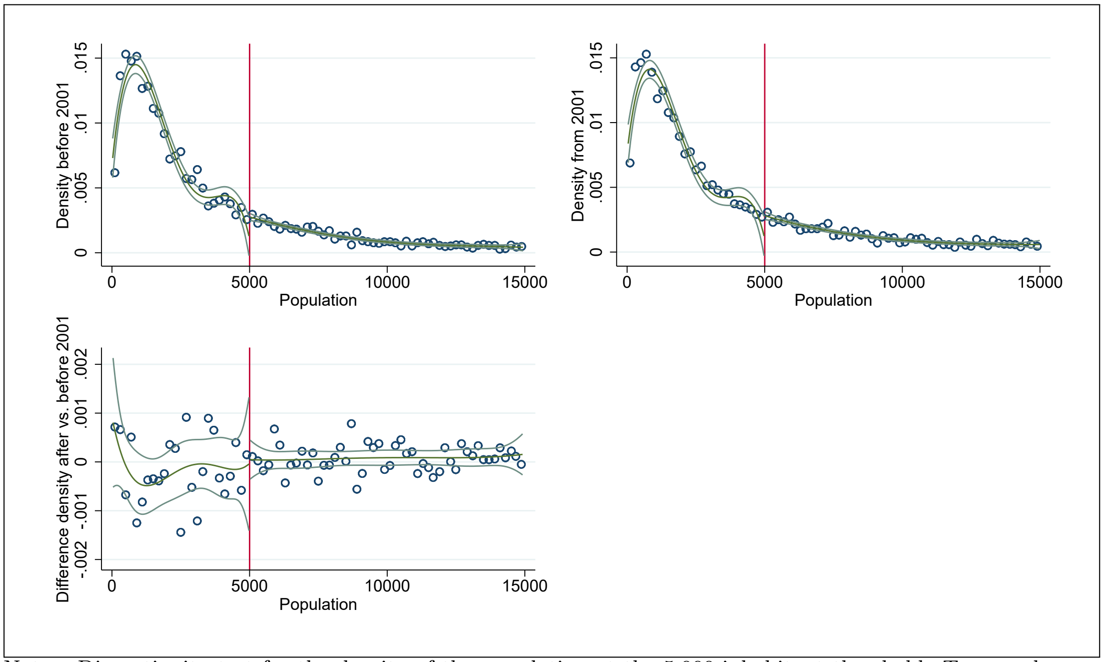
Notes. Diff-in-disc estimates of the impact of fiscal rules on the education of politicians. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2000. Variables in the Table: (>= 1999)*(> 5000)= interaction between dummy = 1 for electoral years 1999-2000 and dummy = 1 for municipalities with more than 5,000 inhabitants. The outcome variable is the share of mayoral candidates with a university degree in column 1, while it is equal to 1 for mayors with a university degree in column 2. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

Table A7: Heterogeneity analysis based on municipal budget rigidity
Alternative measures

	(1)	(2)	(3)	(4)
Dependent variable	*Mayoral candidates with university degree*		*Mayors with university degree*	
Control Function	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT
Sample	*Rigidity<median*	*Rigidity>median*	*Rigidity<median*	*Rigidity>median*
Panel A: personnel expenditures as share of current revenues				
(Post)*(> 5000)	-0.221	0.008	-0.224	-0.006
	(0.051)	(0.061)	(0.085)	(0.108)
Observations	1,959	2,031	2,022	1,632
Bandwidth	1127	1500	1160	1217
Mean outcome	0.425	0.516	0.462	0.533
Panel B: debt repayment expenditures as share of current revenues				
(Post)*(> 5000)	-0.125	-0.076	-0.188	-0.066
	(0.052)	(0.091)	(0.089)	(0.107)
Observations	2,214	944	2,110	1,570
Bandwidth	1369	663	1299	1077
Mean outcome	0.474	0.431	0.493	0.486

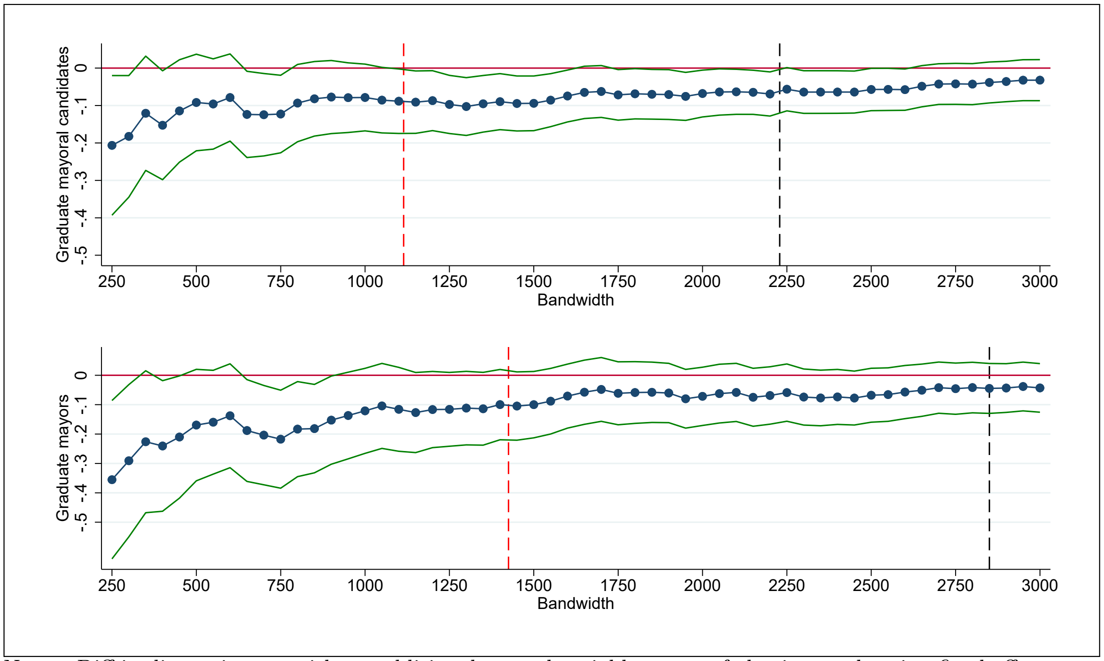
Notes. Diff-in-disc estimates of the impact of fiscal rules on the education of politicians. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Sub-samples: 1) *(Rigidity < median)* = municipalities with a below-median level of personnel (Panel A) or debt (Panel B) expenditures as a fraction of total current revenues; 2) *(Rigidity > median)* = municipalities with an above-median level of personnel (Panel A) and debt (Panel B) expenditures as a fraction of total current revenues. Personnel expenditures as a share of current revenues have an average value of 30.1 percent and debt repayment expenditures as a share of current revenues have an average value equal to 8.2 percent. Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The outcome variable is the share of mayoral candidates with a university degree in columns 1-2, and a dummy variable equal to 1 for mayors with a university degree in columns 3-4. Year of election and region fixed effects not included. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

Figure A1: Density test on the running variable



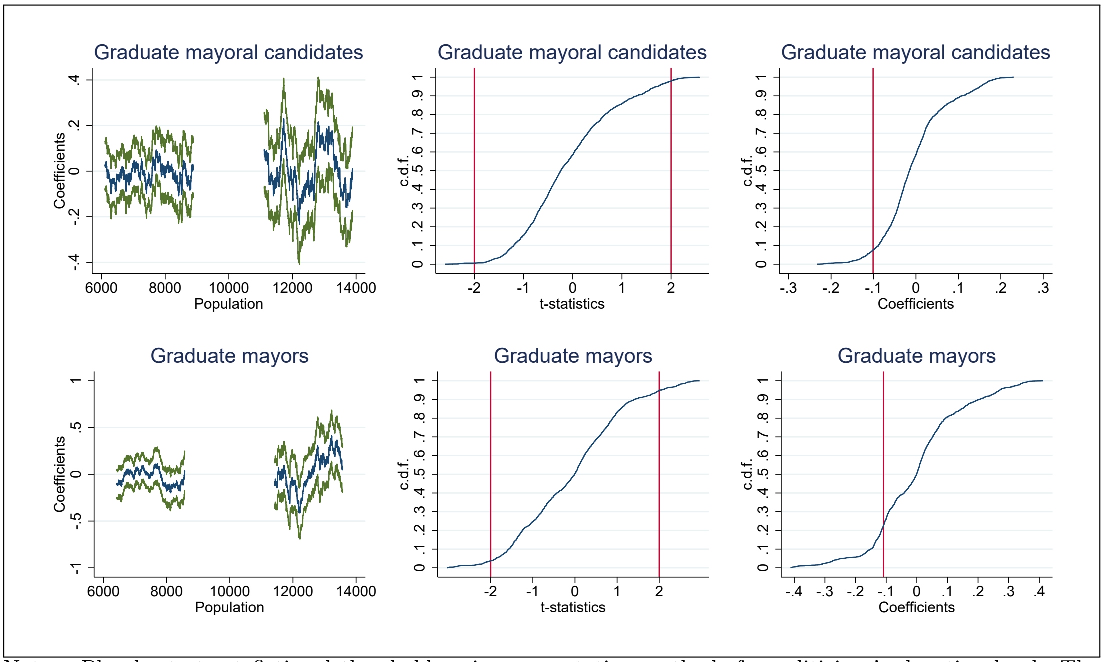
Notes. Discontinuity test for the density of the population at the 5,000-inhabitant threshold. Top graphs: (1) density test for R_{it} before 2001; (2) density test for R_{it} from 2001. Bottom graph: (1) discontinuity test for the difference between the density of average R_{it} from 2001 and the density of average R_{it} before 2001. The central green line represents a split fourth-order polynomial of the outcome variable in the normalized population, fitted separately on each side of the threshold. The grey lines represent the 95 percent confidence interval.

Figure A2: Diff-in-disc estimates: different bandwidths



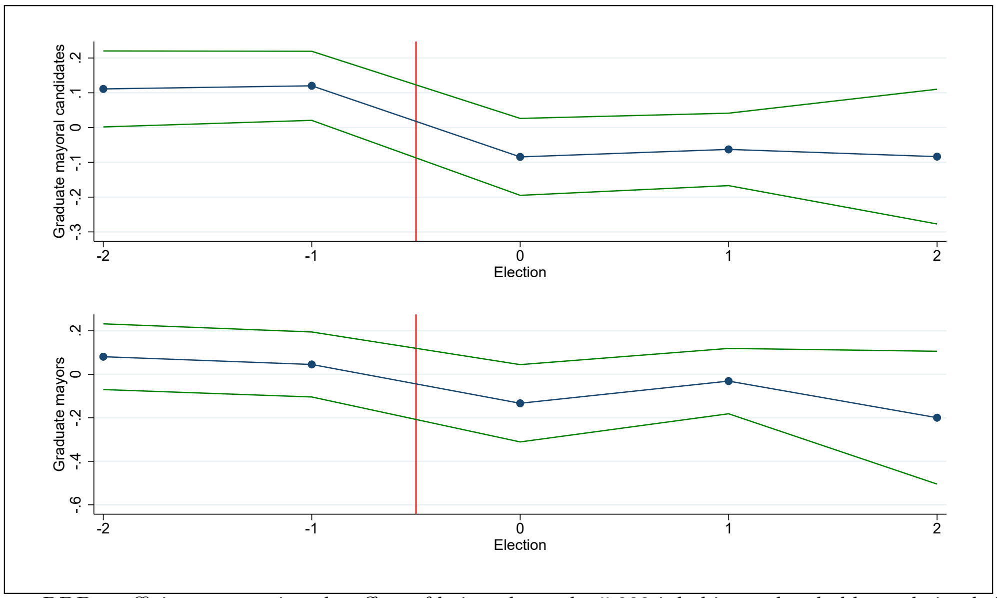
Notes. Diff-in-disc estimates without additional control variables, year of election, and region fixed effects. Horizontal axis: different bandwidths used to estimate the diff-in-disc coefficients. Vertical axis: diff-in-disc coefficients. Dashed red vertical line: optimal bandwidth calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Dashed black vertical line: double the optimal bandwidth. The central blue lines connect the estimated coefficients, while the green lines the 95 percent confidence intervals.

Figure A3: Diff-in-Disc
Placebo thresholds



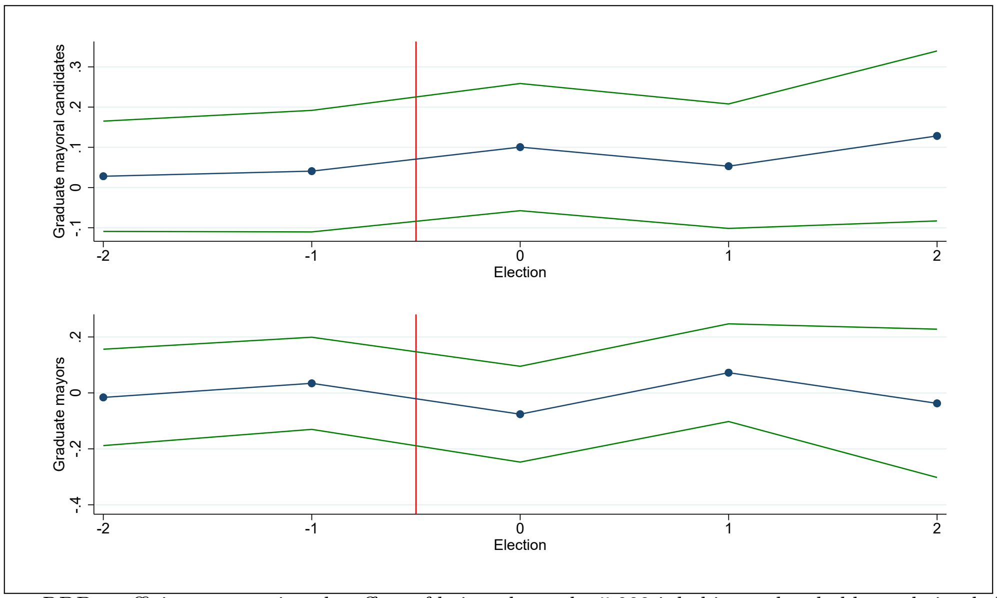
Notes. Placebo tests at fictional thresholds using permutation methods for politicians' education level. The figure reports the estimated coefficients, and c.d.f. of the t-statistics and estimated coefficients of a set of diff-in-disc regressions at 5,542 fictional thresholds for mayoral candidates and 4298 for mayors. The diff-in-disc model is run using a local linear regression with election year and region fixed effects. The graphs on the left report the estimated coefficients at the placebo thresholds with the corresponding population on the x-axis. In these graphs, the central blue lines represent the estimated coefficients, and the green lines the 95 percent confidence intervals. The graphs in the middle report the c.d.f. of the t-statistics associated with these coefficients. The vertical lines in these graphs indicate t-statistics of -2 and 2. The graphs on the right report the c.d.f. of the estimated coefficients. The vertical lines in these graphs indicate the benchmark estimates from Table 4, columns 3.

Figure A4: Cross-sectional RDD coefficients over time (low-rigidity sample)



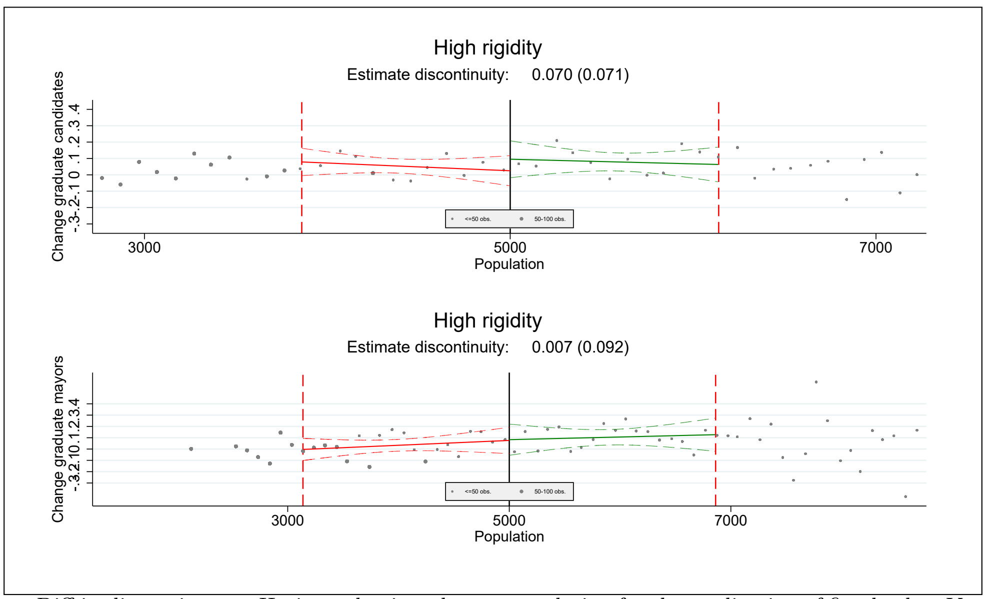
Notes. RDD coefficients capturing the effect of being above the 5,000-inhabitant thresholds vs. being below it. On the x-axis, which goes from -2 to 2, we report the elections before and after the 2001 removal of fiscal rules, where 0 indicates the elections immediately after the relaxation of fiscal rules. We run the cross-section RDD regressions using the optimal CCT bandwidths reported in Table 5. The blue lines connect the estimated coefficients, while the green lines represent the 95 percent confidence interval.

Figure A5: Cross-sectional RDD coefficients over time (high-rigidity sample)



Notes. RDD coefficients capturing the effect of being above the 5,000-inhabitant thresholds vs. being below it. On the x-axis, which goes from -2 to 2, we report the elections before and after the 2001 removal of fiscal rules, where 0 indicates the elections immediately after the relaxation of fiscal rules. We run the cross-section RDD regressions using the optimal CCT bandwidths reported in Table 5. The blue lines connect the estimated coefficients, while the green lines represent the 95 percent confidence interval.

Figure A6: Diff-in-Disc graphical evidence (high-rigidity sample)



Notes. Diff-in-disc estimates. Horizontal axis: relevant population for the application of fiscal rules. Vertical axis: the change over time in the share of mayoral candidates (top graps) and mayors (bottom graph) with a university degree. Scatter points are averaged over bins of 100 inhabitants. The central line represents a linear regression of the outcome variable in the population, fitted separately on each side of the threshold. The other two dashed lines represent 95 percent confidence intervals. The vertical dashed lines indicates the limit of the optimal bandwidth used in the regressions. Number of observations in each graph: 1) top graph: 3339 observations in total, 1510 within the optimal bandwidth; 2) bottom graph: 7367 observations in total, 2578 within the optimal bandwidth.

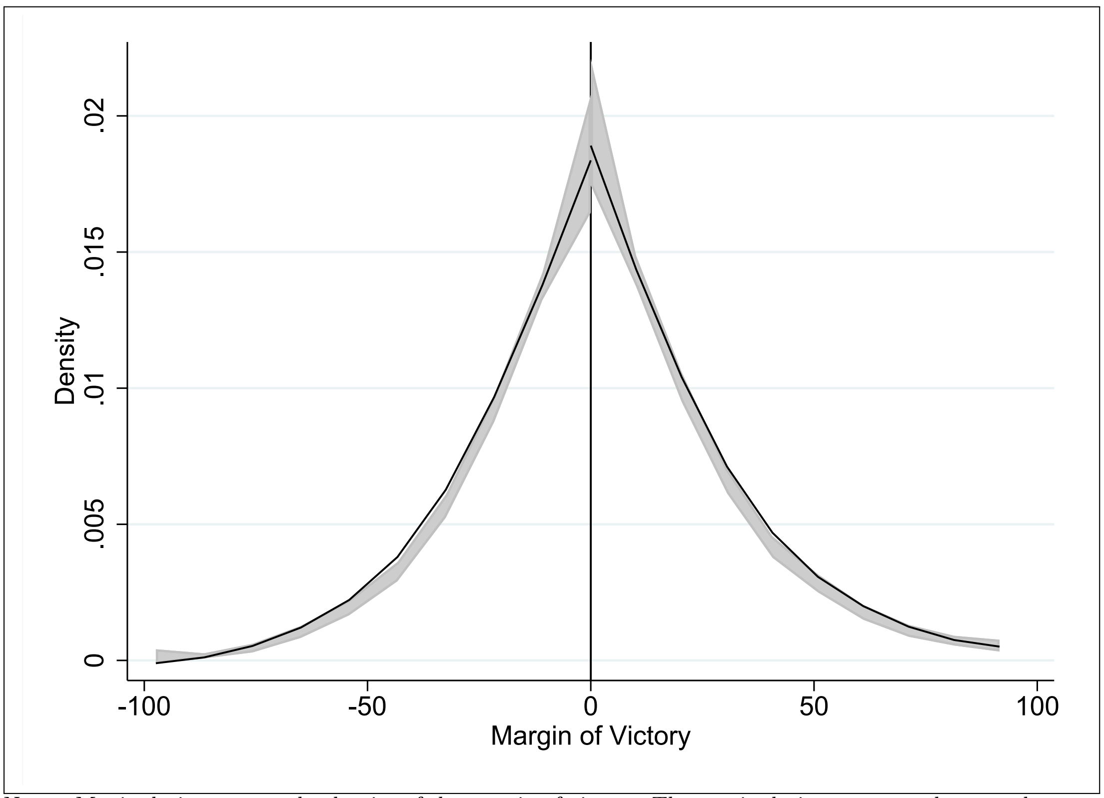
B Appendix: Matching the state of economy

B.1 Regression discontinuity design: main assumptions

There are two main assumptions required for the identification strategy described by model (12) in section IV.C.IV to work correctly. First, there must be no sorting around the threshold $MV_{it} = 0$, such that voters in municipalities with narrow mixed electoral competitions are not able to manipulate the running variable MV_{it}. We test this assumption in Figures B1 and B2, using the test on the continuity of the density of the running variable proposed by Cattaneo, Jansson, and Ma (2018). The evidence in Figures B1-B2 excludes that sorting is happening.

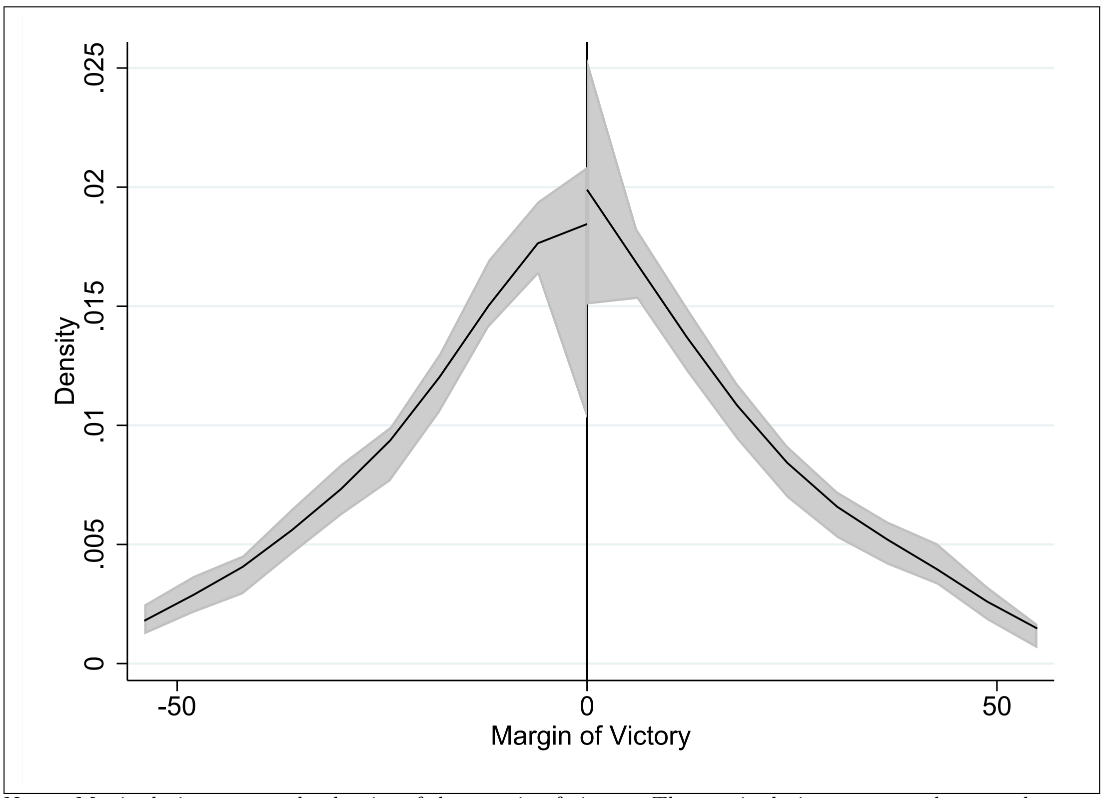
Second, observable municipal characteristics should vary smoothly at the threshold $MV_{it} = 0$. This assumption is required to guarantee that municipalities on one side of the threshold are a proper counterfactual for municipalities on the other side of the cutoff. We test this assumption in Tables B1 and B2, which confirm that municipal covariates are balanced.

Figure B1: Manipulation test on the margin of victory - Municipalities below 5,000



Notes. Manipulation test on the density of the margin of victory. The manipulation test uses the procedure developed by Cattaneo, Jansson, and Ma (2018). T-statistics: the conventional test statistics is 0.501, while the robust one is 0.679.

Figure B2: Manipulation test on the margin of victory - Municipalities above 5,000



Notes. Manipulation test on the density of the margin of victory. The manipulation test uses the procedure developed by Cattaneo, Jansson, and Ma (2018). T-statistics: the conventional test statistics is 0.582, while the robust one is 1.208.

Table B1: Balance test on municipal covariates
RDD, below 5000

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A: Characteristics municipal population							
Dependent variables	% university degree	% 15-64	% 65+	(log) income per capita	# firms	no-profit ass	area	population density
Graduate Mayor	-0.000	0.003	-0.004	-0.015	0.000	0.000	1.547	-6.117
	(0.002)	(0.005)	(0.007)	(0.036)	(0.002)	(0.000)	(2.742)	(22.834)
Observations	2281	2583	2643	2624	2737	3175	2873	3134
Bandwidth	15.24	17.91	18.53	18.32	19.38	23.93	20.79	23.48
	Panel B: Geographical characteristics municipalities							
Dependent variables	NE	NW	CEN	SOU	Past deficit			
Graduate Mayor	0.023	-0.032	-0.048	0.050	0.002			
	(0.036)	(0.061)	(0.043)	(0.060)	(0.005)			
Effective Observations	2897	2766	2336	2525	1850			
Bandwidth	21	19.62	15.63	17.34	17.44			

Notes. RDD estimates of the impact of graduate mayors on municipal covariates. Municipalities below 5,000 inhabitants. Electoral years between 2001 and 2012. Treatment variable: Graduate is a dummy variable equal to 1 when the mayor has a university degree, 0 otherwise. Estimation by RDD using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Municipal dependent variables in Panel A (measured in 2001): 1) share of population with a university degree; 2) share of active population (i.e. population between 15 and 64 years old); 3) share of elderly (i.e. population above 65 years old); 4) log of income per capita; 5) number of firms per capita; 6) number of non-profit associations per capita; 7) area of municipality in square km; 8) population density. In Panel B, the dependent variables are geographical dummy variables for different areas of Italy (i.e. North-West, North-East, Centre, South) and the deficit as a fraction of total revenues from the previous term. Robust standard errors clustered at the local labor area level are in parentheses.

Table B2: Balance test on municipal covariates
RDD, above 5000

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Panel A: Characteristics municipal population				
Dependent variables	% university degree	% 15-64	% 65+	(log) income per capita	# firms	no-profit ass	area	population density
Graduate Mayor	-0.003	-0.001	0.002	-0.009	-0.002	0.000	3.635	-119.001
	(0.003)	(0.005)	(0.006)	(0.058)	(0.004)	(0.000)	(6.463)	(105.907)
Observations	1098	1083	1013	1041	763	1024	1145	1072
Bandwidth	20.95	20.35	18.18	18.87	12.68	18.45	22.24	20.01
				Panel B: Geographical characteristics municipalities				
Dependent variables	NE	NW	CEN	SOU	Past deficit			
Graduate Mayor	0.105	-0.134	0.012	0.001	-0.000			
	(0.077)	(0.098)	(0.059)	(0.074)	(0.004)			
Observations	871	902	1122	1185	683			
Bandwidth	14.95	15.61	21.44	23.60	15.30			

Notes. RDD estimates of the impact of graduate mayors on municipal covariates. Municipalities between 5000 and 15,000 inhabitants. Electoral years between 2001 and 2012. Treatment variable: Graduate is a dummy variable equal to 1 when the mayor has a university degree, 0 otherwise. Estimation by RDD using the MSE-optimal bandwidth h selector per Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018). Municipal dependent variables in Panel A (measured in 2001): 1) share of population with a university degree; 2) share of active population (i.e. population between 15 and 64 years old); 3) share of seniors (i.e. population above 65 years old); 4) log of income per capita; 5) number of firms per capita; 6) number of non-profit associations per capita; 7) area of municipality in square km; 8) population density. In Panel B, the dependent variables are geographical dummy variables for different areas of Italy (i.e. North-West, North-East, Centre, South) and the deficit as a fraction of total revenues from the previous term. Robust standard errors clustered at the local labor area level are in parentheses.

B.2 Additional Tables described in section IV.C.IV (Results on education level and policy choice)

To address potential endogeneity in municipal income growth during a mayor's term, and consequently in the dependent variable shown in Table 6 in section IV.C.IV, we generate an alternative version of this variable. This is done by predicting income growth through regression analysis on pre-determined municipal characteristics, along with regional and year fixed effects. As indicated in Table B3, employing this alternative variable yields similar results.

Table B3: Graduate mayors and matching the state of the economy (with predicted income growth)

	(1)	(2)	(3)
Control Function	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT
Year of election FE	No	No	Yes
Region FE	No	No	Yes
Mayoral covariates	No	Yes	Yes
Panel A: municipalities below 5000			
Graduate Mayor	0.043	0.050	0.047
	(0.025)	(0.025)	(0.024)
Observations	2790	2756	2642
Bandwidth	19.85	19.50	18.50
Panel B: municipalities above 5000			
Graduate Mayor	-0.041	-0.040	-0.053
	(0.047)	(0.046)	(0.044)
Observations	910	911	785
Bandwidth	15.19	15.20	12.59
P-Value difference Panel A vs. B	0.109	0.086	0.046

Notes. Municipalities below 15,000 inhabitants. Electoral terms between 2001 and 2012. Dependent variable: probability of matching the state of the economy over the electoral mandate. In this Table, we use a predicted version of income growth. Treatment variable: Graduate is equal to 1 when mayor has a university degree, 0 otherwise. Estimation by RDD using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Mayoral covariates in columns 2-3: 1) age of the mayor; 2) political experience: years of past political experience of the mayor at any level of politics; 3) high skills job = 1 if mayor worked in a high skills occupation in the past; 4) female = 1 if mayor is a woman; 5) left = 1 for a center-left mayor. Robust standard errors clustered at the local labor area level are in parentheses. The last row presents the p-value for the test comparing whether the coefficients in panel A are the same to that in panel B.

The different results in Table 6 for municipalities below and above the

5,000-inhabitant threshold may be due to the different wages paid to the mayors. To rule out this possibility, we expand the initial dataset, including the 2013-2015 period, and repeat the RDD exercise using only those years, during which fiscal rules applied equally across the threshold. Conversely, during these years, the wage increase across the threshold was in place. Table B4 shows that the differences in matching the state of the economy disappear when fiscal rules apply in the same way across the threshold, as none of the estimated coefficients in the Table is statistically different from zero.

Table B4: Graduate mayors and matching the state of the economy
Years 2013-2015

	(1)	(2)	(3)
Control Function	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT
Year of election FE	No	No	Yes
Region FE	No	No	Yes
Mayoral covariates	No	Yes	Yes
Panel A: municipalities below 5000			
Graduate Mayor	-0.059	-0.053	-0.060
	(0.048)	(0.047)	(0.046)
Observations	1129	1146	1137
Bandwidth	15.52	15.85	15.64
Panel B: municipalities above 5000			
Graduate Mayor	-0.089	-0.061	-0.055
	(0.073)	(0.072)	(0.074)
Observations	476	479	426
Bandwidth	15.21	15.54	13.33
P-Value difference Panel A vs. B	0.736	0.922	0.957

Notes. Municipalities below 15,000 inhabitants. Years 2013-2015. Dependent variable = 1 in the event of above-median deficit coupled with below-median income growth or below-median deficit with above-median income growth. Treatment variable: Graduate is 1 when mayor has a university degree, 0 otherwise. Estimation by RDD using the MSE-optimal bandwidth h selector per Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018). Mayoral covariates in columns 2-3: 1) age of the mayor; 2) political experience = years of past political experience of the mayor at any level of politics; 3) high skills job = 1 if mayor worked in a high-skill occupation in the past; 4) female = 1 if mayor is a woman; 5) left = 1 for a center-left mayor.Robust standard errors clustered at the local labor area level are in parentheses. The last row presents the p-value for the test comparing whether the coefficients in panel A are the same to that in panel B.

In Table B5, we present additional evidence that more educated mayors are likelier to foster successful municipal administrations. This is based on outcomes that include investment expenditures, measures of fiscal sustainability, and the amount of public services provided. Data on investment expenditures and measures of fiscal sustainability comes from the municipal budget outcomes from the Aida PA database, an online archive managed by the Bureau Van Dijk. The data contains information on the fiscal items of the budgets of all Italian municipalities, covering the years 2000-2012. Data on investment expenditures and measures of fiscal sustainability are derived from municipal budget outcomes, as recorded in the Aida PA database. This online archive, managed by the Bureau Van Dijk, contains information on the fiscal items of budgets for all Italian municipalities, spanning the years 2000-2012.

To measure the amount of public services provided, we use an indicator developed through data collected by the Italian Ministry of Finance. This indicator is available through the website Opencivitas (www.opencivitas.it). Lockwood et al. (2021) provide an extensive description. The indicator measures the difference between the amount of services provided by one municipality and the standard level of services that should be provided, which, accordingly to the methodology developed by the Italian Ministry of Finance, corresponds to the average level of services provided by municipalities in the same population bracket. Using this continuous indicator, we build a dummy variable equal to one for municipalities providing a level of public services equal to or greater than the standard level of services.

The results in Panel A of Table B5 show that in municipalities without fiscal rules, graduate mayors are more likely to increase investment expenditures and to provide more public services compared to non-graduate mayors. In addition, we do not find differences in fiscal sustainability measures (i.e., deficit and debt repayment, see Vannutelli, 2022, for more detail) and current expenditures between graduate and non-graduate mayors. This evidence suggests that graduate mayors can produce better outcomes without worsening the sustainability of the municipal administration. Conversely, as shown in Panel B of Table B5, these differences disappear in municipalities with fiscal rules.

Table B5: Performance of graduate mayors

	(1)	(2)	(3)	(4)	(5)
Dep. Variables	Current	Capital	Services	Deficit	Loan
Dep. Variables	Expenditures	Expenditures	Provided		Repayment
Control Function	Linear	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT	CCT
Year of election FE	Yes	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes	Yes
Mayoral covariates	Yes	Yes	Yes	Yes	Yes
Panel A: municipalities below 5000					
Graduate Mayor	8.072	150.462	0.131	-2.241	-4.777
	(36.099)	(79.769)	(0.057)	(5.385)	(11.052)
Observations	2191	2311	1031	1814	3173
Bandwidth	14.47	15.43	24.30	11.50	23.85
Panel B: municipalities above 5000					
Graduate Mayor	-21.218	3.734	0.041	2.689	19.024
	(32.985)	(30.430)	(0.095)	(4.716)	(16.325)
Observations	1079	857	257	949	843
Bandwidth	18.90	13.97	12.43	16.04	13.71
P-Value difference Panel A vs. B	0.549	0.086	0.416	0.491	0.227

Notes. RDD estimates. Municipalities below 15,000 inhabitants. Electoral terms between 2001 and 2012. Dependent variables: 1) current expenditures = municipal current expenditures; 2) capital expenditures = municipal capital expenditures; 3) services provided = dummy variable equal to 1 if municipality provided in 2010 an average level of public services equal or above the standard level; 4) deficit = total revenues - total expenditures; 5) loan repayment = loan repayment expenditures. Treatment variable: Graduate is equal to 1 when mayor has a university degree, 0 otherwise. Estimation by RDD using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Mayoral covariates in all columns: 1) age of the mayor; 2) political experience: years of past political experience of the mayor at any level of politics; 3) high skills job = 1 if mayor worked in a high skills occupation in the past; 4) female = 1 if mayor is a woman; 5) left = 1 for a center-left mayor. Robust standard errors clustered at the local labor area level are in parentheses. The last row presents the p-value for the test comparing whether the coefficients in panel A are the same to that in panel B.

C Appendix: Alternative stories

We offer more detailed insights into how we address the alternative explanations outlined in section IV.C.V. First, the application of fiscal rules may require the selection of more politically experienced politicians, who may be less educated. To rule out the latter explanation, we run the *Diff-in-Disc* model on other personal characteristics of local politicians, such as past professional background, age, gender, and past political experience. It is important to highlight that, due to data limitations, it was only possible to reconstruct the past political experience for elected mayors, and not for mayoral candidates. We report the results of this exercise in Table C1. For characteristics potentially correlated with education, the estimated coefficient goes in the expected direction (i.e., a decline in the share of politicians from skilled occupations). On the other hand, gender and years of political experience do not seem to be affected by fiscal rules. The lack of an effect for political experience rules out the possibility that the application of fiscal rules may require the selection of more politically experienced politicians.

In addition, in Table C2, we check whether fiscal rules negatively affected municipal councilors' education level. Specifically, as described in section I., our expectation about the effect of fiscal rules was that these should affect politicians in powerful positions, like mayors, rather than politicians in less prominent positions, like municipal councilors. In line with this expectation, Table C2 reports coefficients that, even though negative, are small and not statistically significant.

Third, we show that different non-political outside options for individuals with different education levels are unlikely to explain our results. In principle, fiscal rules may affect the value of public office for individuals with different

Table C1: The effect of fiscal rules on other characteristics

	(1)	(2)	(3)	(4)
Control Function	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT
Election Year FE	No	No	No	No
Region FE	No	No	No	No
Dependent Variables	High skill	Age	Female	Pol Experience
Panel A: mayoral candidates				
(Post)*(> 5000)	-0.092	1.090	0.001	
	(0.043)	(0.810)	(0.024)	
Observations	2,944	4,549	3,637	
Bandwidth	952.1	1482	1180	
Mean outcome	0.286	47.95	0.112	
Panel B: mayors				
(Post)*(> 5000)	-0.089	1.277	0.011	-0.592
	(0.062)	(1.445)	(0.034)	(0.750)
Observations	3,510	3,554	3,596	4,156
Bandwidth	1158	1168	1172	1339
Mean outcome	0.309	47.89	0.087	8.182

Notes. Diff-in-disc estimates of the impact of fiscal rules on politicians' characteristics. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The outcome variables are 1) high skill: politicians from high-skill occupations; 2) Age: age of the politicians; 3) Female = 1 for female politicians; 4) Pol Experiences = years of political experience at any level of politics (for mayors only). The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

levels of education in the same way. This homogeneous effect could then affect the entry into politics of individuals with different levels of education heterogeneously, given their different outside options. If higher-educated individuals have a better outside option in the labor market compared to less educated ones, the overall effect could be a reduction in the quality of candidates. Table C3 appears to rule out this alternative story. Specifically, in Table C3, we use data on the municipal shares of employed individuals divided by income

Table C2: The effect of fiscal rules on the education of municipal councilors

	(1)	(2)	(3)
Control Function	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT
Election Year FE	No	Yes	Yes
Region FE	No	No	Yes
(Post)*(> 5000)	-0.026	-0.030	-0.032
	(0.021)	(0.020)	(0.020)
Observations	3,742	3,742	3,742
Bandwidth	1221	1221	1221
Mean outcome	0.263	0.263	0.263

Notes. Diff-in-disc estimates of the impact of fiscal rules on the education level of municipal councilors. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting after 2001. The outcome variable is the share of municipal councilors with a university degree. The bandwidth is calculated using the MSE-optimal bandwidth h selector following Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

brackets to measure how concentrated opportunities in the labor market are. To do so, we calculate a Herfindahl index of these income brackets' share to measure whether employed individuals are concentrated in one or more specific income brackets. Higher values of this index suggest a greater concentration in one specific bracket and, thus, more homogeneous opportunities for individuals in that municipality, independently of the level of education. The results in Table C3 indicate that the findings are similar across municipalities with low vs. high values of the Herfindahl index, and, if anything, municipalities with a higher Herfindahl index (i.e., where outside options are homogeneous) present stronger results. The fact that the results are stronger in municipalities where outside options are homogeneous suggests that it is unlikely that different options outside of politics across individuals with different levels of education explain our results.

Table C3: Effect of fiscal rules and outside option in the private sector

	(1)	(2)	(3)	(4)
Dependent Variables	Share mayoral candidates with university degree		= 1 for Mayors with university degree	
Control Function	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT
Election Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Sample	Herfindal index > median	Herfindal index < median	Herfindal index > median	Herfindal index < median
(Post)*(> 5000)	-0.125	-0.080	-0.131	-0.083
	(0.044)	(0.066)	(0.069)	(0.103)
Observations	2,304	1,384	2,713	1,526
Bandwidth (h)	1245	1162	1510	1263
Mean outcome	0.406	0.571	0.422	0.574

Notes. Diff-in-disc estimates. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5,000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The outcome variable is the share of mayoral candidates with a university degree in columns 1-2, and a dummy variable equal to 1 for mayors with a university degree in columns 3-4. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018). Election year and region fixed effects added in all columns. Robust standard errors clustered at the local labor area level are in parentheses.

Fourth, fiscal rules may affect politicians' political orientation, which in turn is correlated with their level of education. As an example, fiscal rules may make political office less attractive for left wing perspective candidates, and this could be positively correlated with income and education (Gethin et al., 2022). Table C4 excludes any effect of fiscal rules on politicians' political orientation.

Fifth, fiscal rules may change the desirability of electing a highly-educated mayor, from the voters' perspective. In particular, they may make competence less important, hence reducing the advantage of highly educated politicians. If true, this effect would also be a potential channel for our result, as highly educated politicians would be discouraged to run, with fiscal rules, because they anticipate the reduction in their electoral advantage. However, it does

Table C4: The effect of fiscal rules on ideology

	(1)	(2)	(3)	(4)
Control Function	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT
Election Year FE	No	No	No	No
Region FE	No	No	No	No
Dependent Variables	Left	Right	Center	Civic List
Panel A: mayoral candidates				
(Post)*(> 5000)	0.026	-0.008	0.005	-0.047
	(0.036)	(0.034)	(0.012)	(0.049)
Observations	3,537	4,067	4,778	3,287
Bandwdith	1151	1311	1549	1062
Mean outcome	0.201	0.214	0.021	0.566
Panel B: mayors				
(Post)*(> 5000)	0.033	-0.037	0.010	-0.099
	(0.058)	(0.052)	(0.012)	(0.070)
Observations	4,023	3,841	5,194	3,261
Bandwdith	1305	1245	1680	1060
Mean outcome	0.256	0.175	0.012	0.550

Notes. Diff-in-disc estimates of the impact of fiscal rules on the ideology of politicians. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The outcome variables are: 1) Left = share of center-left candidates in Panel A, =1 for center-left mayors in Panel B; 2) Right = share of center-right candidates in Panel A, =1 for center-right mayors in Panel B; 3) Center = share of center candidates in Panel A, =1 for center mayors in Panel B; 4) Civic lists = share of independent candidates in Panel A, =1 for independent mayors in Panel B. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

not seem to be the case in our data. More in detail, in Table C5, we use data at the mayoral candidate level and OLS to show that graduate mayoral candidates have better electoral performances than non-graduate ones independently of whether fiscal rules apply or not, in races where at least one highly educated candidate is running. Specifically, graduate candidates receive more votes, reach a better final ranking position, and are more likely to be elected. The results go in the same direction irrespective of whether we consider municipalities and electoral years with fiscal rules or without them.

In addition, in Table C6 we look at the effect of fiscal rules on the probability of having a mayor with a university degree, splitting the sample between municipalities with a pre-treatment share of high education candidates above and below the median. The effect seems to be stronger in the latter group, although the coefficients are not statistically significant. This suggests that voters may partially "correct" for the reduced number of highly educated candidates, by voting for them when available. However, this compensation appears less feasible when the reduction induced by fiscal rules implies that no highly educated candidates are running. Those pieces of evidence should be seen as suggestive, rather than causal, as fiscal rules may change not only the number of high-education candidates, but also their type (in dimensions other than education), and this may be endogenous as well. However, the fact that voters do not change their behaviour seems to suggest that any endogenous selection process on characteristics different from education is not too relevant. One possible explanation for the fact that voters do not seem to change their behaviour with fiscal rules is that fiscal policies are just one of the several tasks a mayor is supposed to do. Hence, voters may think that human capital has a positive impact on other tasks as well, hence keeping (roughly) the same preferences even when fiscal rules constrain fiscal policies.

Sixth, educated mayors may be more corrupt than non-graduate ones. Daniele and Giommoni (2020) show that the introduction of fiscal rules should make it more challenging to extract rents, reducing the office value for individuals attracted by them. If graduates are more corrupt than non-graduates, the introduction of fiscal rules may make them less interested in entering politics. However, this does not seem to be the case. Using the Mafia index built by Calderoni (2011), which quantifies the presence of Mafia-style criminal organizations in Italian provinces, we run model (11) splitting the sample between municipalities in provinces below vs. above the median of mafia presence. As shown in table C7, the negative effect of fiscal rules on the education of mayoral candidates is driven by municipalities in provinces with low mafia presence. These are the municipalities where corruption is less of an issue. Furthermore, as we can see from Table C8, graduate mayors do not appear to be more corrupt than non-graduate ones. More in detail, to measure corruption, we use the web archive of one of the leading Italian newspapers (La Repubblica) to find episodes of corruption linked to the mayors in the analysis. Using an algorithm based on the mayor's first and last names, the name of the city, the years of the legislature, and a series of keywords related to episodes of corruption, we create a database of newspaper articles reporting episodes of corruption linked to the mayors in the dataset. We use this database to create a dummy variable equal to 1 for mayors found to be corrupt, and 0 otherwise. The coefficients reported in Table C8 are estimated using this dummy variable as the dependent variable.

Finally, as described and tested in section IV.C.I and Table A6, we do not find interactive effects between the 1999 introduction of fiscal rules and the differential wage paid across the 5000 inhabitants threshold. To further check that this is the case, in Table C9, we replicate the main analysis of Tables 4 and 5 by keeping only the electoral years from 1999 (i.e., excluding prior elections in which fiscal rules were not implemented in any municipality) and the sub-sample of municipalities that effectively held an election in either the electoral years 1999 or 2000 (i.e., election years in which fiscal rules applied uniformly across the 5000 inhabitants threshold). The idea of this exercise is to repeat the analysis by keeping a pre-treatment period in which the application of fiscal rules is constant over time and across the threshold. As we can see from Panel A of Table C9, the results for the mayoral candidates (i.e., the main focus of our theoretical and empirical analysis) are essentially unchanged, even though less precisely estimated due to the lower number of observations. The results for mayors in Panel B are somehow weaker (i.e., smaller coefficients and not statistically significant), but they are qualitatively similar (i.e., negative and economically significant coefficients in the entire and low rigidity samples, small and positive coefficients in the high rigidity sample).

Table C5: Candidate level regressions: graduate vs. non-graduate candidates

	(1)	(2)	(3)
Dependent Variables	Vote Shares	Ranking Position	=1 if elected Mayor
	Panel A: all elections		
Graduate	5.951	-0.222	0.086
	(0.239)	(0.011)	(0.007)
Observations	41,086	41,185	41,185
	Panel B: fiscal rules applied		
Graduate	6.016	-0.242	0.088
	(0.384)	(0.020)	(0.011)
Observations	14,080	14,092	14,092
	Panel C: fiscal rules did not applied		
Graduate	5.891	-0.210	0.085
	(0.280)	(0.013)	(0.008)
Observations	27,006	27,093	27,093

Notes. OLS estimates. Municipalities below 15,000 inhabitants. Electoral terms between 1993 and 2012. Only electoral races with at least one graduate candidate. Dependent variables: 1) vote shares = vote share taken by mayoral candidate; 2) ranking position = position of the candidate in the final ranking of mayoral candidates; 3) =1 if elected mayor = 1 if candidate elected mayor. Independent variable reported in the Table is = 1 for mayoral candidates with a university degree, 0 otherwise. Election year and region fixed effects included in all columns. Mayoral candidate covariates included in all columns: 1) high skills job = 1 if candidate worked in a high-skill occupation in the past; 2) female = 1 if candidate is a woman; 3) age = age of the mayoral candidate; 4) independent = 1 if candidate is not affiliated to a national political party; 5) unemployed = 1 if candidate is unemployed. Municipal covariates in all columns (measured in 2001, except for numbers 5 and 6, which are measured in 2005): 1) share of population with a university degree; 2) share of active population (i.e. population between 15 and 64 years old); 3) share of seniors (i.e. population above 65 years old); 4) log of income per capita measured in 2001; 5) number of firms per capita; 6) number of non-profit associations per capita; 7) area of municipality in square km; 8) population density. Robust standard errors clustered at the municipality level are in parentheses.

Table C6: Effect of fiscal rules and pre-treatment share of graduate candidates

	(1)	(2)	(3)	(4)
Dependent Variable:	= 1 mayors with university degree			
Control Function	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT
Election Year FE	No	No	Yes	Yes
Region FE	No	No	Yes	Yes
Sample	Low share graduate	High share graduate	Low share graduate	High share graduate
(Post)*(> 5000)	-0.126	-0.057	-0.115	-0.077
	(0.101)	(0.085)	(0.102)	(0.085)
Observations	1,119	2,323	1,119	2,323
Bandwidth	1075	1143	1075	1143
Mean outcome	0.224	0.634	0.224	0.634

Notes. Diff-in-disc estimates of the impact of fiscal rules on the education of mayoral candidates. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. High and low share of graduate candidates before treatment is measured using elections between 1993 and 2000. The sample split is at the median. Variables: 1) (> 5000) = 1 for municipalities with more than 5,000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The outcome variable is the share of mayoral candidates with a university degree in all columns. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018). Election year and region fixed effects added in columns 3 and 4. Robust standard errors clustered at the local labor area level are in parentheses.

Table C7: The role of criminal organizations

	(1)	(2)	(3)	(4)
Dependent variable	*Mayoral candidates with university degree*		*Mayors with university degree*	
Control Function	Linear	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT	CCT
Election Year FE	Yes	Yes	Yes	Yes
Region FE	Yes	Yes	Yes	Yes
Sample	*Mafia index<median*	*Mafia index>median*	*Mafia index<median*	*Mafia index>median*
(Post)*(> 5000)	-0.141	0.001	-0.274	0.051
	(0.062)	(0.048)	(0.096)	(0.086)
Observations	1,476	2,468	1,722	1,994
Bandwidth	927.8	1613	1078	1304
Mean outcome	0.409	0.513	0.432	0.545

Notes. Diff-in-disc estimates of the impact of fiscal rules on the education of politicians. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. Sub-samples: Mafia index<median if municipality located in a province with a low presence of Mafia-style criminal organizations; Mafia index>median if municipality located in a province with a high presence of Mafia-style criminal organizations. The mafia index comes from Calderoni (2011). Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001. The outcome variable is the share of mayoral candidates with a university degree in column 1-2 and is equal to 1 for mayors with a university degree in column 3-4. The bandwidth is calculated using the MSE-optimal bandwidth h selector per Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

Table C8: The effect of graduate mayors on corruption

		(1)	(2)	(3)	(4)
Control Function		Linear	Linear	Linear	Linear
Bandwidth		CCT	CCT	CCT	CCT
Year of election FE		No	Yes	No	Yes
Region FE		No	Yes	No	Yes
Covariates		No	Yes	No	Yes
Municipalities		Below 5000		Above 5000	
		Dependent variable = 1 if mayor corrupt			
Graduate Mayor		-0.008	0.006	-0.008	-0.027
		(0.015)	(0.014)	(0.041)	(0.038)
Effective Observations		2654	2319	1015	907
Bandwidth		18.60	15.49	17.52	15.12
	Descriptive statistics dummy variable for corruption				
	Mean	St. deviation	Min	Max	Observations
	0.098	0.206	0	1	6694

Notes. Municipalities below 15,000 inhabitants. Electoral terms between 2001 and 2012. Treatment variable: Graduate is a dummy variable =1 when the mayor has a university degree, 0 otherwise. Estimation by RDD using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Mayoral covariates included in columns 2 and 4: 1) female = 1 if mayor is a woman; 2) age = age of the mayor at the beginning of the term; 3) political experience = years of past political experience of the mayor at any level of politics; 4) left = 1 for center-left mayor; 5) high skills job = 1 if mayor worked in a high-skill occupation in the past. Robust standard errors clustered at the local labor area level are in parentheses.

Table C9: The effect of fiscal rules on the education of politicians
Election years from 1999

	(1)	(2)	(3)
Control Function	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT
Election Year FE	Yes	Yes	Yes
Region FE	Yes	Yes	Yes
Sample	*All sample*	*Rigidity<median*	*Rigidity>median*
Panel A: mayoral candidates with university degree			
(Post)*(> 5000)	-0.084	-0.182	0.032
	(0.067)	(0.072)	(0.111)
Observations	1,675	1,167	686
Bandwidth	1114	1279	1140
Mean outcome	0.445	0.411	0.494
Panel B: mayors with university degree			
(Post)*(> 5000)	-0.057	-0.084	0.025
	(0.083)	(0.093)	(0.122)
Observations	2,171	1,267	1,196
Bandwidth	1425	1386	1862
Mean outcome	0.449	0.434	0.468

Notes. Diff-in-disc estimates of the impact of fiscal rules on the education level of politicians. Original sample: municipalities between 0 and 15,000 inhabitants. Electoral terms between 1999 and 2012 and only municipalities that voted in election year 1999 or election year 2000. Variables in the table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting after 2001. The outcome variable is the share of mayoral candidates with a university degree in Panel A, and a dummy variable equal to 1 for mayors with a university degree in Panel B. The bandwidth is calculated using the MSE-optimal bandwidth h selector following Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). Robust standard errors clustered at the local labor area level are in parentheses.

D Additional details on empirical models

D.1 Difference-in-discontinuity model

We estimate the Difference-in-Discontinuity (*Diff-in-Disc*) model described in equation (11) with a local linear regression (Gelman and Imbens, 2018), using the subsample of observations that lie within the interval $R_{it} \in [-h, +h]$ around the threshold. The optimal bandwidth h is calculated using the MSE-optimal bandwidth following Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018). More in detail, to leverage the panel structure of our dataset, which includes multiple electoral years and time observations for each municipality, we follow the approach of Grembi et al. (2016).

Specifically, we estimate model (11) using the statistical software Stata and the command "regress", assigning equal weight to all observations within the optimal bandwidth h. Accordingly, the MSE-optimal bandwidth h is calculated in Stata using the "rdrobust" command (Calonico et al., 2017), with the option set for a rectangular kernel.

D.2 Regression discontinuity design model

We estimate model (12) using local linear regression (Gelman and Imbens, 2018) on a subsample of municipalities within the interval $MV_{it} \in [-h, +h]$. The optimal bandwidth h is determined based on the MSE-optimal bandwidth criteria from Calonico, Cattaneo, and Titiunik (2014), and Calonico, Cattaneo, and Farrell (2018). This estimation is conducted in Stata with the "rdrobust" command (Calonico et al., 2017). In line with the guidance of Calonico, Cattaneo, and Titiunik (2014), and Cattaneo, Idrobo, and Titiunik (2020), we employ robust inference methods and we weight observations by their proximity to the cutoff using a triangular kernel. The "rdrobust" command provides RDD estimates with a conventional variance estimator (Conventional), bias-corrected RDD estimates with a conventional variance estimator (Bias-corrected), and bias-corrected RDD estimates with a robust variance estimator (Robust). For simplicity, in all tables that present estimates from model (12), we report RDD estimates using the conventional variance estimator (Conventional). We have confirmed that the results and evidence from both bias-corrected RDD estimates with a conventional variance estimator (Bias-corrected) and those with a robust variance estimator (Robust) are essentially identical.

E Appendix to the theoretical framework

E.1 Formal analysis of the model

To ease the notation, we define the expected payoffs from being in office as follows:

$$h := E + k((1-\tau)\phi^H + \tau) \tag{E.1}$$

$$l := E + k((1-\tau)\phi^L + \tau) \tag{E.2}$$

$$f := E + k(1-\tau)(1-p) \tag{E.3}$$

E.1.1 No fiscal rules

When there are no fiscal rules, any elected politician is free to choose the policy once in office. As a consequence, at the policy stage biased politicians choose $x = 1$, unbiased politicians choose $x = s$ and hence they pick the correct policy with probability ϕ^Γ.

Lemma E1. *Without fiscal rules, there is a PBNE whose policy choice is as follows*

- *Biased politicians always choose* $x = 1$*;*
- *Unbiased politicians choose* $x = s$*.*

Proof of Lemma E1. Once in office, politicians learn their bias and there is no trade off with respect to their favourite policy. Hence, biased politicians choose $x = 1$ irrespective of the state. Unbiased politicians always choose $x = s$, because $\phi^L > max[p, 1-p]$, hence the signal realization always indicates the most likely state of the world.

To see this, note that if $\phi > p$

$$Pr(\theta = 0|s = 0) = \frac{\phi(1-p)}{\phi(1-p) + (1-\phi)p} > \frac{(1-\phi)p}{(1-\phi)p + \phi(1-p)} = Pr(\theta = 1|s = 0)$$

by Bayes' rule, and if $\phi > 1 - p$

$$Pr(\theta = 1|s = 1) = \frac{\phi p}{\phi p + (1-\phi)(1-p)} > \frac{(1-\phi)(1-p)}{\phi p + (1-\phi)(1-p)} = Pr(\theta = 0|s = 1)$$

Ex ante (i.e. before observing the signal realization), a politician with signal precision ϕ expects to choose the policy that matches the state, if she follows the signal realization (i.e. if she chooses $x = s$), with probability ϕ. To see this, note that, from an ex ante perspective,

$$Pr(s = \theta) = (1-p)Pr(s = 0|\theta = 0) + pPr(s = 1|\theta = 1) = \phi$$

■

The voter anticipates the equilibrium choices described above. Since higher educated unbiased politicians behave in a better way, in expectation, V prefers to elect the candidate with $\Gamma = H$ when the election is contested.

Lemma E2. *If there are two candidates of different education level, $\gamma_H^i = 1$.*

Proof of Lemma E2. At the voting stage, V anticipates the policy choices outlined in Lemma E1. Suppose two candidates of different education level run: from V's point of view, the expected utility of choosing the H candidate is $\mathbb{E}u^V(\Gamma = H) = \tau p + (1-\tau)\phi^H$, because the biased politician matches the state with probability p and the unbiased one with probability ϕ^Γ. It is easy to see that $\mathbb{E}u^V(\Gamma = H) > \mathbb{E}u^V(\Gamma = L) = \tau p + (1-\tau)\phi^L$ because $\phi^H > \phi^L$. ■

Combining these results, we can derive the relevant thresholds in w.

Lemma E3. *Without fiscal rules, there exists a symmetric PBNE where the entry threshold of politicians is defined by*

$$\bar{w}_H = \frac{4W^H\left(E + ((1-\tau)\phi^H + \tau)k\right)}{4W^H + (E + ((1-\tau)\phi^H + \tau)k)}$$

$$\bar{w}_L = \frac{4W^H - \left(E + ((1-\tau)\phi^H + \tau)k\right)}{4W^H + (E + ((1-\tau)\phi^H + \tau)k)} \frac{4W^L\left(E + ((1-\tau)\phi^L + \tau)k\right)}{4W^L + (E + ((1-\tau)\phi^L + \tau)k)}$$

Proof of Lemma E3. Start from an H politician. She compares (1) and (2), choosing to enter iff $w^i \leq \gamma^i_H(E + k\mathbb{E}_{b,\theta,s}u^P_H)$. Given Lemma E1, it is clear that $\mathbb{E}_{b,\theta,s}u^P_H = ((1-\tau)\phi^H + \tau)$. Given Lemma E2, it is clear that $\gamma^i_H = 1 - \frac{1}{2}\frac{1}{2}p^c_H$, where p^c_H is the conjectured probability that an opponent perspective candidate of H type chooses to run. Moving to an L politician, the logic on $\mathbb{E}_{b,\theta,s}u^P_L$ is the same. However, she knows she can win office only if H does not run, hence with probability $\left(1 - \frac{1}{2}\frac{1}{2}p^c_L - \frac{1}{2}p^c_H\right)$, because the L type loses for sure against an H opponent and with probability 0.5 against a L opponent. In a symmetric equilibrium, strategies must be the same for players of the same type and conjectured probabilities of running must be correct, thus $p^c_L = \frac{\bar{w}_L}{W^L}$ and $p^c_H = \frac{\bar{w}_H}{W^H}$. As a consequence, the symmetric equilibrium thresholds are the solution of the following system of equations:

$$\bar{w}_H = \left(1 - \frac{1}{2}\frac{1}{2}\frac{\bar{w}_H}{W^H}\right)\left(E + ((1-\tau)\phi^H + \tau)k\right)$$
$$\bar{w}_L = \left(1 - \frac{1}{2}\frac{1}{2}\frac{\bar{w}_L}{W^L} - \frac{1}{2}\frac{\bar{w}_H}{W^H}\right)\left(E + ((1-\tau)\phi^L + \tau)k\right)$$

We solve the system starting from $\bar{w}_H$ and using (E.1) and (E.2) to ease the notation.

$$
\begin{aligned}
\bar{w}_H &= \left(1 - \frac{1}{2}\frac{1}{2}\frac{\bar{w}_H}{W^H}\right) h \\
\bar{w}_H &= \left(\frac{4W^H - \bar{w}_H}{4W^H}\right) h \\
\bar{w}_H(4W^H + h) &= 4W^H h \\
\bar{w}_H &= \frac{4W^H h}{4W^H + h}
\end{aligned}
$$

Substituting in the second equation, we solve for $\bar{w}_L$:

$$
\begin{aligned}
\bar{w}_L &= \left(1 - \frac{1}{2}\frac{1}{2}\frac{\bar{w}_L}{W^L} - \frac{1}{2}\frac{\bar{w}_H}{W^H}\right) l \\
\bar{w}_L &= \left(1 - \frac{\bar{w}_L}{4W^L} - \frac{1}{2W^H}\frac{4W^H h}{4W^H + h}\right) l \\
\bar{w}_L \left(1 + \frac{l}{4W^L}\right) &= \left(1 - \frac{2h}{4W^H + h}\right) l \\
\bar{w}_L \left(\frac{4W^L + l}{4W^L}\right) &= \left(\frac{4W^H - h}{4W^H + h}\right) l \\
\bar{w}_L &= \frac{4W^H - h}{4W^H + h}\frac{4W^L l}{4W^L + l}
\end{aligned}
$$

Hence, we find that in our symmetric equilibrium (which is unique conditional on our indifference breaking assumptions)

$$
\begin{aligned}
\bar{w}_H &= \frac{4W^H h}{4W^H + h} \\
\bar{w}_L &= \frac{4W^H - h}{4W^H + h}\frac{4W^L l}{4W^L + l}
\end{aligned}
$$

■

Lemma E4. *Conditional on the assumptions on the tie-breaking rules, there are no symmetric PBNE leading to strategies different than those described in*

Lemma E3.

Proof of Lemma E4. To prove the statement, note the following:

- Policy choices are strictly dominant strategies for the different types of politicians, once the type is realized, hence they are the sole sequentially rational strategy and they must be the same in every equilibrium;

- Given the anticipated and uniquely defined policy choices, the voting choice of the voter is uniquely defined, meaning that in every equilibrium the voter would have a unilateral profitable deviation with any alternative choice than opting for the H candidate whenever available. When there is only one candidate the voter does not play any role.

- Move now to perspective candidates' entry decision. For every conjectured strategy, every candidate's strategy is a threshold strategy. First, consider the H candidate. For any conjectured strategy of the opponents, given the way we assume the voter breaks indifferences, her expected payoffs from running are uniquely defined as $(1 - \frac{1}{2}\frac{1}{2}p_H^c)(E + ((1-\tau)\phi^H + \tau)k)$. Furthermore, $(1 - \frac{1}{2}\frac{1}{2}p_H^c)(E + ((1-\tau)\phi^H + \tau)k) > 0$ and $(1 - \frac{1}{2}\frac{1}{2}p_H^c)(E + ((1-\tau)\phi^H + \tau)k) < W^H$, hence in every equilibrium there must exist a unique type of w^i, strictly between 0 and W^H, such that $w^i = (1 - \frac{1}{2}\frac{1}{2}p_H^c)(E + ((1-\tau)\phi^H + \tau)k)$. For every type above it, the unique best response is not to run. For every type below it, the unique best response is to run. By assumption, type $w^i = (1 - \frac{1}{2}\frac{1}{2}p_H^c)(E + ((1-\tau)\phi^H + \tau)k) := \bar{w}_H$ chooses to run.

- Consider now the L perspective candidate. She knows she will win only if H does not run. Furthermore, she wins with probability $\frac{1}{2}$ against a low education opponent. In every equilibrium, given the way we assume the voter breaks indifferences, this happens with probability $1 - \frac{1}{2}p_H^c - \frac{1}{2}\frac{1}{2}p_L^c \in (0,1)$, and this conjecture must be correct. Furthermore, in every equilibrium her expected payoff from being in office is uniquely defined as $(E+((1-\tau)\phi^L+\tau)k)$. Furthermore, $\left(1 - \frac{1}{2}p_H^c - \frac{1}{2}\frac{1}{2}p_L^c\right)(E+((1-\tau)\phi^L+\tau)k) > 0$ and $\left(1 - \frac{1}{2}p_H^c - \frac{1}{2}\frac{1}{2}p_L^c\right)(E+((1-\tau)\phi^L+\tau)k) < W^L$, hence in every equilibrium there must exists a unique type of w^i, strictly between 0 and W^L, such that $w^i = \left(1 - \frac{1}{2}p_H^c - \frac{1}{2}\frac{1}{2}p_L^c\right)(E+((1-\tau)\phi^L+\tau)k)$. For every type above it, the unique best response is not to run. For every type below it, the unique best response is to run. By assumption on the tie breaking rule, type $w^i = \left(1 - \frac{1}{2}p_H^c - \frac{1}{2}\frac{1}{2}p_L^c\right)(E+((1-\tau)\phi^L+\tau)k) := \bar{w}_L$ chooses to run.

- In every symmetric equilibrium, it must be that conjectures are correct and candidates with the same education level choose the same strategy. Hence, it must be that $p_H^c = \frac{\bar{w}_H}{W^H}$ and $p_L^c = \frac{\bar{w}_L}{W^L}$. As a consequence, $\gamma_H^i = 1 - \frac{1}{2}\frac{1}{2}\frac{\bar{w}_H}{W^H}$ and $\gamma_L^i = 1 - \frac{1}{2}\frac{1}{2}\frac{\bar{w}_L}{W^L} - \frac{1}{2}\frac{\bar{w}_H}{W^H}$. This leads to the system of equations described in lemma E3, whose solution is unique.

The same logic applies to the equilibrium in case of fiscal rules. ■

E.1.2 Fiscal rules

If fiscal rules are present, all politicians in office are constrained to choose $x = 0$. As a consequence,

Lemma E5. *When fiscal rules are in place, equilibrium entry thresholds are as follows:*

$$\bar{w}_H^{FR} = \frac{4W^H\left(E + (1-\tau)(1-p)k\right)}{4W^H + \left(E + (1-\tau)(1-p)k\right)}$$

$$\bar{w}_L^{FR} = \frac{4W^H - \left(E + (1-\tau)(1-p)k\right)}{4W^H + \left(E + (1-\tau)(1-p)k\right)}\frac{4W^L\left(E + (1-\tau)(1-p)k\right)}{4W^L + \left(E + (1-\tau)(1-p)k\right)}$$

Proof of Lemma E5. Given our assumption that, even in case of fiscal rules, the H candidate is chosen over an L candidate, the proof for this Lemma follows the same logic as the proof of Lemma E3. The sole difference is that now $\mathbb{E}_{b,\theta,s}u_H^P = \mathbb{E}_{b,\theta,s}u_L^P = (1-\tau)(1-p)$. The reason is that now both types of politicians, being constrained to play $x = 0$, derive utility only if $\theta = 0$ and they are unbiased. ■

E.2 Proof of the main proposition

Proof of Proposition 1. The proposition implies a comparison between $\hat{\lambda}$ and $\hat{\lambda}^{FR}$, defined using equation (5) and replacing the relevant p_H and p_L. We have:

$$\begin{aligned} &\hat{\lambda} > \hat{\lambda}^{FR} \\ &\frac{p_L^{FR}}{p_H^{FR}} > \frac{p_L}{p_H} \\ &\frac{\bar{w}_H}{\bar{w}_L} > \frac{\bar{w}_H^{FR}}{\bar{w}_L^{FR}} \end{aligned} \tag{E.4}$$

Substituting the relevant thresholds (7), (8), (9) and (10), and using (E.1), (E.2) and (E.3) to ease the notation, we have that:

$$\frac{\bar{w}_H}{\bar{w}_L} = \frac{4W^H h}{4W^H + h} \frac{4W^H + h}{4W^H - h} \frac{4W^L + l}{4W^L l} > \frac{4W^H f}{4W^H + f} \frac{4W^H + f}{4W^H - f} \frac{4W^L + f}{4W^L f} = \frac{\bar{w}_H^{FR}}{\bar{w}_L^{FR}} \tag{E.5}$$

$$\frac{(4W^L + l)h}{(4W^H - h)l} > \frac{(4W^L + f)}{(4W^H - f)}$$

Furthermore, the RHS of (E.5) is increasing in f and $f \leq l$ because $\phi^L >$

$max[p, 1-p]$. Hence, the RHS is below $\frac{(4W^L+l)}{(4W^H-l)}$. Note, however, that:

$$\frac{(4W^L + l)h}{(4W^H - h)l} > \frac{(4W^L + l)}{(4W^H - l)}$$
$$\frac{h}{(4W^H - h)l} > \frac{1}{(4W^H - l)}$$
$$(4W^H - l)h > (4W^H - h)l$$
$$h > l$$

that always holds because $\phi^H > \phi^L$. ■

Equation (E.5) is useful to capture the two channels through which fiscal rules act. Each side is composed by two elements whose comparison, individually taken, points toward $\hat{\lambda} > \hat{\lambda}^{FR}$. First, we have that $\frac{h}{l} > \frac{f}{f} = 1$, because $\phi^H > \phi^L$ and fiscal rules shut down the difference in expected payoffs from office between the two types of politicians. In words, the ratio between expected payoffs from being in office for H over L types is higher without fiscal rules, implying that their presence should discourage H types relatively more (note that fiscal rules reduce both h and l). Furthermore, the condition $\frac{h}{l} > \frac{f}{f}$ can be re-written as $\frac{h-f}{h} > \frac{l-f}{l}$. In other words, fiscal rules reduce $\hat{\lambda}$ through the first channel as long as the expected cost they imply for politicians in office, relative to their payoff from office without fiscal rules, is higher for H than for L types. In our model, this is always the case.

Second, we have that $\frac{(4W^L+l)}{(4W^H-h)} > \frac{(4W^L+f)}{(4W^H-f)}$, because $h > f$ and $l > f$. This part is a consequence of the strategic considerations of different types related with the running probability of the opponent. More in detail, fiscal rules unambiguously decrease the equilibrium p_H. However, a reduction in p_H is good news for L types, because they may win with higher chances.

Finally, we show that $\hat{\lambda}$ corresponds to the expected share of H candidates in any given municipality where at least one candidate runs.

Lemma E6. *The municipality-level expected share of H candidates conditional, on having at least one candidate, is* $\hat{\lambda} = \frac{p_H}{p_H + p_L}$.

Proof of Lemma E6. Define the expected share of H candidates conditional on having at least one candidate running as:

$$\hat{S} := \frac{[0.25 * 1 * (1-(1-p_H)^2) + 0.5 * 1 * p_H(1-p_L) + 0.5 * 0.5 * p_H p_L]}{0.25(1-(1-p_H)^2) + 0.25(1-(1-p_L)^2) + 0.5(1-(1-p_H)(1-p_L))} \tag{E.6}$$

To see that this is the expected share conditional on at least one perspective candidate running, note that at municipality level the share can be 1 with probability $\frac{1}{4}(1-(1-p_H)^2)+\frac{1}{2}p_H(1-p_L)$, i.e. when there are two H perspective candidates and at least one of them run or when there are one H and L candidate and only the H candidate runs. The share is 0.5 with probability $\frac{1}{2}p_H p_L$ (i.e. there are one H and L perspective candidate and both of them run), zero with probability $\frac{1}{4}(1-(1-p_L)^2)+\frac{1}{2}p_L(1-p_H)$ and undefined (define it as $S = \emptyset$) when no perspective candidate runs, i.e. with probability $\frac{1}{4}(1-p_H)^2+\frac{1}{4}(1-p_L)^2+\frac{1}{2}(1-p_H)(1-p_L)$. Then, the expected share conditional on $S \neq \emptyset$ is

$$\begin{aligned}\hat{S} = E(S|S \neq \emptyset) &= 1 * Pr(S = 1|S \neq \emptyset) + 0.5 * Pr(S = 0.5|S \neq \emptyset) \\ &= 1 * \frac{Pr(S = 1 \cap S \neq \emptyset)}{Pr(S \neq \emptyset)} + 0.5 * \frac{Pr(S = 0.5 \cap S \neq \emptyset)}{Pr(S \neq \emptyset)}\end{aligned}$$

Substituting the relevant probabilities, we obtain (E.6), where the denominator is the total probability of having at least one candidate running.

To complete the proof, note that:

$$
\begin{aligned}
\hat{S} &= \frac{0.25(1-(1-p_H)^2)+0.5p_H(1-p_L)+0.5*0.5p_Hp_L}{0.25(1-(1-p_H)^2)+0.25(1-(1-p_L)^2)+0.5(1-(1-p_H)(1-p_L))} \\
&= \frac{1-(1-p_H)^2+2p_H(1-0.5p_L)}{1-(1-p_H)^2+1-(1-p_L)^2+2-2(1-p_H)(1-p_L)} \\
&= \frac{1-(1-2p_H+p_H^2)+2p_H-p_Hp_L}{4-((1-p_H)+(1-p_L))^2} \\
&= \frac{4p_H-p_H^2-p_Hp_L}{4-(2-(p_H+p_L))^2} \\
&= \frac{p_H(4-p_H-p_L)}{4-4-(p_H+p_L)^2+4(p_H+p_L)} \\
&= \frac{p_H(4-p_H-p_L)}{(p_H+p_L)(4-(p_H+p_L))} \\
&= \frac{p_H}{p_H+p_L} := \hat{\lambda}
\end{aligned}
$$

■

E.3 Rigid municipalities

With respect to the baseline model, we add a second group of municipalities, those that are characterized by a high share of rigid expenditures, such as personnel and debt repayment expenditures, which cannot be adjusted in the short run. This implies that they cannot adjust their policy choice quickly. For simplicity, we model this as a constraint to keep the policy constant irrespective of the state of the world. We show that the introduction of fiscal rules is always expected to have a bigger effect on the probability that a candidate is an H type in non-rigid municipalities. Intuitively, the ability to get a better signal about the state of the world does not matter in case of rigidity and in case of fiscal rules. As long as choosing the right policy is valuable for motivating perspective candidates, the constrain imposed by fiscal rules has a stronger discouraging effect on highly educated perspective candidates in previously unconstrained municipalities.

Rigidity as $x = 0$. Assume that rigid municipalities are constrained to the policy $x = 0$ irrespective of the true state of the world, even in the absence of fiscal rules, as they cannot adapt their expenditures quickly when they should respond to negative shocks. Given the above, it is easy to see that, in rigid municipalities, the expected payoff conditional on being in office is the same for every education level, and it is $(E + (1 - \tau)(1 - p)k) = f$, irrespective of whether fiscal rule are in place or not.

Proposition E1. *When rigidity implies $x = 0$, the probability that a candidate is highly educated in rigid municipalities is the same with or without fiscal rules.*

Proof of Proposition E1. Define $\hat{\lambda}^{R0}$ the probability that a candidate is highly educated in those municipalities. Given the exogenous constraint to $x = 0$ irrespective of fiscal rules, we have $\bar{w}_H^{R0} = \bar{w}_H^{FR}$ and $\bar{w}_L^{R0} = \bar{w}_L^{FR}$, hence if we substitute in equation (5) we obtain $\hat{\lambda}^{R0} = \hat{\lambda}^{FR}$. ■

Rigidity as $x = 1$. Assume that those rigid municipalities are constrained to the policy $x = 1$ irrespective of the true state of the world. It is easy to see that, in those rigid municipalities, the expected payoff conditional on being in office is $E + k(\tau + (1 - \tau)p) := r1$. Equilibrium thresholds are the solution of the same system of equations as above, where h and l are both replaced by $r1$. We first show that the comparison between the probability that a candidate is an H type in those municipalities and in municipalities with fiscal rules is in general ambiguous. Second, we show that the probability that a candidate is an H type in non-rigid municipalities is always higher than in rigid municipalities, implying that any negative effect of fiscal rules is stronger in non-rigid municipalities. Define $\hat{\lambda}^{R1}$ the probability that a candidate is an H type in rigid municipalities.

Proposition E2. *When rigidity implies $x = 1$, the probability that a candidate is highly educated in rigid municipalities is higher than the probability that a candidate is highly educated with fiscal rules if $\tau > (1-\tau)(1-2p)$.*

Proof of Proposition E2. The proposition implies a comparison between $\hat{\lambda}^{R1}$ and $\hat{\lambda}^{FR}$. We have:

$$
\begin{aligned}
&\hat{\lambda}^{R1} > \hat{\lambda}^{FR} \\
&\frac{\bar{w}_H^{R1}}{\bar{w}_L^{R1}} > \frac{\bar{w}_H^{FR}}{\bar{w}_L^{FR}} \\
&\frac{4W^L + r1}{4W^{H-r1}} > \frac{4W^L + f}{4W^{H-f}} \\
&4W^H(r1 - f) > 4W^L(f - r1)
\end{aligned}
$$

The inequality is true iff $r1 > f$, i.e. iff $E+k(\tau+(1-\tau)p) > E+k(1-\tau)(1-p)$. This simplifies to $\tau > (1-\tau)(1-2p)$. ■

Proposition E3. *When rigidity implies $x = 1$, the probability that a candidate is highly educated in non-rigid municipalities is always higher than the probability that a candidate is highly educated in rigid municipalities.*

Proof of Proposition E3. The proposition implies a comparison between $\hat{\lambda}$ and $\hat{\lambda}^{R1}$. We have:

$$\hat{\lambda} > \hat{\lambda}^{R1}$$

$$\frac{\bar{w}_H}{\bar{w}_L} > \frac{\bar{w}_H^{FR}}{\bar{w}_L^{FR}}$$

$$\frac{h(4W^L + l)}{l(4W^H - h)} > \frac{4W^L + r1}{4W^H - r1}$$

Note that the RHS is increasing in $r1$, and that $r1 < l$. Hence, the LHS is higher than the upper bound of the LHS. To see this, note that

$$\frac{h(4W^L + l)}{l(4W^H - h)} > \frac{4W^L + l}{4W^H - l}$$

$$\frac{h}{l(4W^H - h)} > \frac{1}{4W^H - l}$$

$$h(4W^H - l) > l(4W^H - h)$$

$$h > l$$

■

E.4 Discussion on the theoretical framework

In this appendix we further discuss some of the assumptions and implications of the model.

E.4.1 The necessity of policy-motivated politicians

This section shows that some degree of policy motivation (irrespective of its direction) is necessary for our result.

Corollary E1. *If $k = 0$, the probability that a candidate is highly educated is the same with and without fiscal rules.*

Proof of Corollary E1. Substituting $k = 0$ in the LHS of equation (E.4) and on the relevant equations of Lemma E3 and E5, we obtain

$$h := E$$

$$l := E$$

$$f := E$$

Hence, $\hat{\lambda} = \hat{\lambda}^{FR}$. ■

Intuitively, when $k = 0$, fiscal rules have no effect on the incentives of H politicians: they get E for being in office irrespective of the policy they choose. Hence, their probability of running is the same, and nothing changes for L politicians as well.

On the other hand, the observed effect of fiscal rules holds if politicians are purely policy motivated and if the winning probability does not enter in their decision. In particular:

Corollary E2. *Assume that $E = 0$ and $k > 0$. In this case, $\hat{\lambda} > \hat{\lambda}^{FR}$.*

Proof of Corollary E2. Substituting $E = 0$ in (E.4) and on the relevant equations of Lemma E3 and E5, it is still true that $h := k(\tau + (1 - \tau)\phi^H) > l := k(\tau + (1 - \tau)\phi^L)$. Therefore, we can follow the same steps as in the proof of proposition 1 and conclude that $\hat{\lambda} > \hat{\lambda}^{FR}$. ■

We also show that the effect of fiscal rules, driven by $\phi^H > \phi^L$, survives even if we assume that candidates keep their salary if they lose, so effectively they do not take into account γ^i_Γ in their decision.

Corollary E3. *Assume that politicians receive w^i, instead of 0, when they run and lose. In this case, $\hat{\lambda} > \hat{\lambda}^{FR}$.*

Proof of Corollary E3. With this assumption, equations (1) and (3) become $\gamma^i \left(E + k\mathbb{E}_{b,\theta,s} u_H^P\right) + (1-\gamma^i)w^i$ and $\gamma^i \left(E + k\mathbb{E}_{b,\theta,s} u_L^P\right) + (1-\gamma^i)w^i$ respectively. This means that $\bar{w}_\Gamma$ does not depend on γ anymore. Therefore, it is straightforward to see that in this case $\bar{w}_H = E + ((1-\tau)\phi^H + \tau)k$, $\bar{w}_L = E + ((1-\tau)\phi^L + \tau)k$ and $\bar{w}_H^{FR} = \bar{w}_L^{FR} = E + (1-\tau)(1-p)$. Replacing in equation (E.4), we obtain

$$\frac{\left(E + ((1-\tau)\phi^H + \tau)k\right)}{\left(E + ((1-\tau)\phi^L + \tau)k\right)} > 1$$

Hence the result holds. ■

E.4.2 Education and bias

Suppose bias is correlated with education, i.e. we have τ_H and τ_L. We show that it is always possible to find a range of values in τ_H, τ_L where the main result of the paper holds. We keep assuming that H politicians are preferred by V.[1]

Proposition E4. *Assume $\tau_H \neq \tau_L$. For every τ_H, it is always possible to find a range of values of τ_L such that $\hat{\lambda} > \hat{\lambda}^{FR}$.*

[1] This translates into the assumption that $(1-\tau_H)\phi^H + \tau_H p > (1-\tau_L)\phi^L + \tau_L p$, i.e. $\tau_L > \tau_H \frac{\phi^H - p}{\phi^L - p} - \frac{\phi^H - \phi^L}{\phi^L - p}$.

Proof of Proposition E4. To ease the notation, we define the following:

$$\begin{aligned} h_\tau &:= E + k((1-\tau_H)\phi^H + \tau_H) \\ l_\tau &:= E + k((1-\tau_L)\phi^L + \tau_L) \\ f_H &= E + k(1-\tau_H)(1-p) \\ f_L &= E + k(1-\tau_L)(1-p) \end{aligned}$$

Using (E.4), but noticing that we cannot simplify the RHS as before, we have that $\hat{\lambda} > \hat{\lambda}^{FR}$ iff

$$\frac{h_\tau(4W^L + l_\tau)}{l_\tau(4W^H - h_\tau)} > \frac{f_H(4W^L + f_L)}{f_L(4W^H - f_H)}$$
$$h_\tau(4W^L + l_\tau)f_L(4W^H - f_H) > f_H(4W^L + f_L)l_\tau(4W^H - h_\tau)$$

Note that $h > f_H$ and $l > f_L$, therefore a sufficient condition for the inequality to hold is

$$h_\tau f_L > f_H l_\tau \tag{E.7}$$

We now show that, for every τ_H and every combination of parameters, there exists a set of values of τ_L where (E.7) holds. First, note that if $\tau_H > \tau_L$ then $h_\tau > l_\tau$ and $f_L > f_H$, therefore the inequality is always satisfied. Consider now the case of $\tau_H < \tau_L$. Noticing that the LHS of (E.7) is increasing in τ_H and the RHS is decreasing in τ_H, we set τ_H to zero and look for a condition on τ_L such that the inequality holds. Higher τ_H are only going to relax this condition. Substituting $\tau_H = 0$ in (E.7) and using the definitions outlined above, we have that the inequality holds iff

$$
\begin{aligned}
&(E + k\phi^H)(E + k(1-\tau_L)(1-p)) > (E + k(1-p))(E + k(\phi^L + (1-\phi^L)\tau_L)) \\
&(E + k(1-p))k(\phi^H - \phi^L) > \tau_L \left[(E + k(1-p))k(1-\phi^L) + (E + k\phi^H)(1-p)k\right] \\
&\tau_L < \frac{(E + k(1-p))(\phi^H - \phi^L)}{E(2-\phi^L - p) + k(1-p)(1+\phi^H - \phi^L)} := \bar{\tau}_L
\end{aligned}
$$

Note that $\bar{\tau}_L$ is strictly positive for every combination of parameters, therefore we have a non-empty set of values of τ_L such that the main result of the paper holds. ■

References cited only in the Online Appendix

Calderoni F. (2011). "Where is the mafia in Italy? Measuring the presence of the mafia across Italian provinces." *Global Crime*.

Calonico, S., Cattaneo, M.D., Farrell, M. H. and Titiunik R. (2017). "rdrobust: Software for Regression Discontinuity Designs." *The Stata Journal* 17(2): 372–404.

Cattaneo, M. D., Jansson, M., and Ma, X. (2018). "Manipulation Testing based on Density Discontinuity." *Stata Journal*, 18(1): 234-261.

Cattaneo, M. D., Idrobo, N. and Titiunik R. (2020). "A Practical Introduction to Regression Discontinuity Designs: Foundations." Elements in Quantitative and Computational Methods for the Social Sciences. Cambridge: Cambridge University Press.

Lockwood, B., Porcelli, F., Redoano, M., Bracco, E., Liberini, F., and Sgroi, D. (2021). "The Effects of Social Capital on Government Performance and Turnover: Theory and Evidence from Italian Municipalities." Working paper.