WARWICK

ECONOMICS

QAPEC

Quantitative and Analytical Political Economy Research Centre

Discussion Paper Series

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

Matteo Gamalerio & Federico Trombetta

Date December 2021

No: 10

QAPEC

Quantitative and Analytical Political Economy Research Centre

Department of Economics
University of Warwick, Coventry,
CV4 7AL, United Kingdom

https://warwick.ac.uk/fac/soc/economics/research/centres/qapec/discussionpapers

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

Matteo Gamalerio† Federico Trombetta‡

December 13, 2021

Abstract

Fiscal rules, i.e., constraints to the policymaking discretion of elected officials, are widely used tools to regulate fiscal policies. We build a simple model that combines fiscal rules with endogenous entry into politics, showing how fiscal rules can negatively affect the quality (in terms of education) of candidates running for office. The mechanism behind this effect is due to the differential ability in choosing the correct policy between high and low education politicians and the fact that fiscal rules constraint politicians' actions, creating a relatively higher entry cost for high education politicians. Consistent with the model, the empirical analysis developed with data from Italian municipalities and a difference-in-discontinuity design shows that fiscal rules negatively affect candidates' education. We also show that municipalities where fiscal rules meaningfully restrict the action space of politicians (i.e., those not ex-ante financially constrained) drive the effect. In addition, we provide evidence that high education politicians are more likely to choose the correct policy when fiscal rules do not apply. These results highlight a new "general equilibrium" effect of fiscal rules. Reducing policymaking discretion may alleviate inter-jurisdictional externalities and pork-barrel spending. However, it may also lower the quality of the political class.

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

JEL Classification: D72, H62, H70, H72.

*We would like to thank Mohit Agrawal, Matteo Alpino, Pablo Beker, Dan Bernhardt, Christopher Berry, Jeremy Boccanfuso, Massimo Bordignon, Fernanda Brollo, Estelle Cantillon, 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, Marta Reynal-Querol, 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 seminar at 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, the seminar at Università Cattolica, the XV GRASS workshop and the 5th Bruneck-Brunico workshop on Political Economy for insightful comments. Special thanks to Mario Luca for helping with the collection of the corruption data. All mistakes are ours. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

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

‡DISEIS, Università Cattolica del Sacro Cuore, Milano, federico.trombetta@unicatt.it

1 Introduction

This paper shows an important side effect of a reduction in policymaking discretion due to fiscal rules: it negatively affects the quality of candidates willing to run for political offices and elected politicians.

Fiscal rules are constraints on fiscal policies widely adopted worldwide,[1] introduced to reduce the incentives of national and local governments to accumulate debt and run deficits. Both the theoretical (Battaglini and Coate, 2008; Halac and Yared, 2014; Azzimonti, Battaglini, and Coate, 2016; Halac and Yared, 2018; Halac and Yared, 2019) and the empirical (Grembi et al., 2016; Daniele and Giommoni, 2020) literature describe fiscal rules as a reduction in policymaking discretion, with a trade-off behind their application between commitment and discretion. On one side, fiscal rules provide commitment, limiting the incentives toward excessive spending. On the other side, there is a cost in terms of reduced flexibility and discretion, given that fiscal rules limit governments' ability to respond to shocks. The existing literature on fiscal rules has mainly studied their effect on fiscal stability and economic outcomes (Alesina and Perotti, 1996; Wyplosz, 2012; Grembi et al., 2016). 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 policymaking discretion and, potentially, in the value of holding office, can lower the quality of the political class.

Our main contribution is empirical. Specifically, we use data on Italian municipalities from 1993 to 2012 to estimate the effect of fiscal rules on the quality of the political class, measured as politicians' education level.[2] Italy is an interesting framework for the research question investigated in this paper. In 1999, the Italian government introduced fiscal rules to limit the incentives to accumulate debt and run deficits. These rules initially applied to all municipalities and were introduced under the name of the "Domestic Stability Pact" (DSP). In 2001, the central government removed the rules for all towns with less than 5000 inhabitants. This relaxation remained in place until 2013 when the cutoff changed from 5000

[1] As Grembi et al. (2016) reported, many countries have adopted rules to constrain local governments' fiscal policies in recent years. These include 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] As described in section 2, education is an indicator often used in the political selection literature (Galasso and Nannicini, 2011; Gagliarducci and Nannicini, 2013).

to 1000 inhabitants.

This institutional framework would be ideal for a Regression Discontinuity Design if fiscal rules were the only policy change at the 5000 inhabitants threshold. However, at the same cutoff, there is a sharp increase in the wage paid to the mayor and the municipal ministers, based on a policy introduced by the Italian government in the 1960s (Gagliarducci and Nannicini, 2013; Grembi et al., 2016). This policy represents a confounding factor, as Gagliarducci and Nannicini (2013), using data on Italian municipalities between 1993 and 2001, have shown that this wage increase attracts more educated individuals into politics. Hence, we exploit the 2001 removal of fiscal rules for municipalities below 5000 inhabitants to estimate a Difference-in-Discontinuity (*Diff-in-Disc*) model, which allows estimation of the effect of fiscal rules on political selection separately from that of the wage increase (Grembi et al., 2016). The main results of the empirical analysis show that fiscal rules negatively affect the level of education of politicians.[3] More in detail, fiscal rules induce a 10 % points reduction in the share of graduate mayoral candidates. At the same time, we find that fiscal rules reduce the probability of electing a graduate mayor, with a reduction similar in magnitude to the one observed for mayoral candidates.[4]

To rationalize this empirical evidence and shed light on the mechanisms behind it, we propose a simple model of fiscal rules and political selection. Building on the existing theoretical literature on political selection (see Dal Bo et al. 2018 for a recent review), our model adds two elements. First, we consider the role of fiscal rules, modeled as a constraint to the action space of politicians. Second, politicians' type is multidimensional, incorporating both their education (observable) and their pro-deficit bias (unobservable).

In a nutshell, we incorporate explicitly fiscal rules (henceforth, FR) and their well-known flexibility-commitment trade-off (e.g., Halac and Yared 2014). FR avoid the improper use of public money from biased politicians, but they also make it more difficult to choose the correct fiscal policy. This impacts the value of holding office for different types of perspective mayors if they are policy-motivated. We assume that high education unbiased politicians can better understand the economy's state correctly. This different ability is the main driving force of the model because it causes FR to have a more negative impact on high education than on low education unbiased politicians. In other words, tied hands are costlier for high

[3]The analysis below also studies how fiscal rules affect other politicians' characteristics.

[4]In an additional exercise, we extend the data to 2015 and exploit the 2013 variation in the application of fiscal rules. The results of this exercise go in the same direction and confirm the evidence about the negative effect of fiscal rules on politicians' quality.

education politicians, given their higher likelihood of choosing the correct policy. This first, direct effect, which discourages high education politicians from running for office, has the additional effect of making the run more profitable for low education politicians, increasing their chances of winning. Because of those channels, the overall effect of fiscal rules on the education level of the pool of candidates is predicted to be negative.

Together with the main prediction, the model suggests that one important channel through which FR affect political selection is heterogeneity in ability, mirrored by education, combined with some degree of policy motivation of politicians and the fact that FR constitute a restriction to what can be achieved. We cannot test it directly, but we can bring two important elements of this mechanism to the data. First, the mechanism requires FR to be a meaningful restriction to what can be done once in office. Hence, we should observe no effect if the action space is already *de facto* restricted. For example, if a municipality is highly financially constrained because of a pre-existing high debt. Second, the mechanism requires high education politicians to be better, on average, in understanding the state of the economy and hence in choosing the correct policy. Consequently, there has to be a difference between them and low education politicians in their probability of choosing the correct policy when fiscal rules are not in place. However, this difference should disappear when fiscal rules are in place.

The empirical analysis confirms both those elements. More in detail, we build a proxy for the degree of pre-existing financial constraints faced by a municipality to show that, consistent with the model, places with lower pre-treatment deficit drive the negative effect of fiscal rules on politicians' education. Conversely, the negative effect disappears in the sample of municipalities where a more binding financial constraint is in place. Furthermore, we can show that, without fiscal rules, high education politicians are more likely to choose the correct policy, given the state of the economy, *vis-à-vis* low education politicians. This is a direct consequence of how we model fiscal rules: they are assumed to constrain politicians' actions. Hence, matching the state should no longer be correlated with education once fiscal rules are in place. We test this implication using a regression discontinuity design based on mixed electoral competitions between graduate and non-graduate mayors. We show that graduate mayors are more likely to match the state of the economy in municipalities where fiscal rules do not apply. Finally, we show that alternative stories do not explain our results

and demonstrate how the results survive a series of robustness checks.[5]

2 Related literature

This paper aims to contribute to three strands of literature. The first is the political economy literature on the selection of politicians (Besley, 2005; Braendle, 2016; Dal Bo and Finan, 2018). Our paper contributes to this literature from both a theoretical and empirical point of view. On the theory side,[6] 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 is similar to 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 those models consider the theoretical implications of fiscal rules on political selection, modelling them as a flexibility-commitment trade-off consistent with Halac and Yared (2014). Similar to us, Le Borgne and Lockwood (2002) and Izzo (2020) consider the effect of economic-related dimensions on political selection. However, their focus is on political budget cycles and economic crises, respectively.

The empirical literature on political selection has analyzed many different institutions that can affect the selection of politicians.[7] As far as we know, none of the papers in the literature has studied the role of fiscal rules. We contribute to this literature by showing how a reduction in policymaking discretion and the value of holding office due to the application

[5]As described below, we deal with various alternative stories. First, we show that the results are not due to a reduction in real terms of the difference in wages paid to mayors across the threshold. Second, we show that different out-of-politics options for individuals with different levels of education are not sufficient to explain our empirical results. Third, we provide evidence that the effect of fiscal rules on the education of politicians is not an indirect consequence of the potential effect on other variables like political experience, political orientation, and re-election probability. If anything, we show that fiscal rules do not affect these variables. 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).

[6]See among the various papers: 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.

[7]Among the institutions and determinants studied, we find 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), electoral rules (Beath et al., 2015, De Benedetto, 2018), 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).

of fiscal rules can negatively affect the quality of the political class. Within this literature, our results relate to the ones by Gagliarducci and Nannicini (2013). They use data from Italian municipalities around the 5000 inhabitants threshold from 1993-2001 to show how a higher wage 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 this positive selection due to the higher wage paid to local politicians. These results suggest that, while paying politicians high wages may be a good idea, as these attract skilled individuals, competent persons may decide to enter politics for many different reasons.

The second literature is the one that analyzes the effect of fiscal rules on fiscal stability and economic and political outcomes in general. The evidence on fiscal rules' effect on fiscal stability is not definitive, as some studies (Alesina and Perotti, 1996, and Wyplosz, 2012) indicate that fiscal rules may not work for reasons of commitment. From this perspective, the most recent advancement in the literature is Grembi et al. (2016). Using data on Italian municipalities, they have shown that fiscal rules can effectively reduce the deficit run by local governments. On this line of research, Vannutelli (2021) shows that fiscal rules are even more effective when their enforcement is associated with independent auditors. Among the studies on the effect of fiscal rules on other economic outcomes, we find Daniele and Giommoni (2020), who show that fiscal rules reduce corruption. Besides, Coviello et al. (2020) study the impact of fiscal rules on public infrastructure expenditures and the size of firms, while Revelli (2016) studies the impact of tax limits on electoral turnout and local election outcomes. Alpino et al. (2020) study the effect of fiscal rules on distributional policies. Carreri and Martinez (2021) also look at political outcomes of fiscal rules. However they focus on a different type of fiscal rule ("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 this literature by investigating an unexplored consequence of the application of fiscal rules on the quality of the political class.

Third, and more broadly, this paper contributes to the literature on incentives and selection into public sector roles[8] and beyond,[9] with a focus on intrinsic incentives. Deserrano (2018) shows that higher wages may deter prosocial candidates from applying for a government job; Ashraf et al. (2019) find similar results with respect to career perspectives. Bartling et al. (2012) show experimentally the complementarity between the possibility of

[8]See Finan et al. (2015) for a review focused on field experiments.
[9]See Oyer and Schaefer (2010) for a review.

screening employees' past performance and the amount of on-the-job discretion allowed by employers. The quasi-experimental setting of this paper allows us to document the direct effect of a reduction in agents' discretion on the quality of "applicants" for a public office, holding financial returns fixed. We show that more discretionary power attracts better candidates, at least for this type of "executive" jobs. In this respect, the relevance of intrinsic incentives we highlight in this paper is consistent with the results of Gulzar and Khan (2021). They find that increasing the salience of prosocial incentives motivates prosocial people to run for office and behave in a way more aligned with citizens' preferences.

Finally, to empirically measure the quality and competence of politicians, we use the level of education of mayoral candidates and elected mayors. The focus on education is justified by the literature, which shows that education positively affects socio-economic outcomes such as wages (Card, 1997) and measures of citizenship (Dee, 2004). More importantly for this paper, the literature has provided evidence that electing more educated political leaders can positively affect economic growth (Besley et al., 2011) and the production of public goods (Martinez-Bravo, 2017). Besides, education is also an indicator extensively used in the political selection literature (e.g., Gagliarducci and Nannicini, 2013; Galasso and Nannicini, 2011), which positively correlates with measures of administrative competence (Carreri, 2020). There are two justifications for the focus on elected mayors and mayoral candidates. First, as described in section 4.1, Italian mayors are powerful at the municipal level. Second, the seminal paper on political selection by Besley (2005) suggests that the scope of authority enjoyed by elected politicians should affect the selection of directly elected chief executives such as presidents, governors, and mayors rather 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 on politicians in less prominent positions, like municipal councilors.

3 Theoretical model

3.1 Model set up

There is a large number n of municipalities. Each of them 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.

3.1.1 Politicians' types and payoffs

Politicians' education level is denoted by $\Gamma \in \{H, L\}$ and it is observable. We assume for simplicity that in each municipality party 1 is able to find a high education candidate and party 2 finds a low education candidate.[10]

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.[11]

When in office, politicians derive an office rent $E > 0$ capturing the direct office motivation (e.g. salary) and a policy related utility, weighted by $k > 0$. More in detail, some politicians are biased in favour of spending, i.e. they receive a payoff of 1 when they choose $x = 1$. We denote the bias $b \in \{0, 1\}$, with common prior $Pr(b = 1) = \tau \in (0, 1)$. We assume that the bias is uncorrelated with education. Moreover, politicians learn their bias once in office. This assumption simplifies the game, as it avoids to transform the entry decision in a signalling game. This idea of bias toward deficit spending captures many things: politicians may be just present-biased, as in Halac and Yared (2014) or in Piguillem and Riboni (2015); they may enjoy public spending more when they are in office, and they take into account the fact that they may not be in office tomorrow; they may be dishonest and use those money for private purposes; they may be willing (and able) to manipulate voters' inference on their competence (as in Murtinu et al. 2021). The model is sufficiently general to capture all those different motivations.[12] The remaining fraction of politicians is unbiased, meaning that they want to choose the correct policy. This may be intrinsic motivations, re-election incentives, ways to enhance their career perspective. Once again, we are agnostic on the precise motivation. Finally, note that τ can be interpreted also 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 and so on).

[10] This assumption is similar to Izzo (2020).

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

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

3.1.2 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. 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. 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 high education politician has a better understanding of the state of the world than the low education one. This is a way to capture the advantage of high ability politicians once in office, consistently with Caselli and Morelli (2004).

3.1.3 Running decision

In each municipality, there are two political parties. One member for each party is selected to become a candidate. For simplicity, we assume that party 1 selects a high education member and party 2 selects a low education one,[13] and this is known to the players.[14] Selected candidates simultaneously chooses whether to accept and run, as in Brollo et al. (2014) and Dal Bo and Finan (2018). If none of the two randomly picked party members run, a default mayor is in place.[15]

If the chosen politician decides not to run, she keeps her salary (w^i, private information of 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 big to ensure an interior solution. 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 $d^i \in \{r, nr\}$ the decision on whether to run or not from politician i, selected to run. γ^i_Γ is his probability of winning the elections (in equilibrium, this is endogenous as it depends on the decision of the opponent and its education level). As a

[13] Equivalently, we can assume that party 2 does not have high education candidates in the municipality

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

[15] This particular timing is identical to Dal Bo and Finan (2018). However, results are qualitatively unchanged if we assume that first potential politicians self-select into the pool of potential candidates and then parties pick one of those potential candidates, who runs for sure.

consequence, if i is a high education type who runs for office, we have the following objective function:

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

The term in parenthesis is the payoff in case of victory: she gets the ego rent E and the expectation on 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. If instead she chooses not to run once selected, the politician receives

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

In case of a low education potential politician 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 the different informativeness of the signal. In equilibrium, all unbiased politicians follow it, but H politicians are more likely to choose the correct policy.

3.1.4 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, among others, by Halac and Yared (2014). 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. the consonant politician 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, in terms of election choices, voters are indifferent between high and low education politicians when fiscal rules are in place. The way we break this indifference does not matter for our result. However, in order to capture the fact that education can be correlated with in-office performance on other issues that are different from the budget policy, we assume that, if indifferent, the voter chooses the high education candidate.

3.1.5 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;
2. If there are two candidates, voters vote sincerely. If there is only one candidate, he wins directly. If there are no candidates, a default mayor is in place;
3. The winning politician privately learns b. Then he observes s and chooses x.
4. u^V is realized, payoffs are paid and the game ends;

The game is one shot and our solution concept is Perfect Bayesian Nash Equilibrium (PBNE). As a tie-breaking rule, we assume that indifferent politicians choose to run.

3.1.6 "Outcome variable"

We are interested in the effects of fiscal rules on ex ante selection into politics. One way to measure it is to look at the probability of a candidate being a high education type, 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) the probability that a randomly picked high (low) education member of a party chooses to run.

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

3.2 Analysis

3.2.1 No fiscal rules

We solve the game by backward induction. 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 as a consequence they pick the correct policy with probability ϕ^Γ.

Lemma 1. *Without fiscal rules, there is a unique PBNE whose policy choice is as follows*

- *Biased politicians always choose $x = 1$;*

- *Unbiased politicians choose $x = s$.*

All proofs are in the relevant appendix. The voter anticipates the equilibrium choices described above. As high education unbiased politicians behave in a better way, in expectation, V prefers to elect the candidate with $\Gamma = H$, when the election is contested.

Lemma 2. *If both candidates run, $\gamma_H^i = 1$.*

When choosing whether to enter into politics or not, politicians 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 behaviour described in lemma 1. This implies that, in the unique PBNE,

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

In either case, the decision is captured by a threshold in w^i.

Lemma 3. *In the unique PBNE equilibrium without fiscal rules the entry threshold of politicians is defined by*

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

Lemma 3 is straightforward. H politicians know that, if they run, they will surely win. On the opposite side, L politicians know that they will win only if the H candidate does not run. The rest is their expected payoff from being in office. As a consequence, in equilibrium, $p_H = \frac{\bar{w}_H}{W^H}$ and $p_L = \frac{\bar{w}_L}{W^L}$.

3.2.2 Fiscal rules

In case of fiscal rules, all politicians in office are constrained to choose $x = 0$. As a consequence,

Lemma 4. *When fiscal rules are in place, equilibrium entry thresholds are as follows:*

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

3.2.3 Comparison

We are now in the position to compare the share of high education candidates with and without fiscal rules. Proposition 1 summarizes our findings:

Proposition 1. *The share of high education politicians among those willing to run for office is higher without fiscal rules.*

Proposition 1 (and its proof) contains several insights. First note that we can focus just on the ratio between $\bar{w}_H$ and $\bar{w}_L$, with or without fiscal rules, for our comparison.

Second, fiscal rules have two distinct effects, both pushing toward a decrease in the share of H candidates. The first effect is a reduction in the expected policy payoff from office for unbiased politicians, but this reduction is larger for H candidates. To see this, note that without FR they are able to match the state of the world with probability $\phi^H > \phi^L$. With FR, this happens with probability $1 - p$ irrespective of the education level. Hence, the reduction is relatively larger for H politicians. This fact has a second effect, because it implies that H politicians are less likely to run, with FR. This in turn increases the incentive of L politicians to run, as they can win when they run unopposed.

Third, the model allows us to stress that the first effect, going through different ability and policy-related incentives, is crucial. As shown in Appendix D, if we shut it down (i.e. if we assume purely office motivated politicians), the effect of fiscal rules disappears. Conversely, the effect remains if we shut down the change in probability of re-election.

3.2.4 Additional testable implications

Together with its main result, the model suggests an important mechanism[16] for the observed effect. This relies on FR meaningfully restricting the set of available policies, combined with a differential in ability between high and low education politicians, once in office. Both those implications can be brought to the data.

Financially constrained municipalities Consider the case of a municipality that is "financially constrained", i.e. such that high levels of pre-existing debt force the incumbent mayor to balance the budget, irrespective of whether fiscal rules are in place. In those cases, the action space is effectively constrained to $x = 0$ irrespective of fiscal rules. In those places,

[16] In the sense of Ashworth et al. (2021).

FR are not meaningfully restricting the action space of politicians in office. Hence, if the channel implied by the model is driving the behaviour of politicians, we should observe no effect of FR on the composition of the pool of candidates. Appendix C discusses this case formally.

Education level and policy choice The main driving force of the model lies in the fact that high education politicians have a better ability in matching the state of the world, are (sometimes) willing to use it and may be restricted in doing so by FR. This implies that, without fiscal rules, we should observe better policy choices from high education politicians, on average. But, once fiscal rules are in place, this difference should disappear.

3.3 Summary

In a nutshell, the model makes one main prediction: fiscal rules decrease the share of high education candidates willing to run for office. Furthermore, there are two testable implications of the suggested mechanism:

1. The effect of fiscal rules on the quality of candidates should be visible where they meaningfully restrict the action space of politicians in office, hence it should disappear in financially constrained municipalities;

2. High education politicians are more likely to choose the correct policy than low education ones, without fiscal rules. There is no longer a difference when fiscal rules are introduced.

4 Institutional Setting

4.1 Italian municipalities

In Italy, there are 8047 municipalities, of which 70.5 % have less than 5000 inhabitants. Municipalities are responsible for municipal police, infrastructure, transport, welfare, housing, garbage collection, and water supply. They manage 10 % of total public expenditures, and around 20 % of their revenues come from local taxes. The rest of the revenues are discretionary transfers from higher levels of government, like provinces, regions, and the central state. Among local taxes, the most important taxes are the 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 in 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 used above the same threshold. Mayors are elected for five years and, since 1993, a maximum of two consecutive terms, i.e., they face a two-term limit. Within the municipal government, mayors are powerful, as they can choose and dismiss the ministers that form part of the municipal government. Besides that, if the municipal council wants to dismiss the mayor, new elections must be held.

4.2 The "Domestic Stability Pact" (DSP)

Italy introduced fiscal rules for municipal governments in 1999, following the European Stability and Growth Pact (SGP), signed in 1997 by different European countries. In Italy, these rules were called the "Domestic Stability Pact" (DSP) [17]. The DSP goal was to reduce the incentives for local governments to accumulate debt and run deficits. Table A1 describes the temporal evolution of the target imposed by the DSP for the years 1999-2015. The target has not been constant over time, even though, excluding the years 2005-2006, it has been balancing local governments' budgets for all years.[18] The initial penalties introduced by the central government for not complying with the rules were a 5 % cut in grants transferred by the national government, a cut on reimbursement and non-absenteeism bonuses for municipal employees, and a ban on new municipal hires. Municipalities complying with the rules were rewarded with a cut in the expenses for interests on loans received by the central government. In 2008, as described by Coviello et al. (2020), the penalties for not complying with the rules have been made harsher. The new penalties introduced in 2008 included a reinforced cut in central government grants and an automatic 30 % cut in the wage paid to mayors and municipal councilors.

As we can see from Table A1, fiscal rules applied to all municipalities in the first two years (1999-2000). In 2001, the central government removed the fiscal rules for all the municipalities below 5000 inhabitants to lift onerous constraints on municipalities disadvantaged

[17] Domestic Stability Pact stays for the Italian *Patto Interno di Stabilita'*. The Law that introduced the DSP in Italy is the number 448, 23 December 1998, article 28.

[18] The definition of budget balance used as a target for most of the years has been the so-called fiscal gap (Grembi et al., 2016). The fiscal gap is the municipal deficit net of transfers and debt service. The limits on the target have been changing over time. In some years, municipalities were asked to apply a cap to the growth of the target. In other years, municipal governments were asked to cut the target. The limits imposed on the target have always been calculated with reference to past values of the target in specific reference years.

by economies of scale. Specifically, following the law (Legislative Decree n. 267, article 156), the population used to identify the municipalities subject to fiscal rules is the resident population measured by the Italian Statistical Office (Istat) at the end of the previous penultimate year (e.g., for the year 2002, the population of reference was the one on the 31st of December 2000). In 2002, Regions with Special Statute (i.e., Sardegna, Sicilia, Valle d'Aosta, Trentino-Alto Adige, Friuli-Venezia Giulia) were allowed to establish their own fiscal rules. For this reason, we exclude from the analysis the municipalities in these regions. The 5000 inhabitants threshold remained in place until 2013. In that year, the threshold was reduced from 5000 to 1000 inhabitants for 2013-2015 (Daniele and Giommoni, 2020). Finally, the DSP was abolished in 2016 and replaced by a new set of balanced budget rules for all municipalities.

Finally, Table A2 reports the legislative population thresholds applicable to municipalities with less than 15,000 inhabitants. As we can see, the wages paid to the mayor and the municipal ministers change at the 5000 inhabitants threshold (Gagliarducci and Nannicini, 2013). This wage increase is a policy that dates back to the 1960s (Gagliarducci and Nannicini, 2013), and it has remained constant in real terms until today. The population used to identify the wage paid to the mayor is the one measured in the last available population Census.

5 Empirical Evidence

5.1 Empirical Strategy

5.1.1 Difference-in-discontinuity

We test the main prediction of the theoretical model, together with the first implication, exploiting the variation over time in the application of fiscal rules around the 5000 inhabitants threshold. In the absence of other policies changing across the threshold, this institutional setup would be appropriate for a regression discontinuity design (RDD) approach applied to the electoral terms between 2001 and 2012. However, the higher wage paid to the mayor in municipalities above the threshold represents a confounding policy that would invalidate the RDD approach, as it would not be possible to disentangle its effect from that of fiscal rules. Hence, a standard RDD approach is not appropriate in this context. However, as described by Grembi et al. (2016), the removal of fiscal rules in 2001 for municipalities below 5000

can be exploited to implement a Difference-in-Discontinuity (*Diff-in-Disc*) approach, which allows estimation of the effect of fiscal rules separately from that of the wage increase.[19]

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) which combines the *pre/post treatment* variation typical of a Difference-in-Differences design with a *just below/just above a threshold* variation that characterizes an RDD approach. In the Italian context, this strategy enables estimation of the effect of fiscal rules on the selection of politicians while controlling for the wage increase, which is constant in real terms over time. Besides, we can also extend this approach to the electoral years after 2012, exploiting the 2013 reduction of the fiscal rules application threshold to 1000 inhabitants. This extension allows us to evaluate whether the effect of fiscal rules disappears across the 5000 inhabitants threshold once fiscal rules are applied again.

More in detail, we estimate the following empirical model, using data at the municipality and electoral year level:

$$Y_{it} = \rho_0 + \rho_1 R_{it} + (> 5000_{it}) * (\beta_0 + \beta_1 R_{it}) + (Post_t) * [\pi_0 + \pi_1 R_{it} + (> 5000_{it}) * (\phi_0 + \phi_1 R_{it})] + \eta_{it} \tag{7}$$

where Y_{it} is the level of education of politicians. The variable $R_{it} = P_{it-1} - P_{5000}$ is the normalized population which measures the distance of municipality i from the 5000 inhabitants threshold P_{5000} at time t. As described in section 4.2, the population P_{it-1} is the resident population measured at the end of the previous penultimate year (e.g., for the year 2002, the population of reference was the one on the 31st of December 2000). In practice, P_{it-1} is the resident population measured at the beginning of the previous year, which explains the subscript (t-1).[20] The dummy variable $(> 5000_{it})$ is 1 if municipality i is above the 5000 inhabitants threshold. The dummy variable $(Post_t)$ is equal to 1 for elections starting from 2001.[21] The temporal dummy variable $(Post_t)$ has been built in this way because the

[19] As described in section 4.2, the population of reference used for the application of fiscal rules is not the same as the population used for the wage. However, the dummy variable for fiscal rules treatment and the one for a higher wage overlap for most municipalities, and only in a very small number of cases the two variables do not coincide. Specifically, the correlation between the two treatments is equal to 0.97. For this reason, an RDD approach would not allow us to disentangle between the effect of fiscal rules and the effect of a higher wage.

[20] In previous versions of this paper, we have used as the population of reference the one measured in the most recent Census produced by the Italian Statistical Office (Istat), either in 1991 or 2001. We get similar results using the Census population. Results can be made available upon request.

[21] For example, for a municipality that voted in 1995, 1999, 2004 and 2009, $(Post_t)$ is 0 for the electoral terms 1995 and 1999 and equal to 1 for 2004 and 2009.

selection of (new) politicians can happen only during electoral years and not during the electoral mandate (i.e., far away from elections). Also, for this reason, differently from Grembi et al. (2016), who use yearly data, we estimate the model using data at the electoral year level. 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 municipalities that are exempt from their application starting from 2001. The reason for selecting the electoral years from 2001 as the treatment period is that fiscal rules started to apply differently across the 5000 inhabitants threshold only from this year. Therefore, we can expect a differential behavior in terms of political selection due to fiscal rules to emerge across the cutoff only from 2001.

We estimate model 7 by local linear regression (Gelman and Imbens, 2018), using the subsample of observations which lie in the interval $R_{it} \in [-h, +h]$ around the threshold, where the optimal bandwidth h is calculated following the Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth. We cluster standard errors at the local labor market area level.[22] 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 5000 inhabitants threshold before and after 2001, such that municipalities must not be able to self-select themselves and decide on which side of the cutoff to stay. We test this assumption in Figure A1, which shows that there is not any discontinuity in R_{it} at the 5000 inhabitants threshold.[23] Second, other potential outcomes and municipal characteristics must be balanced around the threshold before and after 2001. We test this assumption by running model 7 using municipal characteristics as dependent variables. Table A4 shows that municipal and geographical characteristics are balanced around the threshold before and after 2001. Finally, municipalities just below and just above the 5000 inhabitants 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

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

[23]In Figure A1, we present scatters and 4th-order polynomial estimates to test the null hypothesis of the continuity of the density of the population around the 5000 inhabitants threshold. We apply this test to both R_{it} measured in the years before 2001 and the years from 2001. We also test the continuity of the difference between the density of the average population from 2001 and the average population before 2001. As we can observe in the bottom graph, there is no evidence of sorting.

5.3.1.

5.1.2 Regression discontinuity design

To test implication 2 of the model, we use a Regression Discontinuity Design (RDD) strategy based on close mixed electoral competitions, in which graduate mayors compete against non-graduate ones. We exploit the fact that, in mixed races decided by a narrow margin, the election outcomes are determined by random factors and not by systematic municipal characteristics that could also affect policy outcomes. Hence, under certain assumptions, municipalities where mayors with a university degree barely lost can be used as a counterfactual for municipalities where they barely won. Following the recent developments introduced by Calonico, Cattaneo, and Titiunik (2014), Calonico, Cattaneo and Farrell (2018), and Gelman and Imbens (2018), we estimate the following RDD strategy by local linear regression (LLR) using data at the municipality and electoral year level:

$$Y_{it} = \rho_0 + \rho_1 MV_{it} + \beta_0 Graduate_{it} + \beta_1 Graduate_{it} \cdot MV_{it} + \eta_{it} \tag{8}$$

where the dependent variable Y_{it} is a dummy variable equal to 1 if a mayor chooses the correct policy to match the state of the economy.[24] 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. At the threshold $MV_{it} = 0$ the level of education of the mayor sharply changes from 0 to 1.

We run model 8 on the sub-sample of municipalities in the interval $MV_{it} \in [-h, +h]$, where the optimal bandwidth h is calculated following Calonico, Cattaneo, and Titiunik (2014), and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth. The coefficient of interest is β_0, which identifies the average treatment effect (ATE) of mayors with a university degree at the threshold $MV_{it} = 0$. In the analysis below, we report conventional 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). We cluster standard errors at the local labor market area level. There are two main assumptions required for this identification to

[24] As described in more detail in section 5.3.3, Y_{it} is 1 if the mayor runs a deficit above the median when the economic growth is low or if the mayor runs a deficit below the median when economic growth is high.

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 A4 and A5, using the test on the continuity of the density of the running variable proposed by Cattaneo, Jansson, and Ma (2018). The evidence in Figures A4-A5 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 A12 and A13, which confirm that municipal covariates are balanced.

5.2 Data

We use data from Italian municipalities with less than 15,000 inhabitants for the period 1993-2015. There are various reasons for this choice of the sample. First, municipalities with less than 15,000 inhabitants use a single-ballot majoritarian electoral system, while municipalities above the threshold use a run-off system (Gamalerio, Morelli, and Negri, 2021). To keep electoral institutions constant, we exclude municipalities with more than 15,000 inhabitants. 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). At the same time, the DSP, introduced in 1999 by the central government, has remained in place until 2015. For these reasons, we have collected data on municipal politicians and municipalities' characteristics for the 1993-2015 period. Finally, municipalities from Special Regions (i.e., Sardegna, Sicilia, Valle d'Aosta, Trentino-Alto Adige, Friuli-Venezia Giulia) are excluded, given that they have different political and fiscal institutions, and since 2002 apply a different set of fiscal rules.

The dataset contains information on the characteristics of elected municipal politicians and mayoral candidates for the years 1993-2015. The main observable characteristics are gender, age, years of past political experience at all levels of politics, political orientation (i.e., left, right or independent), past professional background, education. The Italian Home Office provides information about these characteristics. Information on municipalities' characteristics comes from the Italian Statistical Office (Istat). It includes the share of the population with a university degree measured in 2001, the share of the active population (i.e., the population between 15 and 64 years old) measured in 2001, the share of elderly (i.e.,

population above 65 years old) measured in 2001, the income per capita measured in 2001, the number of firms per capita measured in 2005, the number of no-profit associations per capita measured in 2005, the area of the municipality in square kilometers, and population density measured in 2001. We use all these variables as control variables. We have collected the data on municipal budget outcomes from the Aida PA database, an online archive managed by the Bureau Van Dijk. Data contains information on the fiscal items of the budget of all Italian municipalities, and it covers the year from 2000 up to 2015. Finally, data on average income and income growth rate at the municipal level are provided by the Italian Ministry of Economics and Finance and covers 2000-2016.

The final sample consists of 26,005 electoral terms and 6170 municipalities. Table A3 reports the summary statistics of this sample, distinguishing between municipalities below and municipalities above the cutoff.

5.3 Results

We divide the description of the main results into four parts. First, we present empirical evidence related to the main prediction of the theoretical model. Second, we provide evidence on implication 1. Third, we describe the empirical results that refer to implication 2. Finally, we investigate alternative stories.

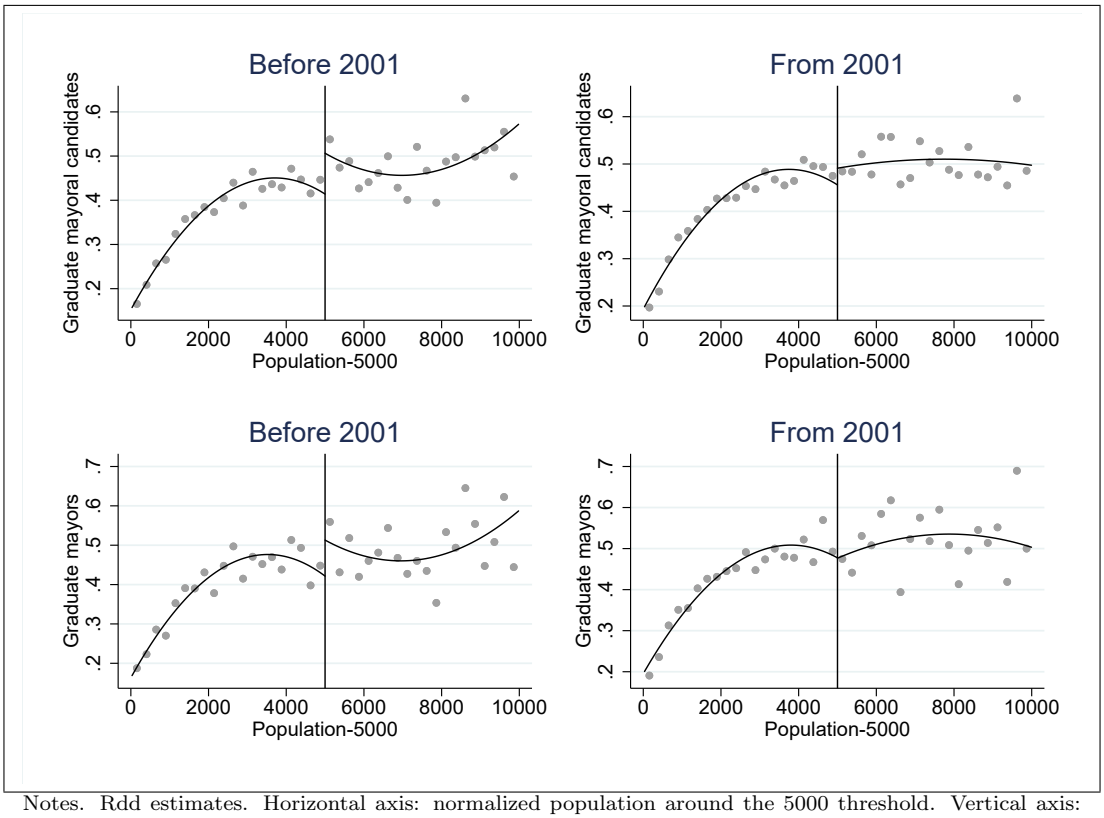
5.3.1 The effect of fiscal rules on the selection of politicians

In this sub-section, we test the main prediction of the theoretical model. We produce this evidence studying the 1993-2012 period and exploiting the 2001 fiscal rules relaxation. Figure 1 gives a preliminary idea about how politicians' education evolved between the years before 2001 - i.e., during which the only difference across the 5000 inhabitants threshold was the wage paid to politicians - and the years after the 2001 relaxation of fiscal rules - i.e., when also fiscal rules application differed across the cutoff. We produce this descriptive introductory evidence running a regression discontinuity design (RDD) based on a second-order split polynomial of politicians' education in the municipal population[25] for the electoral years

[25]One of the main purposes of this picture is to provide evidence of the discontinuity in politicians' education due to the wage paid to politicians in the years before 2001 (Gagliarducci and Nannicini, 2013). For this reason, we have produced this picture using the running variable based on the Census population, which is the population of reference for identifying mayors' wages. Using the running variable based on the resident population measured at the end of the previous penultimate year, the one that determines the application of fiscal rules (see section 4.2 for more detail), we get a similar picture. Results can be made available upon request.

before and after 2001.[26] Two facts emerge from Figure 1. First, there is a discontinuity across the 5000 inhabitants threshold for the electoral years before 2001. In those years, during which fiscal rules did not apply differently across the threshold, the share of graduate mayoral candidates and elected mayors was higher in municipalities just above the threshold. This result is consistent with the fact that municipalities above 5000 inhabitants pay a higher wage to the mayor, which enables them to attract more skilled politicians (Gagliarducci and Nannicini, 2013). Second, the discontinuity around the 5000 inhabitants threshold disappears in the years from 2001 up to 2012. The effect of the wage disappears from 2001, since when fiscal rules apply differently across the cutoff.

Figure 1: Rdd graphical evidence



Notes. Rdd estimates. Horizontal axis: normalized population around the 5000 threshold. Vertical axis: share of graduate mayoral candidates in the top graphs and dummy=1 for graduate mayors in the bottom graphs. Scatter points are averaged over bins of 250 inhabitants. The central black line represents a split second-order polynomial of the outcome variable in the normalized population, fitted separately on each side of the threshold.

[26]The same evidence emerges if we use different polynomial orders or a more local approach, e.g., using the Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth estimator and a linear polynomial. Results can be made available upon request.

To confirm this evidence, we implement the *Diff-in-Disc* analysis, and we run model 7 on the sample of municipalities individuated by the Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018) MSE-optimal bandwidth selector. The baseline results from the *Diff-in-Disc* analysis are in Table 1. We report the estimates obtained running model 7 using the optimal bandwidth with a linear function in the running variable in all columns of Table 1. We do not control for covariates in column 1, while we control for the year of election fixed effects in column 2. In column 3, we also add municipal covariates and Regions fixed effects. Two main results emerge from Table 1. First, the positive coefficients in front of the dummy variable ($> 5000_{it}$) indicate that in the years before 2001 (i.e., when fiscal rules applied in the same way across the cutoff), in municipalities just above 5000 inhabitants, individuals with higher education were running as mayoral candidates. This evidence is consistent with Figure 1 and with the analysis implemented by 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 in municipalities above 5000 inhabitants offsets the positive selection effect induced by the higher wage paid. The results indicate that fiscal rules induced a reduction in the share of graduate mayoral candidates, which is approximately around 10 % points. They also indicate that fiscal rules negatively affected the probability of electing a graduate mayor with a reduction of approximately the same magnitude.[27]

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 municipalities just below and just above the 5000 inhabitants threshold must have been on parallel trends before the 2001 reform (see section 5.1.1 for the evidence on the other two assumptions). We provide evidence on the parallel trends assumption in two ways. First, we run a series of cross-sectional RDD regressions comparing the level of education of politicians across municipalities just below and just above the 5000 inhabitants thresholds. We report the estimated coefficients of these cross-sectional RDD regressions in Figure 2. In this Figure, we

[27]Figure A2 shows how the estimated coefficients change with the bandwidth used. As typical in an RDD setup, moving toward smaller bandwidths produces bigger coefficients (i.e., smaller bias) and larger confidence intervals (i.e., more inefficiency). Besides, Figure A3 shows that the results are not due to random chances. Specifically, we run a series of *Diff-in-Disc* local linear regressions at 500 fake thresholds below and 500 fake thresholds above the 5000 inhabitants threshold (i.e., thresholds from 4900 to 4400, and from 5100 to 5600). The c.d.f. of the t-statistics from these regressions in Figure A3 shows that most of the t-statistics lie in the interval (-2,2), suggesting that it is not possible to find statistically significant results at fake thresholds.

Table 1: The effect of fiscal rules on the education of politicians

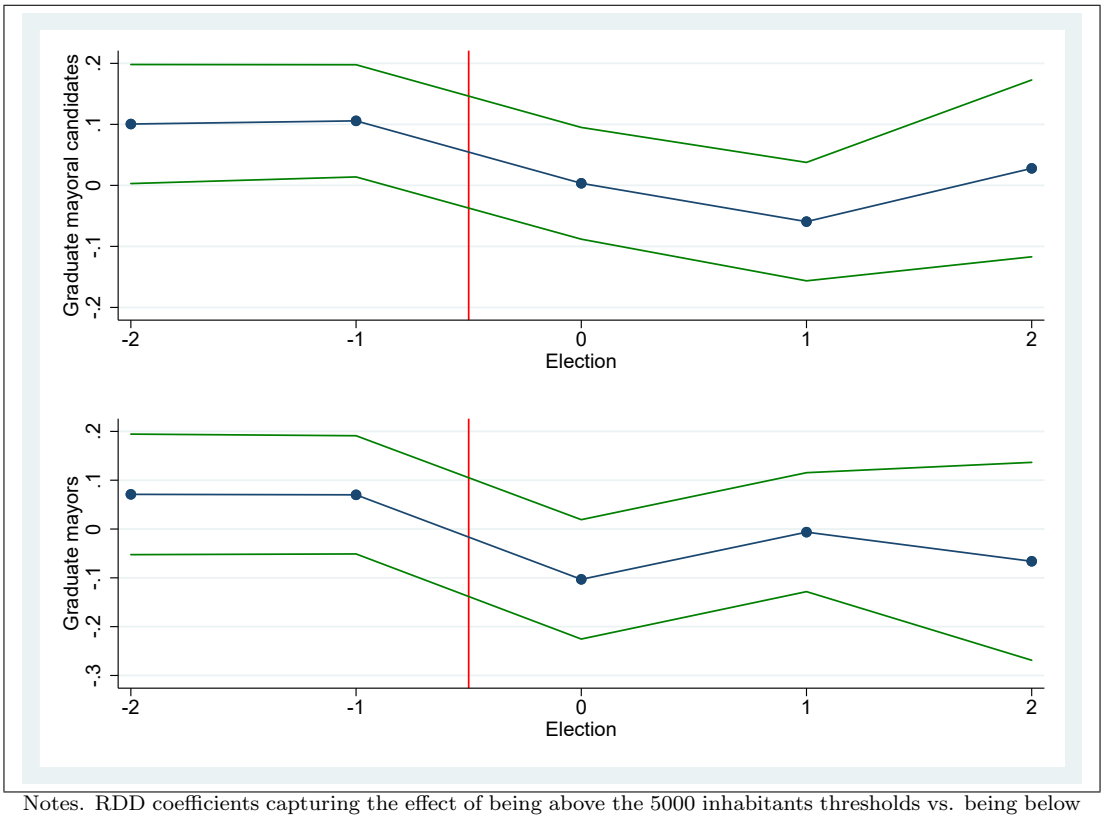
	(1)	(2)	(3)
Control Function	Linear	Linear	Linear
Bandwidth	CCT	CCT	CCT
Election Year FE	No	Yes	Yes
Region FE	No	No	Yes
Controls	No	No	Yes
Panel A: mayoral candidates with university degree			
(> 5000)	0.091***	0.088**	0.071**
	(0.034)	(0.034)	(0.033)
(Post)	0.059**		
	(0.030)		
(Post)*(> 5000)	-0.102**	-0.107**	-0.107**
	(0.044)	(0.044)	(0.042)
Observations	3,576	3,576	3,576
Bandwidth (h)	1166	1166	1166
Mean outcome	0.472	0.472	0.472
Panel B: mayors with university degree			
(> 5000)	0.055	0.054	0.039
	(0.048)	(0.048)	(0.048)
(Post)	0.057		
	(0.047)		
(Post)*(> 5000)	-0.107*	-0.114*	-0.113*
	(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 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; 3) (Post)*(> 5000) = interaction term between (> 5000) and Post. The outcome variable is the share of mayoral candidates with a university degree in Panel A, and it is = 1 for mayors with a university degree in Panel B. The bandwidth is calculated using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Year of election FE in columns 2 and 3, and Regions fixed effects in column 3. Control variables in column 3: 1) share of population with a university degree measured in 2001; 2) share of active population (i.e. population between 15 and 64 years old) measured in 2001; 3) share of elderly (i.e. population above 65 years old) measured in 2001; 4) log of income per capita measured in 2001; 5) number of firms per capita; 6) number of no-profit associations per capita; 7) area of municipality in square km; 8) population density measured in 2001. Robust standard errors clustered at the local labor area level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

run different cross-sectional RDD regressions, grouping in the same regression observations from different municipalities depending on the distance of the electoral year from the 2001 fiscal rules relaxation. For example, as reported on the x-axis of the two graphs in Figure 2, a value of 0 indicates the first elections run immediately after the 2001 relaxation. The value -1 refers to the elections run immediately before the 2001 reform, while -2 indicates elections

run before the elections at time -1. These cross-sectional RDD coefficients allow studying the evolution over time of the effect of being above the 5000 inhabitants thresholds on politicians' education. As we can observe, in the two elections before the 2001 fiscal rules removal (i.e., -2 and -1, for which fiscal rules applied in the same way across the threshold), municipalities above the thresholds were attracting more educated individuals into politics. 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, this effect disappears in the elections from 2001 (i.e., elections 0,1, and 2, for which fiscal rules applied differently across the threshold).

Figure 2: Cross-sectional RDD coefficients over time



Notes. RDD coefficients capturing the effect of being above the 5000 inhabitants thresholds vs. being below the threshold. On the x-axis, which goes from -2 to 2, we report the elections before and after the 2001 fiscal rules remove, where 0 indicates the elections immediately after the fiscal rules relaxation. We run the cross-section RDD regressions using the optimal CCT bandwidths reported in Table 1.

Second, we run model 7 on the pre-2001 period only (i.e., elections -1 and -2 in Figure

2), introducing a placebo treatment for the elections run immediately before the 2001 fiscal rules relaxation. This placebo treatment is equal to the interaction term between a dummy variable that we call (Pre_t), which is equal to 1 for the electoral terms immediately before the 2001 fiscal rules removal,[28] and the dummy variable ($> 5000_{it}$). The parameter in front of the interaction term (Pre_t)*($> 5000_{it}$) allows testing for parallel trends between the treatment and the control groups during the pre-treatment period. We report the results of this regression in Table 2. As shown in the Table, we do not find any evidence of differential pre-trends across municipalities below and above the 5000 inhabitants threshold.[29] This result indicates that municipalities just below and just above the 5000 inhabitants threshold were following the same trend in the pre-treatment period between the first election available in the dataset since 1993 and the election immediately before the 2001 fiscal rules relaxation.[30]

5.3.2 Evidence on financially constrained municipalities

This section provides evidence on implication 1 of the theoretical model. Specifically, we show that the effect of fiscal rules disappears when we consider municipalities that are already financially constrained. The intuition is that a pre-existing financial constraint reduces politicians' discretion and policy space, making political office less attractive for higher educated individuals as fiscal rules do. To provide this evidence, we repeat the *Diff-in-Disc* analysis distinguishing between municipalities that, at the time of the first election since 2001, inherited from the previous municipal government a higher level of debt compared to those that inherited a lower level of debt. The intuition of this heterogeneity analysis is that more public debt inherited from the previous government represents a constraint on what a mayor can do, which can reduce the discretion in setting fiscal policies as fiscal rules can

[28]For example, for a municipality that voted in 1995, 1999, 2004, and 2009, (Pre_t) is equal to 1 for the electoral term 1999 and to 0 otherwise. For a municipality that voted in the years 1994, 1998, 2003 and 2008, (Pre_t) is equal to 1 for the electoral term 1998 and to 0 otherwise.

[29]In addition, Table A5 provides further evidence on the evolution of the fiscal rules' effect during the pre-treatment period. Specifically, Table A5 shows that there has not been interaction between fiscal rules and the differential wage paid across the cutoff. This evidence is required to demonstrate that towns across the threshold did not react differently to the introduction of fiscal rules (Grembi et al., 2016). To test that this is the case, we interact the dummy variable ($> 5000_{it}$) with ($>= 1999$), which is equal to 1 for the electoral years 1999-2000. We run model 7 using this interaction as the main variable of interest. Table A5 shows that the interaction ($>= 1999$) $*$ ($> 5000_{it}$) is never statistically different from zero.

[30]Conversely, in Table A10, we report the estimates obtained running a simple difference-in-differences model on the entire original sample. As we can see, while the results go in the same direction of the *Diff-in-Disc* model, the parallel trends assumption does not seem to apply in this context. The violation of this assumption provides a justification for the use of the *Diff-in-Disc* model (Grembi et al., 2016), which allows us to compare more similar municipalities.

Table 2: Parallel trends test

	(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
Controls	No	No
(Pre)*(> 5000)	0.026	0.042
	(0.054)	(0.067)
Observations	1,966	2,210
Bandwidth	1364	1534
Mean outcome	0.450	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: $(Pre)*(> 5000)$= interaction between dummy = 1 for election immediately before 2001 fiscal rules removal and dummy = 1 for municipalities with more than 5000 inhabitants. The outcome variable is the share of mayoral candidates with a university degree in column 1, while it is = 1 for mayors with a university degree in column 2. The bandwidth is calculated using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Robust standard errors clustered at the local labor area level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

do. This intuition of a high inherited public debt as a financial constraint is also consistent with the literature on the political economy of fiscal deficits and public debt (Persson and Svensson, 1989; Alesina and Tabellini, 1990; Pettersson Lidbom, 2001; Alt and Lassen, 2006; Eslava, 2010). This literature explains how public debt and fiscal deficits produced by previous governments can represent a constraint on what the current government can implement, thus reducing the level of policymaking discretion. Hence, if implication 1 is correct, we should find that municipalities that are not constrained by fiscal rules but inherited more public debt behave as municipalities limited by fiscal rules. Besides, we should expect the baseline effect to be driven by the group of towns that inherited less debt.

To test this idea, in columns 1 and 3 of Table 3, we run the *Diff-in-Disc* using only the subsample of municipalities with a level of the inherited debt below the median. In columns 2 and 4, we run the regressions keeping only the municipalities with a level of municipal inherited debt above the median. As a measure for inherited debt, we use the deficit per capita produced in the final year of the electoral term run just before the 2001 fiscal rules relaxation (i.e., the new debt issued by the previous government and passed to the new government for which we evaluate the effect of fiscal rules in terms of political selection). The negative effect of fiscal rules on the education of politicians is driven by municipalities

that inherited less debt. At the same time, there are no differences between municipalities just above and below the threshold in the subsample of towns that inherited a high level of debt. This evidence further confirms that more educated politicians enter politics with a lower probability if they cannot enjoy a high level of discretion in setting fiscal policies.

Table 3: The role of past deficit

	(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
Controls	Yes	Yes	Yes	Yes
Sample	*Deficit<median*	*Deficit>median*	*Deficit<median*	*Deficit>median*
(Post)*(> 5000)	-0.176***	-0.028	-0.238**	-0.007
	(0.065)	(0.053)	(0.093)	(0.097)
P-value difference	0.079		0.080	
Observations	1,637	2,509	1,741	2,131
Bandwidth	1013	1653	1094	1405
Mean outcome	0.500	0.449	0.502	0.483

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) *(Deficit < median)* = municipalities with a level of past deficit as a fraction of total revenues below the median; 2) *(Deficit > median)* = municipalities with a level of past deficit as a fraction of total revenues above the median. Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001; 3) (Post)*(> 5000) = interaction term between > 5000 and Post. The outcome variable is the share of mayoral candidates with a university degree in column 1-2 and it is = 1 for mayors with a university degree in column 3-4. The bandwidth is calculated using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Year of election and Regions fixed effects added in all columns. Control variables in all columns: 1) share of population with a university degree measured in 2001; 2) share of active population (i.e. population between 15 and 64 years old) measured in 2001; 3) share of elderly (i.e. population above 65 years old) measured in 2001; 4) log of income per capita measured in 2001; 5) number of firms per capita; 6) number of no-profit associations per capita; 7) area of municipality in square km; 8) population density measured in 2001. Robust standard errors clustered at the local labor area level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

5.3.3 Matching the state of the economy

In this subsection, we provide empirical evidence on implication 2 of the theoretical model. To test implication 2, we run the RDD model described by equation 8 to investigate whether educated politicians have a higher probability of matching the state of the economy compared to the less educated ones. Besides, we study whether the differences in behavior between more and less educated politicians disappear where fiscal rules apply. We implement this analysis using the subsample of mixed electoral competitions between graduate and non-graduate mayoral candidates, considering only electoral terms between 2001 and 2012. We use this

sample of mixed electoral competition and then we calculate the Calonico, Cattaneo, and Titiunik (2014) and Calonico, Cattaneo, and Farrell (2018) MSE-optimal bandwidth around the $MV_{it} = 0$ threshold. We run the RDD regression keeping the observations within this optimal bandwidth. We report the results of this analysis in Table 4. Following the procedure developed by Calonico, Cattaneo and Farrell (2018), we report conventional 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).

To run the analysis reported in Table 4, we build a dependent variable which is equal to 1 if a mayor chooses the correct policy to match the state of the economy in a specific year and 0 otherwise.[31] After building this variable at the year level, we collapse the data at the municipality and electoral term level. Hence, the dependent variable obtained is equal to the probability of matching the state of the economy over the electoral mandate. Panel A 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 8 without control variables. In column 2, we check whether the estimated coefficients remain unchanged if we control for other personal characteristics of the mayors, such as age, gender, political experience, professional background, and political orientation. In column 3, we control for municipal characteristics, besides the regional and year of election fixed effects. The results of Table 4 show that, where fiscal rules do not apply, educated mayors have a higher probability of matching the state of the economy compared to less educated ones. Conversely, we do not find differences between graduate and non-graduate mayors in municipalities affected by fiscal rules.

[31] More in detail, to build this dependent variable, we create two dummy variables. The first is equal to 1 for mayors that run a deficit above the median in a specific year and 0 otherwise. The second dummy variable is equal to 1 for years characterized by a growth in municipal income below the median and 0 otherwise. Hence, the dependent variable for a mayor correctly matching the state of the economy in a specific year is equal to 1 when these two dummy variables are equal and 0 otherwise. Besides, to deal with the potential endogeneity of municipal income growth during a mayor's term, we create an alternative version of the dependent variable using the income growth predicted regressing income growth on pre-determined municipal characteristics and regional and years fixed effects. We get similar results if we use this alternative dependent variable for mayors correctly matching the state of the economy. Results can be made available upon request.

Table 4: 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
Municipal covariates	No	No	Yes
Panel A: municipalities below 5000			
Conventional	0.062**	0.076**	0.066**
	(0.030)	(0.030)	(0.029)
Bias-corrected	0.074**	0.087***	0.077***
	(0.030)	(0.030)	(0.029)
Robust	0.074**	0.087***	0.077**
	(0.033)	(0.034)	(0.032)
Observations	2094	2018	2005
Bandwidth	13.69	13.08	12.94
Panel B: municipalities above 5000			
Conventional	-0.021	-0.011	-0.022
	(0.038)	(0.038)	(0.036)
Bias-corrected	-0.016	-0.006	-0.018
	(0.038)	(0.038)	(0.036)
Robust	-0.016	-0.006	-0.018
	(0.046)	(0.046)	(0.042)
Observations	1165	1147	1111
Bandwidth	21.37	21.05	19.72

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: Graduate is a dummy variable =1 when the mayor has a university degree, 0 otherwise. Estimation by RDD-LLR using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Estimates reported: conventional RD estimates with a conventional variance estimator (Conventional), bias-corrected RD estimates with a conventional variance estimator (Bias-corrected), and bias-corrected RD estimates with a robust variance estimator are reported (Robust). Mayoral covariates included in columns 2-3: 1) age = 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 center-left mayor. Municipal covariates in column 3: 1) share of population with a university degree measured in 2001; 2) share of active population (i.e. population between 15 and 64 years old) measured in 2001; 3) share of elderly (i.e. population above 65 years old) measured in 2001; 4) log of income per capita measured in 2001; 5) number of firms per capita; 6) number of no-profit associations per capita; 7) area of municipality in square km; 8) population density measured in 2001; 9) past deficit from previous term. Year of election and Regional FE included in column 3. Robust standard errors clustered at the local labor area level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

5.3.4 Alternative stories

We deal with several potential alternative explanations to our main story. First, one potential alternative story is that the effect is due to a reduction in real terms of the difference in wages

paid to mayors across the threshold. To rule out this possibility, in Table A6, we repeat the analysis using all the electoral years between 1993 and 2015 and exploit the fact that the fiscal rules' application threshold was reduced from 5000 to 1000 inhabitants for the years 2013-2015 (Daniele and Giommoni, 2020). More in detail, we rerun model 7 adding an interaction term between a dummy variable equal to 1 for elections from 2013 ($>= 2013_t$) and the dummy variable ($> 5000_{it}$). Suppose the effect was due to a reduction in real terms of the difference in wages. In that case, we should expect the negative effect across the threshold to persist even after fiscal rules were reintroduced for municipalities below 5000 inhabitants. The results in Table A6 show that this is not the case, given that, consistent with our story, the coefficients in front of $(>= 2013_t) * (> 5000_i)$ are small and not statistically different from zero.

Second, we show that the presence of different out-of-politics outside options for individuals with different levels of education is not sufficient to explain our results. In principle, it may be possible that fiscal rules affect the office value of individuals with a different level of education in the same way. This homogeneous effect could affect the entry into politics of individuals with different levels of education heterogeneously given their different outside options. If high-educated individuals have a better outside option in the labor market compared to low-educated ones, the overall effect could be a reduction in the quality of candidates. Table A7 seems to rules out this alternative story. Specifically, in Table A7, we use data on the municipal shares of employed individuals divided by income brackets to measure the probability that highly-educated and low-educated individuals have different opportunities in the labor market. To do so, we calculate a Herfindahl index of these income brackets share to measure whether employed individuals are concentrated in one or few specific income brackets. Higher values of this index suggest a bigger concentration in one specific bracket and thus more homogeneous opportunities for individuals with different levels of education. The results in Table A7 indicate that the results do not change across municipalities with low vs. high values of the Herfindahl index, suggesting that different out-of-politics outside options do not seem to play the main role in this context.

Third, the application of fiscal rules may require the selection of more politically experienced politicians, who may be less educated. To rule out the second potential explanation, in Table A8, 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. For

characteristics potentially correlated with education, the estimated coefficient goes in the expected direction (i.e., a decline in the share of politicians from high skills 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.[32] Fourth, fiscal rules may affect politicians' political orientation, which in turn is correlated with their level of education. Table A9 excludes any effect of fiscal rules on politicians' political orientation. Fifth, fiscal rules may change mayors' probability of being re-elected. Results in Panel B of Table A4 show that this is not the case. In addition, in Table A11, 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. Specifically, graduate candidates receive more votes, reach a better final ranking position, and are more likely to be elected mayors. The results go in the same direction independently on whether we consider municipalities and electoral years characterized or not by the presence of fiscal rules. The evidence in Tables A4 and A11 suggests that voters did not change their electoral behavior due to the introduction of fiscal rules and that they maintained a preference towards more educated politicians.

Moving to the RDD analysis, the different results for municipalities below and above the 5000 inhabitants thresholds may be due to the different wage paid to the mayors (Gagliarducci and Nannicini, 2013). To rule out this possibility, we repeat the RDD exercise using only fiscal outcomes measured in 2013-2015, during which fiscal rules applied equally across the threshold. Conversely, during these years, the wage increase across the threshold was in place. Table A14 shows that the differences in matching the state of the economy disappear when fiscal rules apply in the same way across the threshold.

Finally, educated mayors may be more corrupt than non-graduate ones. As shown by Daniele and Giommoni (2020), the introduction of fiscal rules should make it more challenging to extract rents, reducing the office value for individuals attracted by them. If graduate individuals are more corrupt than non-graduate ones, the introduction of fiscal rules may make them less interested in entering politics. However, as we can see from Table A15, graduate mayors do not appear to be more corrupt than non-graduate ones.[33]

[32]For data limitations, it was possible to reconstruct the past political experience only for elected mayors, and not for mayoral candidates.

[33]More in detail, we use the web archive of one of the leading Italian newspapers (La Repubblica), to find episodes of corruption linked to the mayors in my analysis. Using an algorithm based on the mayor's first

6 Concluding remarks

This paper investigates the effect of a reduction in policymaking discretion on the selection of politicians. Specifically, it shows theoretically and empirically that a reduction in policymaking discretion, induced by the application of fiscal rules, negatively affects politicians' level of education. The results of this paper have four implications for future research. First, it would be interesting to analyze whether electing fewer educated politicians can lead to worse policies. The existing literature (Besley et al., 2011; Martinez-Bravo, 2017) on the implications of educated leaders suggests that this should be the case. For what concerns the Italian context, Daniele and Giommoni (2020) show that fiscal rules reduce corruption more in municipalities with more educated mayors. Besides, Mitra (2020) shows that educated mayors increase public investment in education without worsening the municipality's financial situation. Also, Carreri (2020) uses a measure of executive politicians' administrative competence that positively correlates with education to show that competent mayors are associated with better policies. The evidence in all these papers calls for further research on the policy implications of reducing the share of competent elected politicians.

Second, this paper analyzes the ex-ante quality of the political class, but it does not say anything about the political class's representativeness relative to the electorate. As suggested by the recent literature (Dal Bo et al., 2020; Carreri, 2020), the fact that the election of fewer educated and skilled politicians represents a decrease in quality is only one side of the story. On the other side, the elections of fewer competent individuals may lead to a better representation of marginalized groups in terms of labor market performance and socioeconomic background. Based on the existing literature, we do not know much about whether a better representation of marginalized groups may lead to more targeted policies towards these groups. This intuition calls for more research on the representativeness of the political class.

Third, this paper provides evidence using data from one country and methodologies as *Diff-in-Disc* and RDD. Using local data from one country allows the paper to avoid the limits of cross-countries analysis. The methodologies used in the paper have strong internal

and last names, the name of the city, the years of the legislature, and a series of keywords recalling episodes of corruption, we create a database of newspapers' 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 A15 are estimated using this dummy variable as the dependent variable.

validity and provide casual estimates. However, the limit of the analysis is a potential cost in terms of external validity. This potential limitation calls for further research using data from other countries, which would enable researchers to understand whether the evidence provided in this paper also applies in other contexts. Fourth, the results of this paper could extend beyond the political economy application we consider. We document a positive effect of on-the-job discretion on the quality of "applicants" that could potentially apply to a wide range of employer-employee relationships. More research is needed to better understand applications and limitations of this result in "non-political" job interactions.

In conclusion, there are two policy implications related to our results. First, the paper suggests that the negative selection effect of fiscal rules should be taken into account when designing those rules. Although local, the effect is sizeable and significant, pointing at the fact that the quality of the pool of candidates is endogenous to the rules themselves. Second, this paper shows that the reduction in discretion implied by fiscal rules exactly compensates for the positive selection implied by higher wages paid just above the 5000 inhabitants threshold. Consequently, it suggests that one possible way to compensate for the negative selection effect of fiscal rules would be to combine them with a raise in politicians' wages.

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." *European Economic Review*, 50: 1430–1439.

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.

Ashworth, S., Berry, C. R., and E. Bueno de Mesquita (2021). *Theory and Credibility. Integrating theoretical and empirical social science.* Princeton, Princeton University Press.

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 .

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.

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.

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

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., Marino, I., Nannicini, T., and N. Persico (2020). "Demand Shocks and Firm Investment: Micro-evidence from fiscal retrenchment in Italy." *The Economic Journal*, conditionally 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.

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.

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

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., 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.

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

Gulzar, S., and M. Y. Khan (2021). ""Good Politicians": Experimental Evidence on Motivations for Political Candidacy and Government Performance." *working paper*.

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.

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.

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.

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

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.

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

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. (2021). "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.

A Additional tables

Table A1: 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 decided by the Italian central government which apply year by year to the covered municipalities. Columns definition: "Year" = year of application of the rule; "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 according to their resident population measured at the end of the penultimate previous 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.

Table A2: Legislative population thresholds in Italy: Municipalities below 15,000

Population	Wage Mayor	Wage Ministers	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 15000 inhabitants. Columns definition: Population = municipal population as measured by the last Census; Wage Mayor = it is the wage paid to the mayor, expressed in Euros at 2000 prices; Wage Ministers = wage paid to the ministers as a percentage of the wage of the mayor; Size Government = maximum numer of ministers that can be appointed in the municipal government; Size Council = numer of seats in the municipal council. All the wage thresholds date back to 1960, except the 1000 and 10,000 thresholds, which were introduced in 2000. Sources: Gagliarducci and Nannicini (2013); Grembi et al. (2016).

Table A3: 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	0.095	1334	0.229
Age mayors	48.23	4836	47.78	1334	0.023
High skills job mayors	0.227	4836	0.310	1334	0.000
Graduate mayors	0.373	4836	0.516	1334	0.000
Political experience mayors	8.25	4836	8.16	1334	0.490
Female mayoral candidates	0.104	4836	0.110	1334	0.213
Age mayoral candidates	48.10	4836	48.09	1334	0.936
High skills job mayoral candidates	0.213	4836	0.310	1334	0.000
Graduate mayoral candidates	0.355	4836	0.504	1334	0.000
Municipal characteristics					
South	0.252	4836	0.288	1334	0.008
Centre	0.136	4836	0.165	1334	0.006
North-West	0.504	4836	0.306	1334	0.000
North-East	0.107	4836	0.239	1334	0.000
Population density	145.64	4836	496.30	1334	0.000
Area	25.315	4836	43.145	1334	0.000
No profit associations	9.830	4836	34.327	1334	0.000
Firms per capita	0.075	4836	0.081	1334	0.000
Income per capita	9084	4836	10335	1334	0.000
% elderly	0.228	4836	0.176	1334	0.000
% 15-64 years old	0.643	4836	0.677	1334	0.000
% graduate	0.043	4836	0.051	1334	0.000

Notes. Municipalities between 0 and 15,000 inhabitants. Electoral terms between 1993 and 2012. *Below 5000* = 1 for municipalities below 5000 inhabitants. *Above 5000* = 1 for municipalities above 5000 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 A4: 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)	-0.001	-0.007	-0.002	-0.028	-0.001	-1.466	0.297	-11.640
	(0.002)	(0.004)	(0.006)	(0.049)	(0.003)	(1.358)	(5.241)	(51.774)
Observations	2,919	2,644	2,509	3,490	2,334	2,679	3,504	4,012
Bandwidth	944.6	859.8	818.2	1139	768.2	872.5	1142	1296
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	0.001		
	(0.050)	(0.061)	(0.040)	(0.064)	(0.068)	(0.062)		
Observations	3,865	3,772	5,105	3,182	2,706	3,412		
Bandwidth	1246	1220	1644	1034	879.8	1689		

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 5000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001; 3) (Post)*(> 5000) = interaction term between > 5000 and Post. The bandwidth is calculated using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Robust standard errors clustered at the municipality level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A5: 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
Controls	No	No
(>= 1999)*(> 5000)	-0.040	-0.043
	(0.060)	(0.077)
Observations	1,966	2,210
Bandwidth	1364	1534
Mean outcome	0.450	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 5000 inhabitants. The outcome variable is the share of mayoral candidates with a university degree in column 1, while it is = 1 for mayors with a university degree in column 2. The bandwidth is calculated using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Robust standard errors clustered at the local labor area level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A6: Wage depreciation

	(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
Controls	No	No
(>= 2001)*(> 5000)	-0.102**	-0.107*
	(0.044)	(0.060)
(>= 2013)*(> 5000)	-0.033	0.009
	(0.075)	(0.090)
Observations	4,133	5,062
Bandwidth	1166	1425
Mean outcome	0.475	0.489

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 2015. Variables in the Table: (>= 2001)*(> 5000)= interaction between dummy = 1 for electoral years from 2001 up to 2012 and dummy = 1 for municipalities with more than 5000 inhabitants; (>= 2013)*(> 5000)= interaction between dummy = 1 for electoral years from 2013 up to 2015 and dummy = 1 for municipalities with more than 5000 inhabitants. The outcome variable is the share of mayoral candidates with a university degree in column 1, while it is = 1 for mayors with a university degree in column 2. The bandwidth is calculated using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Robust standard errors clustered at the local labor area level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A7: 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
Controls	Yes	Yes	Yes	Yes
Sample	Herfindal index > median	Herfindal index < median	Herfindal index > median	Herfindal index < median
(Post)*(> 5000)	-0.125***	-0.124*	-0.129*	-0.107
	(0.046)	(0.068)	(0.070)	(0.103)
Observations	2,287	1,599	2,713	1,526
Bandwidth	1233	1308	1510	1263
Mean outcome	0.411	0.569	0.422	0.574

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. Variables in the Table: 1) (> 5000) = 1 for municipalities with more than 5000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001; 3) (Post)*(> 5000) = interaction term between (> 5000) and Post. The outcome variable is the share of mayoral candidates with a university degree in columns 1-2, while it is = 1 for mayors with a university degree in columns 3-4. The bandwidth is calculated using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Year of election and Regions fixed effects added in all columns. Control variables in all columns: 1) share of population with a university degree measured in 2001; 2) share of active population (i.e. population between 15 and 64 years old) measured in 2001; 3) share of elderly (i.e. population above 65 years old) measured in 2001; 4) log of income per capita measured in 2001; 5) number of firms per capita; 6) number of no-profit associations per capita; 7) area of municipality in square km; 8) population density measured in 2001. Robust standard errors clustered at the municipality level are in parentheses. Robust standard errors clustered at the local labor area level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A8: 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
Covariates	No	No	No	No
Dependent Variables	High skill	Age	Female	Pol Experience
Panel A: mayoral candidates				
(Post)*(> 5000)	-0.104**	1.042	0.002	
	(0.047)	(0.823)	(0.025)	
Observations	2,820	4,319	3,997	
Bandwidth	913.3	1400	1290	
Mean outcome	0.290	47.93	0.113	
Panel B: mayors				
(Post)*(> 5000)	-0.089	1.277	0.011	-0.541
	(0.062)	(1.445)	(0.034)	(0.769)
Observations	3,510	3,554	3,596	3,994
Bandwidth	1158	1168	1172	1290
Mean outcome	0.309	47.89	0.086	8.135

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; 3) (Post)*(> 5000) = interaction term between (> 5000) and Post. The outcome variables are: 1) high skill = for politicians from high skill occupations; 2) Age = age of the politicians; 3) Female = 1 for female politicians; 4) Pol Experiences = years of politicial experience at any level of politics (for mayors only). The bandwidth is calculated using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Robust standard errors clustered at the local labor area level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A9: 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
Covariates	No	No	No	No
Dependent Variables	Left	Right	Center	Civic List
Panel A: mayoral candidates				
(Post)*(> 5000)	0.038	-0.016	0.013	-0.067
	(0.038)	(0.033)	(0.012)	(0.053)
Observations	3,653	4,309	4,701	3,394
Bandwdith	1185	1396	1529	1101
Mean outcome	0.203	0.215	0.0143	0.568
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	0.0122	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; 3) (Post)*(> 5000) = interaction term between (> 5000) and Post. 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 Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Robust standard errors clustered at the local labor area level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A10: The effect of fiscal rules on the education of politicians
Difference-in-differences estimates

	(1)	(2)	(3)	(4)
Election Year FE	No	No	No	No
Region FE	No	No	No	No
Covariates	No	No	No	No
Dependent Variables	Share mayoral candidates with university degree		= 1 for Mayors with university degree	
(> 5000)	0.154***	0.171***	0.135***	0.140***
	(0.013)	(0.014)	(0.016)	(0.017)
(Post)	0.036***	0.034***	0.023***	0.020**
	(0.005)	(0.006)	(0.007)	(0.008)
(Post)*(> 5000)	-0.019*	-0.037***	0.007	0.001
	(0.010)	(0.012)	(0.016)	(0.019)
Pre		-0.005		-0.006
		(0.005)		(0.006)
(Pre)*(> 5000)		-0.036***		-0.011
		(0.011)		(0.015)
Observations	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 5000 inhabitants; 2) (Post) = 1 for electoral terms starting from 2001; 3) (Post)*(> 5000) = interaction term between (> 5000) and Post; 4) (Pre) = 1 for election immediately before 2001 fiscal rules removal; 5) (Pre)*(> 5000) = interaction term between (> 5000) and (Pre). The outcome variable is the share of mayoral candidates with a university degree in columns 1 and 2, while it is = 1 for mayors with a university degree in columns 3 and 4. Robust standard errors clustered at the local labor area level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A11: 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	1.982***	-0.087***	0.035***
	(0.215)	(0.009)	(0.005)
Observations	52,436	58,568	58,568
Panel B: fiscal rules applied			
Graduate	2.395***	-0.120***	0.042***
	(0.344)	(0.017)	(0.009)
Observations	19,301	19,318	19,318
Panel B: fiscal rules did not applied			
Graduate	1.739***	-0.070***	0.031***
	(0.263)	(0.010)	(0.006)
Observations	33,135	39,250	39,250

Notes. OLS estimates. Municipalities below 15,000 inhabitants. Electoral terms between 1993 and 2012. 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. Year of election and Region fixed effects included in all columns. Mayoral candidates covariates included in all columns: 1) high skills job = 1 if candidate worked in a high skills 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 national political parties; 5) unemployed = 1 if candidate is unemployed. Municipal covariates in all columns: 1) share of population with a university degree measured in 2001; 2) share of active population (i.e. population between 15 and 64 years old) measured in 2001; 3) share of elderly (i.e. population above 65 years old) measured in 2001; 4) log of income per capita measured in 2001; 5) number of firms per capita; 6) number of no-profit associations per capita; 7) area of municipality in square km; 8) population density measured in 2001. Robust standard errors clustered at the municipality level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A12: 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
Conventional	-0.000	0.003	-0.004	-0.015	0.001	0.712	1.547	-6.117
	(0.002)	(0.005)	(0.007)	(0.036)	(0.003)	(0.611)	(2.742)	(22.834)
Bias-corrected	0.001	0.004	-0.005	-0.013	0.001	0.889	2.210	-5.096
	(0.002)	(0.005)	(0.007)	(0.036)	(0.003)	(0.611)	(2.742)	(22.834)
Robust	0.001	0.004	-0.005	-0.013	0.001	0.889	2.210	-5.096
	(0.002)	(0.006)	(0.008)	(0.041)	(0.003)	(0.691)	(3.124)	(24.582)
Observations	2281	2583	2643	2624	2723	2821	2873	3134
Bandwidth	15.24	17.91	18.53	18.32	19.26	20.28	20.79	23.48
Panel B: Geographical characteristics municipalities								
Dependent variables	NE	NW	CEN	SOU	Past deficit			
Conventional	0.023	-0.032	-0.048	0.050	0.002			
	(0.036)	(0.061)	(0.043)	(0.060)	(0.005)			
Bias-corrected	0.021	-0.026	-0.065	0.068	0.003			
	(0.036)	(0.061)	(0.043)	(0.060)	(0.005)			
Robust	0.021	-0.026	-0.065	0.068	0.003			
	(0.041)	(0.068)	(0.047)	(0.066)	(0.006)			
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 =1 when the mayor has a university degree, 0 otherwise. Estimation by RDD-LLR using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Estimates reported: conventional RD estimates with a conventional variance estimator (Conventional), bias-corrected RD estimates with a conventional variance estimator (Bias-corrected), and bias-corrected RD estimates with a robust variance estimator are reported (Robust). Municipal dependent variables in Panel A: 1) share of population with a university degree measured in 2001; 2) share of active population (i.e. population between 15 and 64 years old) measured in 2001; 3) share of elderly (i.e. population above 65 years old) measured in 2001; 4) log of income per capita measured in 2001; 5) number of firms per capita; 6) number of no-profit associations per capita; 7) area of municipality in square km; 8) population density measured in 2001. 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. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A13: 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
Conventional	-0.002	-0.001	0.002	-0.010	-0.002	3.050	3.715	-122.749
	(0.003)	(0.005)	(0.006)	(0.058)	(0.004)	(2.185)	(6.417)	(106.289)
Bias-corrected	-0.002	-0.001	0.004	-0.003	-0.003	3.597*	4.960	-141.923
	(0.003)	(0.005)	(0.006)	(0.058)	(0.004)	(2.185)	(6.417)	(106.289)
Robust	-0.002	-0.001	0.004	-0.003	-0.003	3.597	4.960	-141.923
	(0.004)	(0.005)	(0.007)	(0.064)	(0.004)	(2.486)	(7.348)	(114.794)
Observations	1101	1102	996	1054	808	1142	1152	1052
Bandwidth	20.73	20.75	17.71	19.24	13.50	21.92	22.33	19.02
Panel B: Geographical characteristics municipalities								
Dependent variables	NE	NW	CEN	SOU	Past deficit			
Conventional	0.111	-0.140	0.014	0.004	0.000			
	(0.079)	(0.098)	(0.060)	(0.074)	(0.004)			
Bias-corrected	0.105	-0.155	0.012	0.000	-0.000			
	(0.079)	(0.098)	(0.060)	(0.074)	(0.004)			
Robust	0.105	-0.155	0.012	0.000	-0.000			
	(0.087)	(0.107)	(0.068)	(0.081)	(0.005)			
Observations	887	907	1104	1236	704			
Bandwidth	15.23	15.64	20.96	24.75	15.75			

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 =1 when the mayor has a university degree, 0 otherwise. Estimation by RDD-LLR using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Estimates reported: conventional RD estimates with a conventional variance estimator (Conventional), bias-corrected RD estimates with a conventional variance estimator (Bias-corrected), and bias-corrected RD estimates with a robust variance estimator are reported (Robust). Municipal dependent variables in Panel A: 1) share of population with a university degree measured in 2001; 2) share of active population (i.e. population between 15 and 64 years old) measured in 2001; 3) share of elderly (i.e. population above 65 years old) measured in 2001; 4) log of income per capita measured in 2001; 5) number of firms per capita; 6) number of no-profit associations per capita; 7) area of municipality in square km; 8) population density measured in 2001. 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. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A14: 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
Municipal covariates	No	No	Yes
Panel A: municipalities below 5000			
Conventional	-0.060	-0.038	-0.051
	(0.048)	(0.049)	(0.047)
Bias-corrected	-0.077	-0.053	-0.064
	(0.048)	(0.049)	(0.047)
Robust	-0.077	-0.053	-0.064
	(0.055)	(0.055)	(0.055)
Effective Observations	1119	1101	1059
Bandwidth	15.34	15.76	14.90
Panel B: municipalities above 5000			
Conventional	-0.087	-0.060	-0.048
	(0.074)	(0.071)	(0.074)
Bias-corrected	-0.083	-0.067	-0.048
	(0.074)	(0.071)	(0.074)
Robust	-0.083	-0.067	-0.048
	(0.089)	(0.084)	(0.088)
Observations	473	465	392
Bandwidth	15.17	16.22	12.64

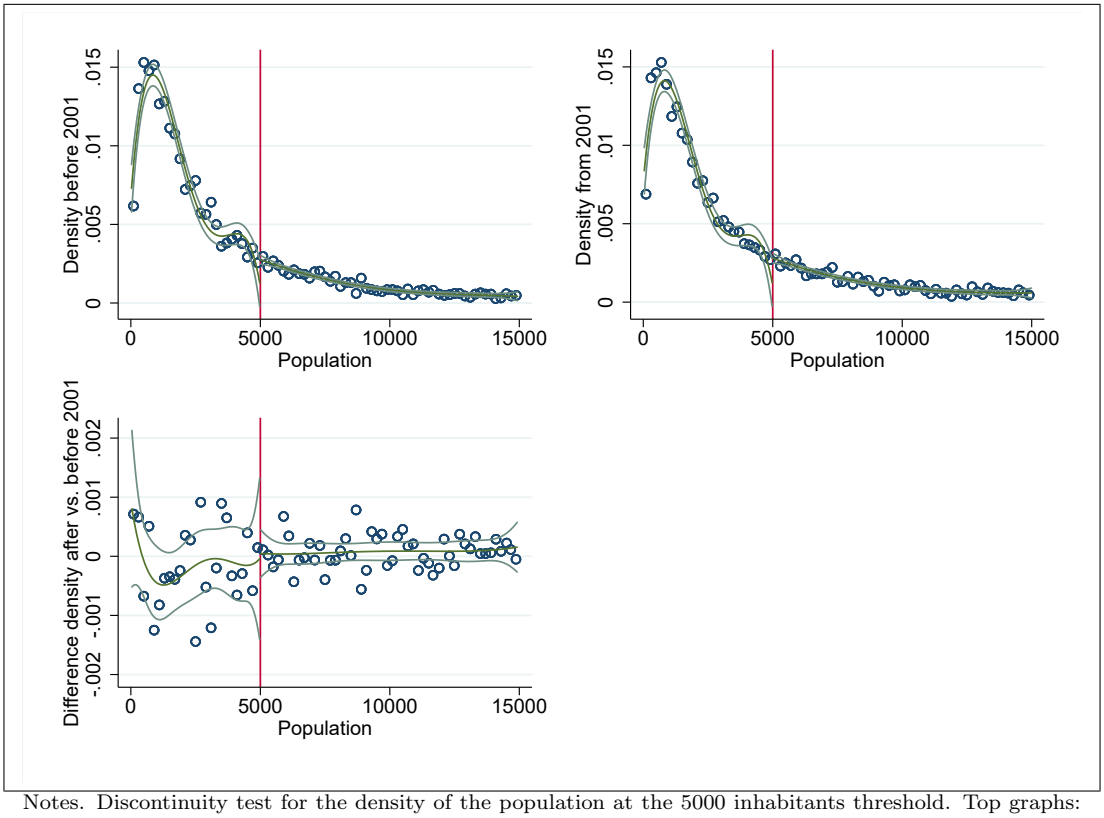
Notes. Municipalities below 15,000 inhabitants. Years 2013-2015. Dependent variable = 1 if deficit above median and income growth below median or deficit below median and income growth above median. Treatment variable: Graduate is a dummy variable =1 when the mayor has a university degree, 0 otherwise. Estimation by RDD-LLR using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Estimates reported: conventional RD estimates with a conventional variance estimator (Conventional), bias-corrected RD estimates with a conventional variance estimator (Bias-corrected), and bias-corrected RD estimates with a robust variance estimator are reported (Robust). Mayoral covariates included in columns 2-3: 1) age = 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 center-left mayor. Municipal covariates in column 3: 1) share of population with a university degree measured in 2001; 2) share of active population (i.e. population between 15 and 64 years old) measured in 2001; 3) share of elderly (i.e. population above 65 years old) measured in 2001; 4) log of income per capita measured in 2001; 5) number of firms per capita; 6) number of no-profit associations per capita; 7) area of municipality in square km; 8) population density measured in 2001; 9) past deficit from previous term. Year of election and Regional FE included in column 3. Robust standard errors clustered at the local labor area level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table A15: 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				
Conventional	-0.008	0.002	-0.005	-0.029
	(0.015)	(0.013)	(0.040)	(0.038)
Bias-corrected	-0.004	0.005	-0.011	-0.029
	(0.015)	(0.013)	(0.040)	(0.038)
Robust	-0.004	0.005	-0.011	-0.029
	(0.018)	(0.015)	(0.049)	(0.046)
Effective Observations	2641	2459	1025	900
Bandwidth	18.48	16.57	17.64	14.77

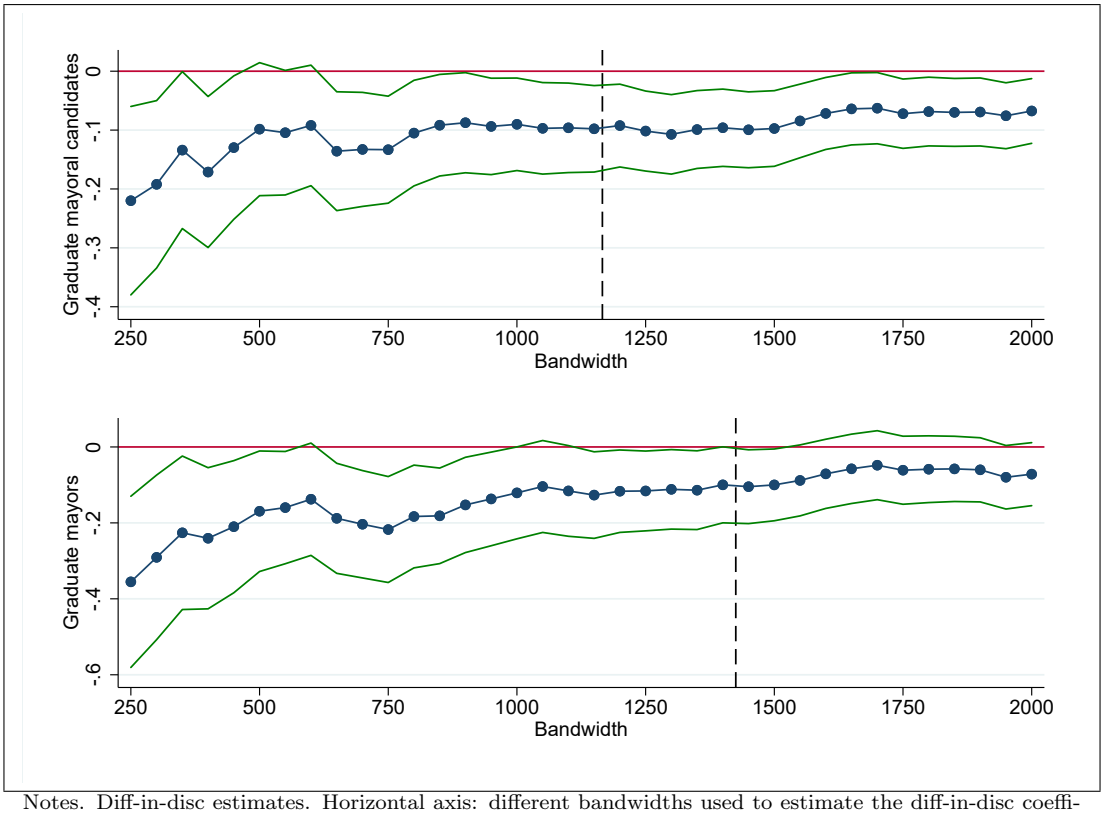
Notes. Municipalities below 15,000 inhabitants. Electoral terms between 2001 and 2012. Treatment variable: Graduate is a dummy variable =1 when the mayor has a university degree, 0 otherwise. Estimation by RDD-LLR using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector. Estimates reported: conventional RD estimates with a conventional variance estimator (Conventional), bias-corrected RD estimates with a conventional variance estimator (Bias-corrected), and bias-corrected RD estimates with a robust variance estimator are reported (Robust). Year of election fixed effects included in even columns. Region fixed effects included in even columns. 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 skills occupation in the past. Municipal covariates in columns 2 and 4: 1) share of population with a university degree measured in 2001; 2) share of active population (i.e. population between 15 and 64 years old) measured in 2001; 3) share of elderly (i.e. population above 65 years old) measured in 2001; 4) log of income per capita measured in 2001; 5) number of firms per capita; 6) number of no-profit associations per capita; 7) area of municipality in square km; 8) population density measured in 2001. Robust standard errors clustered at the local labor area level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Figure A1: Density test on the running variable



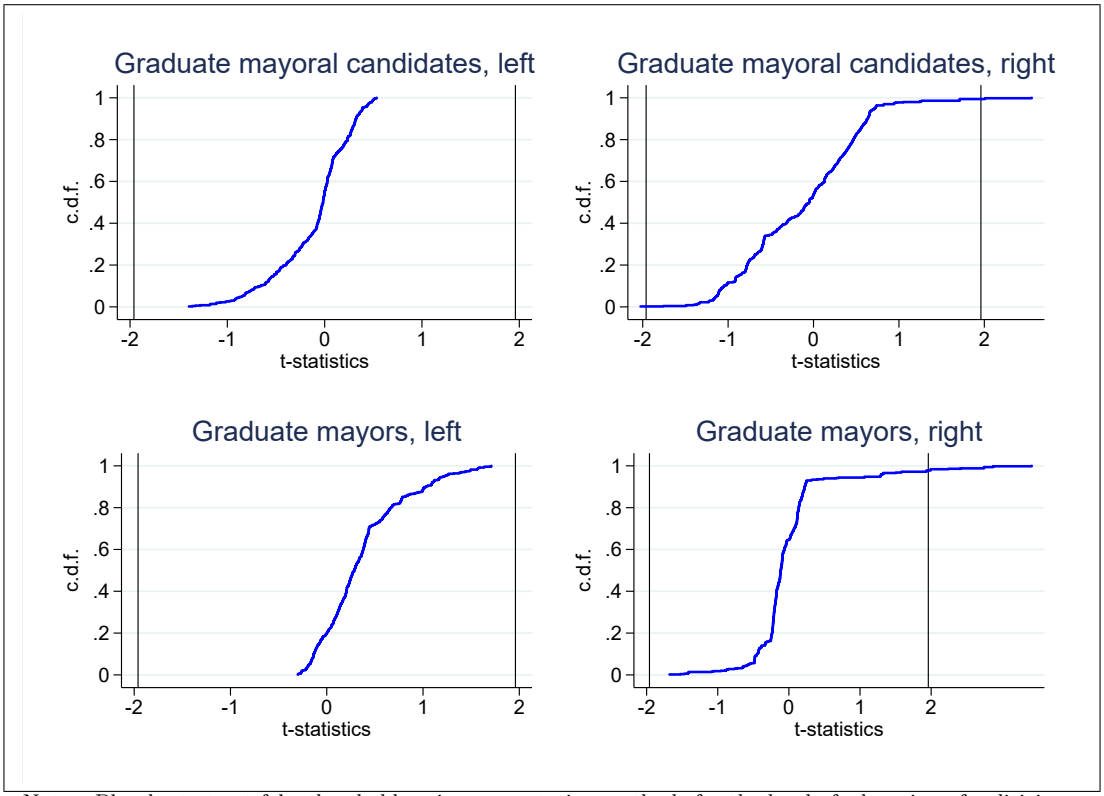
Notes. Discontinuity test for the density of the population at the 5000 inhabitants threshold. Top graphs: (1) density test for R_{it} before 2001; (2) density test for R_{it} from 2001. Bottom graph: (1) discontinuity test for the difference between the density of average R_{it} from 2001 and the density 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



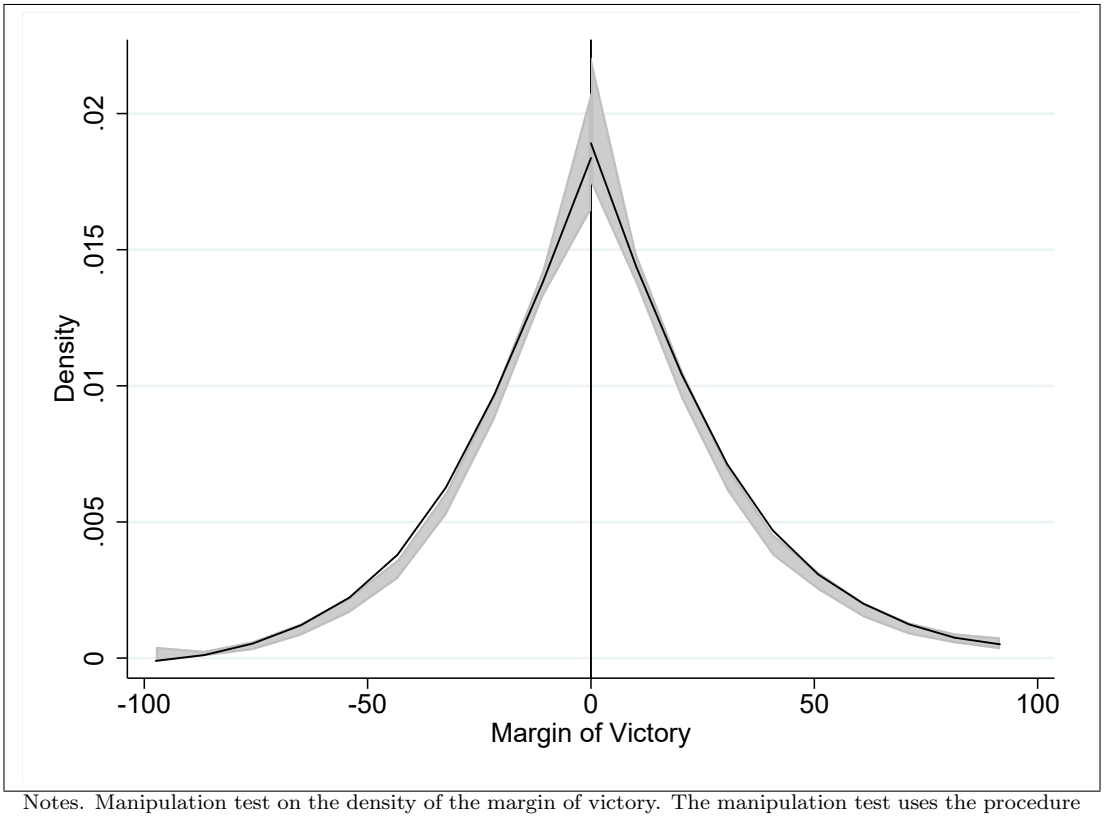
Notes. Diff-in-disc estimates. Horizontal axis: different bandwidths used to estimate the diff-in-disc coefficients. Vertical axis: diff-in-disc coefficients. Dashed vertical line: optimal bandwidth calculated using the Calonico, Cattaneo and Titiunik (2014) and Calonico, Cattaneo and Farrell (2018) MSE-optimal bandwidth h selector.

Figure A3: Diff-in-Disc
Placebo thresholds



Notes. Placebo tests at fake thresholds using permutation methods for the level of education of politicians. The figure reports the c.d.f. of the t-statistics of a set of diff-in-disc regressions at 500 fake thresholds below and 500 fake thresholds above the 5000 threshold (i.e. thresholds from 4900 to 4400, and from 5100 to 5600). The diff-in-disc model is run using a local linear regression. The vertical lines indicate t-statistics of -2 and 2. The top graphs report the c.d.f. of the t-statistics for the share of mayoral candidates with a university degree (respectively to the left and to the right of the 5000 threshold). The bottom graphs report the c.d.f. of the t-statistics for the share of mayors with a university degree (respectively to the left and to the right of the 5000 threshold).

Figure A4: 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 0.500, while the robust one is 0.679.

Figure A5: 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 0.584, while the robust one is 1.244.

B Proofs

Proof of Lemma 1. 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. ■

Proof of Lemma 2. At the voting stage, V anticipates the policy choices outlined in Lemma 1. Suppose both candidates 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 $\phi^H > \phi^L$. ■

Proof of Lemma 3. Start from the H politician. He compares (1) and (2), choosing to enter iff $w^i \leq \gamma^i(E + k\mathbb{E}_{b,\theta,s}u_H^P)$. Given Lemma 1, it is clear that $\mathbb{E}_{b,\theta,s}u_H^P = ((1-\tau)\phi^H + \tau)$. Given Lemma 2, it is clear that $\gamma^i = 1$ irrespective of the choice of the other candidate. Moving to the L politician, the logic on $\mathbb{E}_{b,\theta,s}u_L^P$ is the same. However, he knows he can win office only if H does not run, hence with probability $\left(1 - \frac{\bar{w}_H}{W^H}\right)$. ■

Proof of Lemma 4. Given our assumption that, even in case of FR, the H candidate is chosen whenever he runs, the proof for this Lemma follows the same logic as the proof of Lemma 3. 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. ■

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:

$$
\begin{aligned}
&\hat{\lambda} > \hat{\lambda}^{FR} \\
&\frac{p_L^{FR}}{p_H^{FR}} > \frac{p_L}{p_H} \\
&\frac{\bar{w}_H}{\bar{w}_L} > \frac{\bar{w}_H^{FR}}{\bar{w}_L^{FR}} \\
&\frac{\left(E + ((1-\tau)\phi^H + \tau)k\right)}{\left(1 - \frac{\bar{w}_H}{W^H}\right)\left(E + ((1-\tau)\phi^L + \tau)k\right)} > \frac{(E + (1-\tau)(1-p)k)}{\left(1 - \frac{\bar{w}_H^{FR}}{W^H}\right)\left[E + (1-\tau)(1-p)k\right]} \\
&\frac{\left(1 - \frac{\bar{w}_H^{FR}}{W^H}\right)\left(E + ((1-\tau)\phi^H + \tau)k\right)}{\left(1 - \frac{\bar{w}_H}{W^H}\right)\left(E + ((1-\tau)\phi^L + \tau)k\right)} > 1
\end{aligned}
\tag{B.1}
$$

where the second line follows directly from (5), the third from the definition of p_Γ and the fact that we can cancel out the denominators because of the uniform distribution, and the last one from the equilibrium strategies already defined. Note that the inequality is always verified: first, $E + ((1-\tau)\phi^H + \tau)k > E + ((1-\tau)\phi^L + \tau)k$. Second, $\bar{w}_H > \bar{w}_H^{FR}$ and as a consequence $\left(1 - \frac{\bar{w}_H^{FR}}{W^H}\right) > \left(1 - \frac{\bar{w}_H}{W^H}\right)$ ■

C Financially constrained municipalities

With respect to the baseline model, we add a cost of running deficit. More formally, if $x = 1$ is chosen, then the municipality (including the incumbent politician) faces a cost $c(D_{t-1})$, where D_{t-1} is the level of previous debt. We can interpret c as a form of debt repayment, or as a constraint on the ability to run deficits again, and we assume that $c \in \{0, \bar{c}\}$, where $\bar{c} > 1$ kicks in if D_{t-1} is above a certain threshold. This can be seen as a very high probability of default, or as a very high interest rate necessary to borrow further resources.

This implies that we can divide the set of municipalities into two groups: "ordinary" municipalities (where $c = 0$) and "financially constrained" municipalities (where $c = \bar{c}$). The first subset behaves exactly as in the benchmark model. In financially constrained municipalities, instead, the equilibrium looks different than before.

Lemma C1. *In financially constrained municipalities, the equilibrium policy choice is such that $x = 0$ for every type of politician and every signal realization.*

Proof of Lemma C1. Suppose the incumbent politician is a biased type. By choosing $x = 1$, his payoff is $1 - \bar{c} < 0 = u^P(x = 0)$. Suppose the incumbent politician is an high ability

unbiased type with signal realization $s = 1$. We have

$$\begin{aligned}\mathbb{E}u_H^P(x=1|s=1) &= \frac{p\phi^H}{p\phi^H + (1-p)(1-\phi^H)} - \bar{c} \\ &< 1 - \frac{p\phi^H}{p\phi^H + (1-p)(1-\phi^H)} = \mathbb{E}u_H^P(x=0|s=1)\end{aligned}$$

As this holds for high education politicians receiving the most favourable signal, it holds for every other type and every other signal. ■

This follows from the fact that $\bar{c} > 1$, hence no type of politician has any incentive of choosing $x = 0$. Given lemma C1, it is easy to see that $\bar{w}_H = (E + (1-\tau)(1-p)k) = \bar{w}_H^{FR}$, irrespective of whether FR are in place or not.

Proposition C1. *The share of high education politicians willing to run in financially constrained municipalities is the same with or without fiscal rules.*

Proof of Proposition C1. Anticipating the equilibrium strategies, we have that $\bar{w}_H = \bar{w}_L = (E + (1-\tau)(1-p)k) = \bar{w}_H^{FR} = \bar{w}_L^{FR}$, hence if we substitute in equation (5) we obtain $\hat{\lambda} = \hat{\lambda}^{FR}$ ■

D Discussion on the model

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

D.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 D1. *If $k = 0$, the share of high education politicians is the same with and without fiscal rules.*

Proof of Corollary D1. Substituting $k = 0$ in the LHS of equation (B.1) and on the relevant equations of Lemma 3 and 4, we obtain

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

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

Intuitively, when $k = 0$, FR 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 FR holds if politicians are purely policy motivated and if the winning probability does not enter in their decision. In particular:

Corollary D2. *Assume that $E = 0$ and $k > 0$. In this case, $\hat{\lambda} > \hat{\lambda}^{FR}$.*

Proof of Corollary D2. Substituting $E = 0$ in the LHS of equation (B.1) and on the relevant equations of Lemma 3 and 4, we obtain

$$\frac{\left(1 - \frac{\bar{w}_H^{FR}}{W^H}\right)\left((1-\tau)\phi^H + \tau\right)}{\left(1 - \frac{\bar{w}_H}{W^H}\right)\left((1-\tau)\phi^L + \tau\right)}$$

with $\bar{w}_H = (1-\tau)\phi^L + \tau)k > (1-\tau)(1-p)k = \bar{w}_H^{FR}$. Hence, $\hat{\lambda} > \hat{\lambda}^{FR}$. ■

Corollary D3. *Assume that politicians receive w^i also when he runs and loses. In this case, $\hat{\lambda} > \hat{\lambda}^{FR}$.*

Proof of Corollary D3. 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. However, replacing in the LHS of equation (B.1), 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. ■

D.2 Education and bias

Suppose bias is correlated with education, i.e. we have τ_H and τ_L. We can derive a sufficient condition for our result to hold even in the most extreme case (i.e. $\tau_H = 0$ and $\tau_L = 1$), keeping the condition that H politicians are preferred by V.[1]

Proposition D1. *Assume $\tau_H \neq \tau_L$. If $E \geq \frac{W^H - \phi^H k}{\phi^H + p}$ then $\hat{\lambda} > \hat{\lambda}^{FR}$ for every τ_H, τ_L.*

Proof of Proposition D1. Using (B.1), but noticing that we cannot simplify the RHS as before, we have that $\hat{\lambda} > \hat{\lambda}^{FR}$ iff

$$\frac{\left(E + ((1-\tau_H)\phi^H + \tau_H)k\right)}{\left(1 - \frac{\bar{w}_H}{W^H}\right)\left(E + ((1-\tau_L)\phi^L + \tau_L)k\right)} > \frac{\left(E + (1-\tau_H)(1-p)k\right)}{\left(1 - \frac{\bar{w}_H^{FR}}{W^H}\right)\left[E + (1-\tau_L)(1-p)k\right]} \tag{D.1}$$

[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. $\tau_L > \tau_H \frac{\phi^H - p}{\phi^L - p} - \frac{\phi^H - \phi^L}{\phi^L - p}$.

Note that the LHS of (D.1) is decreasing in τ_L and the RHS is increasing in τ_L. Hence, the condition is least likely to hold if $\tau_L = 1$. Furthermore, the LHS of (D.1) is increasing in τ_H and the RHS is decreasing in τ_H. Hence, the condition is least likely to hold if $\tau_H = 0$. Replacing $\tau_L = 1, \tau_H = 0$ in (D.1), we obtain

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

It is clear that the numerator of the LHS is bigger than the numerator of the RHS. As a consequence, the condition is surely true if $\left(1 - \frac{E+\phi^H k}{W^H}\right)(E+k) \leq \left(1 - \frac{E+(1-p)k}{W^H}\right)E$. Re-arrainging, this simplifies to $E \geq \frac{W^H - \phi^H k}{\phi^H + p}$. ■