

Politicians' Quality and Electoral System. Evidence from a Quasi-experimental Design for Italian Cities

Marco Alberto De Benedetto
Università degli Studi di Messina
Birkbeck College

Abstract

We study the effect of the electoral system on the quality of politicians, as measured by the average years of schooling, at the local level in Italy over the period 1994-2017. Since 1993 municipalities below 15,000 inhabitants vote under a single ballot, whereas cities above 15,000 inhabitants threshold are subject to a double ballot. Exploiting the discontinuous policy change nearby the population cutoff we implement a quasi-experimental design (RDD) and find that runoff elections lead to a decrease in the educational attainment of local politicians within government apparatus by about 2 percent compared to years of schooling of politicians in municipalities voting under a single ballot scheme. The exogenous change in the electoral system negatively affects also mayors and councilors' education, and is present only in those municipalities characterized by a low level of education of the resident population. We speculate that the negative effect is not driven by the higher level of political competition characterizing runoff elections, but on the different selection process of candidates adopted by political parties between runoff and single ballot system. Findings are similar when we measure politicians' quality in terms of previous occupation and when we control for several measures of political closeness.

JEL codes: C31; D72; I20; J42.

Keywords: Regression discontinuity design; Electoral system; Education; Political competition.

1. Introduction

The choice of the electoral system is one of the most crucial institutional decisions for any democracy since it has a profound effect on the future political life of all of countries. For instance, proportional representation and parliamentary systems have been found to lead to higher levels of public expenditures compared to majoritarian and presidential systems (Persson and Tabellini, 2005; Milesi-Ferretti *et al.*, 2002), and under plurality rule two-party systems usually prevail (Bordignon *et al.*, 2016).

The majority of contributions in the literature has focused on the impact of electoral systems on political competitiveness, public policies (see Chamon *et al.*, 2008; Rizzo and Zanardi, 2010; Bracco and Brugnoli, 2012) and turnout (Barone and De Blasio, 2013), even though empirical results are mixed. In particular, on the relationship between electoral rules and political competition, Callander (1999) finds that plurality elections tend to see more candidates running, whereas Cox and Neto (1997), and Chamon *et al.* (2008) highlight how more candidates are observed in runoff elections. The same conclusions are reached by Bordignon *et al.* (2016) for Italian municipalities who find that under runoff elections, the number of political candidates is larger, but the influence of extremist voters on equilibrium policy and hence policy volatility are smaller, because the bargaining power of the political extremes is reduced compared to single round elections.

Hence, these works confirm that in cities affected by double ballot more candidates compete at the electoral race and a greater number of parties are involved, compared to cities voting under a single ballot rule, also at the local level in Italy (Bordignon *et al.*, 2016), leading in turn to an increase in the level of political competition.

The aim of our paper is to provide evidence on the effect of the electoral system on the quality of politicians using data for Italian Municipal administrations over the period 1994-2017. DL 25 March 1993, no. 81 has introduced a different electoral scheme based on population size: municipalities below 15,000 inhabitants vote under a single ballot, whereas cities above 15,000 inhabitants threshold are subject to a double ballot. The Italian local institutional framework allows us to exploit a sharp change in the electoral rule and to implement a quasi-experimental design, namely a regression discontinuity design (RDD). We focus on a narrow sample of municipalities (Local Linear Regression) nearby the cutoff and compare cities just below and just above the threshold. Those municipalities are assumed to be comparable in all their baseline characteristics (we discuss this assumption later in the paper). This implies that municipalities adopting a single ballot provide a valid counterfactual for municipalities voting under runoff and by comparing these outcomes we identify the causal effect of interest. To the best of our knowledge this is the first contribution that explicitly identifies a causal link between electoral rules and politicians' quality, as measured by the average years of schooling.

Our results show a causal negative relationship between the change in the electoral rule and politicians' quality. In fact, in cities affected by a double ballot system, the educational attainment of local politicians decreases on average by about 2 percent compared to local politicians in cities just below the threshold. Findings tend to be the same both when we measure the quality of politicians in terms of their former job/employment status, and when different specifications of our model (polynomials of the forcing variable higher than one and different optimal bandwidths) are used.

The main explanation driving our results relies on the fact that in municipalities below 15,000 inhabitants candidates running for a mayor position are supported only by one list (more homogeneous candidates in terms of education), whereas in cities above the cutoff candidates may be supported by one or more lists. In the latter municipalities, political parties assemble more than one list (with more heterogeneous candidates in terms of education in each of them) exclusively to get more votes, allowing their candidates to win the electoral race and become mayors: the more candidates in each list, the higher the chance that a candidate supported by parties becomes mayor. Moreover, voters are attached more to ideology and party affiliation rather than to candidates' characteristics especially in large cities (we should keep in mind that only 9% of municipalities have a population size larger than 15,000 inhabitants and in small municipalities most of parties do not belong to any right/left wings), and in turn cast their vote in the first round for the political party that represents their expectations the most. Only during the second round, when parties cannot receive votes anymore, voters are likely to focus on the two best leading candidates' features to choose the winner of the electoral competition.

Furthermore, it might be the case that electors vote one of the lists supporting mayor candidates based not on candidates' characteristics (including the educational attainment) but on their potential connections (family ties, friendship, and so on) with candidates. This channel especially works in low information environments where having a high number of candidates or parties running for elections leads more confused electors to cast their votes not considering candidates' skills. We

formally test this channel by splitting the sample based on the median value of population educational attainment, finding a negative effect of the runoff on the local politicians 'quality only in those cities with a low level of education (below the median).

Related to this argument, it might be that political parties rather than selecting high-skilled candidates, choose those who are more able to attract electors to cast their vote in their favor at polls. In fact, Caselli and Morelli (2004) and Messner and Polborn (2004), by considering citizen-candidate models, find that in equilibrium low-quality individuals may be more likely to run for office than high-quality ones, since the value of the outside opportunities of high-quality individuals exceed the rewards from office, thus discouraging them from running for public office. Moreover, by taking into account the role of parties in the selection of politicians, Mattozzi and Merlo (2010) find that parties may find it optimal to attract low-quality politicians, in order to keep the overall level of party service sufficiently high. This scenario is more likely to happen in cities voting under runoff rather than in electoral races held under a single ballot.

No other potential channels are left in explaining the negative effect of the reform on the educational attainment of local politicians. Since in the literature, it has been shown that runoff elections are characterized by a larger number of candidates and political parties compared to single ballot elections (the number of candidates and of parties sharply change in the neighborhood of the 15,000 inhabitants cutoff), as a further robustness check we implement a Fuzzy RDD with optimal bandwidth where the number of candidates and political parties are instrumented by using the 15,000 threshold as instrument and we control for province and year fixed effects. We do not find a statistically significant effect of these two measures of political closeness on the educational attainment of local politicians.

Our paper also relates to the literature focusing on the debate about the selection of good politicians for enhancing adequate policies and improving economic performance (see among others, Besley, Persson, and Sturm, 2007; Jones and Olken, 2005; Besley, 2007). In particular, Besley and Coate (1997) and Osborne and Slivinski (1996) provide a theoretical framework, in which political competition is represented as a game between citizens competing to hold public office, and show the relevance of politicians' identity for policy choices. Following this approach, Pande (2003) and Chattopadhyay and Duflo (2004) investigate the impact of caste and gender reservations on policy choices. Other papers focus, instead, on the effects of politicians' education on economic outcomes. Jones and Olken (2005) and Besley, Montalvo and Reynal-Querol (2011) look at the relationship between politicians' quality and growth and find that growth is higher when leaders are more educated. Gagliarducci and Nannicini (2011) and Ferraz and Finan (2009), investigating the effects of remuneration on the quality and productivity of politicians, show that better paid politicians perform better, mainly due to the selection of more skilled candidates. Finally, Baltrunaite *et al.* (2014) and Daniele and Geys (2015) look at the causal effect of gender quotas and organized crime on local politicians' educational attainment respectively. We add to this huge literature, investigating the effect produced by an exogenous change in the electoral rule on politicians' quality.

The paper is structured as follows. In Section 2 we describe the Italian local institutional framework and the datasets used in the empirical analysis. Section 3 describes the methodology design and its validity. Section 4 and 5 present the main results, whereas in Section 6 some robustness checks are considered. Section 7 concludes.

2. Institutional setting and data

2.1. Municipal Elections in Italy

In Italy, municipal administrations are involved in a number of important functions, such as the management of public utilities (local roads, water, sewage and garbage collection), the provision of public housing, transportation and nursery schools, and the assistance of elderly people. The Municipal Council (*Consiglio Comunale*) is endowed with legislative powers, while the executive authority is assigned to a Mayor (*Sindaco*) heading an Executive Committee (*Giunta Comunale*). Since the services provided by Municipalities have a great impact on citizens' daily lives, voters are generally very interested both in the composition as well as in the performance of Municipal apparatus.

The system currently regulating municipal elections in Italy has been introduced in 1993 (DL 25 March 1993, no. 81). It has established the direct election of the mayor with some differences according to the size of the city. Since 1993, mayors have been subject to a two-term limit, while members of the Executive Committee and of the Municipal Council, endowed with legislative power, can be re-elected indefinitely. Municipal elections in Italy are held every 5 years¹ and municipal governments cannot choose the election schedule. In certain circumstances, the legislature may not survive until the end of its legislative term, e.g. because of a mayor's early resignation. In these cases, elections are held before the natural schedule, and, as a consequence all subsequent elections will be held at different times from other municipalities that have completed the foreseen legislative term. Municipalities have a registry of eligible voters, which is revised whenever there is an election and all citizens aged 18 or above on the election date are automatically registered to vote. Voting takes place in polling stations organized by the local authorities. Elections are organized according to a traditional paper ballot system.

Moreover, in municipalities with a population size fewer than 15,000 inhabitants elections are held with single-ballot and majority rule: the candidate with the highest number of votes is elected (relative majority). Whoever has obtained the highest number of votes becomes a mayor and the connected list wins 2/3 of the council seats; the remaining third is divided among the other parties or coalitions based on the percentages obtained. The first seat of the majority lists belongs to candidates running for a mayor position who are defeated. In the event of parity in terms of votes received by the two leading candidates running for a mayor position, the ballot will be held on the second following Sunday. In case of further parity the oldest candidate is elected.

Conversely, in municipalities with more than 15,000 inhabitants, elections are held with double ballot and majority rule: if in the first round none of the candidates obtains an absolute majority (50% + 1 of the votes), the second ballot is held on the following second Sunday between the two most voted candidates. Each candidate can be connected to *one or more lists* and electors can express two choices: one for the mayor and one for the list of councilors. It is possible to choose a different list from the one related to the candidate to whom elector has expressed his vote (separate vote). In the intermediate phase between the first and second round the two candidates who go to the ballot can link to lists other than those who had supported them in the first round. In the second

¹ With the exception of the years between 1993 and 1999, when the electoral mandate had a duration of 4 years.

round, electors can only express a vote for candidates running for a mayor position, while votes on the lists are not allowed.

As far as the election of councilors within the municipal apparatus is concerned, in municipalities with a population size less than 15,000 inhabitants each mayor candidate is connected to *only one* list and then when a candidate is elected, automatically the list of councilors connected to him/her is chosen; each list must be composed by a number of candidates not higher than the number of councilors to be elected by law and not less than 3/4 of the total. Voters can express a preference among councilors of the list connected to the chosen candidate, writing the surname under the list. Whereas, in municipalities with more than 15,000 inhabitants the lists must include a number of candidates not exceeding the number of councilors to be elected by law and not less than two thirds of the total. Three different situations can arise for the election of the council: 1) the mayor is elected in the first round, but the list or linked lists do not reach 50% of the votes: in this case the seats are distributed according to the proportional d'Hondt rule and it is possible that the board is composed for the majority by members of the opposition in the event that the list or the lists of one of the defeated candidates have obtained more than 50% of the votes; 2) the list or the lists connected to the candidate elected in the first round reaches 40% of the votes and no other list exceeds 50%. In this case the list attached to the elected candidate is awarded the majority prize (60% of the seats). The remaining seats are assigned with the d'Hondt method; 3) the mayor is elected at the ballot, but the supporting list gets already in the first round 40% of the votes and none of the other lists has reached 50%: even in this case the list is awarded the majority prize (60%) and the rest of the seats is assigned with the proportional method (d'Hondt system). Lists that have not reached the 3% threshold of votes are not admitted to the distribution of seats.

Summing up, it is clear that in municipalities with less than 15,000 thousand inhabitants the elected mayor will certainly have a majority in council; this may not happen in municipalities with more than 15,000 inhabitants, where the mayor can be with the majority of the council in hand of opposition parties.

2.2 Data description

In our analysis we use three different sources of data: first, an administrative dataset provided by the Ministry of the Interior containing detailed information on gender, age and education level of all politicians elected in Italian municipalities from 1985 to 2017; second, an electoral dataset provided by the Italian Ministry of the Internal Affairs having information for all municipal elections on the number of voters and the number of people eligible to vote over the period 1994-2017; and third, 1991, 2001 and 2011 Italian Census of Population. Since the electoral reform introducing the double ballot came into force in 1993, we restrict our analysis over the period 1994-2017 for first round elections, and we end up with a sample of 4,165 observations for 1,173 municipalities.

Table 1 reports the descriptive statistics of the variables used in our empirical exercise. We build our main outcome variable *Politicians Education* (with a mean of 14) as the average educational attainment of all members within local municipalities at the time of elections. The use of politicians' educational level follows previous work by, among others, De Paola and Scoppa (2010), and Baltrunaite et al. (2014), where human capital is considered as a valid proxy for individuals' ability and skills. The average education level of local politicians in a given municipality and year is calculated on an annual basis translating degrees into the minimum number of years necessary to

obtain them². Further, given the detailed information of the dataset, we are also able to distinguish between the educational level of candidates running for a mayor position, i.e. *Mayor Education* (with a mean of 15.38 and a standard deviation of 3.08), the educational attainment of councilors, namely *Councilors Education* (with a mean of 13.76 and a standard deviation of 1.31), and of members of the Executive Committee, i.e. *Aldermen Education* (with a mean of 14.21 and a standard deviation of 2.09). Moreover, we also distinguish between the average years of schooling of male and female local politicians, namely *Male Education* and *Female Education*, with a mean of 13.81 and 14.89 respectively.

Table 1. Descriptive Statistics for the Discontinuity Sample

| Variables | Mean | S.D. | Min | Max | Observations ³ |
|-----------------------------------|---------|---------|--------|---------|---------------------------|
| Politicians Education | 14.0018 | 1.3771 | 8 | 18 | 4,165 |
| Mayor Education | 15.3808 | 3.0761 | 5 | 18 | 3,913 |
| Councilors Education | 13.7592 | 1.3066 | 8 | 18 | 3,985 |
| Aldermen Education | 14.2099 | 2.0912 | 7 | 18 | 3,346 |
| Male Education | 13.8084 | 1.4570 | 8 | 18 | 4,149 |
| Female Education | 14.8865 | 2.1852 | 5 | 18 | 3,701 |
| Proportion of Female Politicians | 0.1785 | 0.1324 | 0 | 1 | 4,165 |
| Incumbent | 0.5757 | 0.4929 | 0 | 1 | 4,165 |
| Politicians' Age | 44.2797 | 3.9833 | 30 | 71 | 4,165 |
| Female Politicians | 0.9083 | 0.2886 | 0 | 1 | 4,165 |
| Population Size/1,000 | 12.372 | 3.9431 | 7.553 | 22.389 | 4,165 |
| Voter Turnout | 0.7751 | 0.0787 | 0.4562 | 0.9253 | 4,165 |
| Population Education | 7.7307 | 1.0171 | 4.9863 | 11.7465 | 4,165 |
| Employment | 0.3320 | 0.1302 | 0.0549 | 0.8298 | 4,165 |
| Area (in km2) | 50.0809 | 53.9059 | 1.62 | 551 | 4,165 |
| Candidates Electoral Margin | 0.1371 | 0.1134 | 0 | 0.7144 | 4,152 |
| Candidates Electoral Margin (t-1) | 0.1366 | 0.1147 | 0 | 0.7144 | 3,454 |
| Parties Electoral Margin | 0.1211 | 0.1098 | 0 | 0.7144 | 3,338 |
| Parties Electoral Margin (t-1) | 0.1241 | 0.1117 | 0 | 0.7144 | 2,718 |
| # Parties | 5.4532 | 4.0046 | 1 | 36 | 3,347 |
| # Parties (t-1) | 5.1194 | 3.6787 | 1 | 36 | 2,729 |
| # Candidates | 3.7377 | 1.3171 | 2 | 16 | 4,152 |
| # Candidates (t-1) | 3.7423 | 1.3521 | 2 | 16 | 3,454 |

Source: Local Administrators Data set (1985-2011), Italian Ministry of Internal Affairs; Italian Census of Population (1991- 2001-2011)

Using information about past political experience of local administrators, we build a dummy variable *Incumbent* taking the value of 1 when local politicians have been in charge for a mayor position during previous legislature and zero otherwise. From Table 1, we can notice that in 57% of municipal bodies there is a politician who has already performed the charge of mayor in the previous legislature. Furthermore, exploiting information on candidates' gender we build a dummy variable *Female Politicians* taking the value of 1 when there is at least one woman among local politicians: the proportion of municipal bodies in which there is at least a woman is about 91% with

² In Italy, it takes 5 years to take a primary diploma, 8 years for a Middle-school diploma, 13 years to attain a High-School Degree while 17-18 years are necessary to attain a College Degree. Moreover, the educational attainment of people with a PhD or a Master degree is always 18 years in our sample.

³ Observations refer to the specification in which we regress the average educational attainment of politicians within municipal apparatus on all of covariates, including province-year fixed effect and using an optimal bandwidth as proposed by Imbens and Kalyanaraman (2012).

a standard deviation of 0.29, whereas the proportion of female politicians within municipal bodies is about 18 percent. Finally, local politicians are on average 44 years old.

As far as electoral characteristics are concerned, we measure *Voter Turnout* as the ratio between the number of voters and the number of eligible voters in the first round. As shown in Table 1 Italy is characterized by a quite high electoral turnout compared to many European countries and to US: the average turnout in our sample over the period 1994-2017 has been of 77.5%, with a standard deviation of 0.08. This dataset allows us also to use some proxies of the degree of competition characterizing each electoral race. We have information on the number of votes obtained by each candidate/party and we create the variables *Candidates Electoral Margin* and *Parties Electoral Margin* as the absolute difference between the votes obtained by the two leading candidates/parties (divided by the number of eligible voters). These two proxies represent an inverse measure of expected electoral closeness and are on average equal to 0.137 and 0.136 respectively⁴. The second proxy of the degree of political competition is simply the number of candidates/parties competing at the electoral race (*No. Candidates* with an average of 4 and a standard deviation of 1.43, and *No. Parties* with a mean of 5.45 and a standard deviation of 4). Moreover, we use the margin by which the incumbent candidate/party won the previous election and the lagged value of the number of candidates/parties as further indicators of competitiveness.

Finally, to control for municipalities' demographic characteristics, we use the 1991, 2001 and 2011 Italian Census of Population. Data from the 1991 census are used for elections held in the period 1994-2000, data from the 2001 census are used for elections held between 2001 and 2010, while data from 2011 census are used for elections held over 2011-2017 period. We have information on the size of resident population, the average level of employment, the educational attainment of the population and the municipal area in Km². As shown in Table 1, the average population size in our sample is 12,372, the average educational attainment of population, by considering only people aged 6 or above, is about 8 years, whereas the fraction of employed people in the population is 33%.

3. Empirical framework and validity of the RDD

3.1 Methodology

To recover the causal effect of the electoral system on local politicians' average years of schooling, we use a quasi-experimental design and implement a sharp Regression Discontinuity Design (RDD), since the local electoral system (single ballot vs runoff) exogenously change around the 15,000 inhabitants cutoff. The probability that a municipality has a population just above or just below the threshold is understood to be random, and helps us to overcome the presence of confounding and unobservable factors that affect local politicians' educational attainment. This way, we are able to compare cities adopting a single ballot system (below the 15,000 inhabitants threshold) to municipalities using a double ballot (just above the same cutoff). We decide to estimate a Local Linear Regression (LLR henceforth), by means of OLS with fixed effects at province level, controlling for a linear polynomial of the normalized forcing variable, i.e. population size, along with an interaction term of the first order between the treatment and the

⁴ The value of zero characterizes few elections in which the two leading candidates/parties obtained exactly the same number of votes.

assignment variable, and choosing an optimal bandwidth as suggested by Imbens and Kalyanaraman (2012). The basic econometric specification is therefore the following:

$$Politicians_Education_{it} = \beta_0 + \beta_1 Runoff_{it} + \beta_2 Z_{it} + \beta_3 X_{it} + \beta_4 P_{it} + \varphi_i + \mu_t + \varepsilon_{it}, \quad [1]$$

where $Politicians_Education_{it}$ is a variable measuring the educational attainment of local politicians, as measured by the average years of schooling (De Paola and Scoppa, 2010; Baltrunaite et al., 2014; Daniele and Geys, 2015), whereas P_{it} is a linear polynomial of the normalized assignment variable (population size minus 15,000)⁵. Furthermore, Z_{it} is a vector of controls for local politicians' characteristics. In particular, it contains $Incumbent_{it}$ that takes a value equal to 1 if a local politician has performed the mayor charge during the previous legislature and zero otherwise, the average age of local politicians, and a dummy variable taking the value of 1 when there is at least a woman within municipal bodies. X_{it} is a vector which includes controls for municipal characteristics at the time of elections (the average educational attainment of inhabitants, the proportion of employed people in the population, the municipal area and the electoral participation at the polls), φ_i and μ_t are respectively a province and a year fixed effect, whereas ε_{it} is the stochastic error term of the model. The fixed effects φ_i accounts for time-invariant characteristics of the province, either observable or unobservable.

In all regressions standard errors are robust to heteroskedasticity and are clustered at the municipal level to take into account the fact that voters' behavior in the same municipality may be affected by common shocks.

3.2 Regression discontinuity design validity checks

It is standard in the RD design to demonstrate that treatment and control groups are similar in their observed baseline covariates. To check whether the assumptions of the RD are satisfied, we present a test of the continuity of the distribution of the covariates added in the main specification at the cut-point. The idea behind this kind of test is to regress a covariate on a first or second order polynomial of the forcing variable along with a dummy for the treatment status: a statistically insignificant coefficient for the treatment dummy is taken as evidence in favor of local random assignment (Caughey and Sekhon, 2011; Lee, 2008; Lee, Moretti and Butler, 2004).

In Table 2 we test whether the runoff is predictive of a larger set of municipal and candidates' characteristics, by choosing an optimal bandwidth as suggested by Imbens and Kalyanaraman (2012) in the neighborhood of the 15,000 inhabitants threshold, and by controlling both for a quadratic polynomial of the forcing variable as well as for a second order interaction between the assignment variable and the treatment status, and for province-year fixed effects. Moreover, standard errors are robust to heteroskedasticity and are clustered at the municipal level. Overall, Table 2 shows that the change in the local electoral rule does not predict any of the predetermined characteristics.

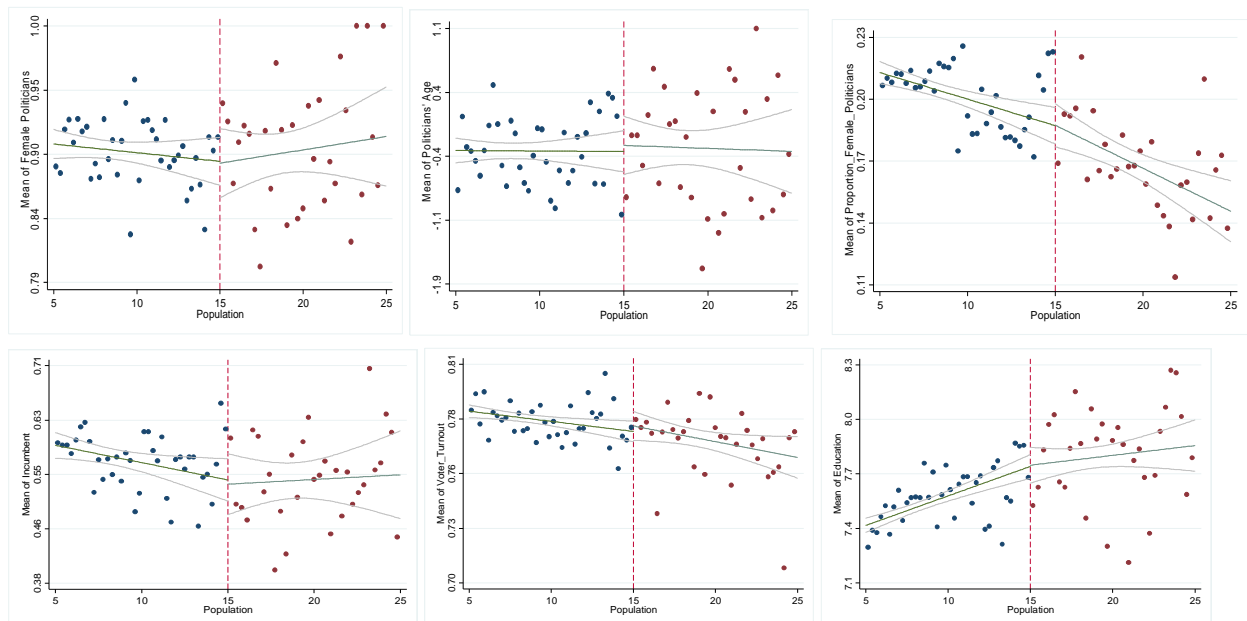
⁵ In the main specification we use a linear polynomial of the forcing variable. As reported in the robustness checks section, we also run a LLR by means of OLS (with province-year fixed effect) with polynomials of population size higher than one.

Table 2: Balance Test on Control Variables

| Variables | Optimal Bandwidth | Obs. |
|----------------------------------|---------------------|-------|
| Politicians 'Age | -0.5076 (0.4098) | 4,165 |
| Female Politicians | 0.0445 (0.0319) | 4,165 |
| Proportion of Female Politicians | -0.0089 (0.0108) | 4,165 |
| Incumbent | 0.0063 (0.0503) | 4,165 |
| Voter Turnout | 0.0011 (0.0058) | 4,165 |
| Population Education | 0.0059 (0.0752) | 4,165 |
| Employment | 0.0097 (0.0135) | 4,165 |
| Area (in Km2) | 3.0719 (6.6248) | 4,165 |

Note: Municipalities between 7,550 and 22,450 inhabitants. Baseline RDD estimates. We control for province and year fixed effects and we focus on the period 1994-2017. Estimation method: LLR with a second order polynomial of the forcing along with an interaction of the second order between the treatment status and the assignment variable. Standard Errors are robust to heteroscedasticity and are cluster at municipal level (reported inside the brackets). Significance at the 10% level is represented by *, at the 5% level by **, and at 1% level by ***.

As a further test, we present in Figure 1, some of the predetermined characteristics, as mentioned above, plotted against population size nearby the 15,000 inhabitants cutoff. Each figure depicts the assignment variable cell means of the predetermined characteristics in the proximity of the population threshold, along with the fitted values of a locally weighted regression which is calculated within each segment. Overall, Figure 1 confirms results highlighted in Table 2.



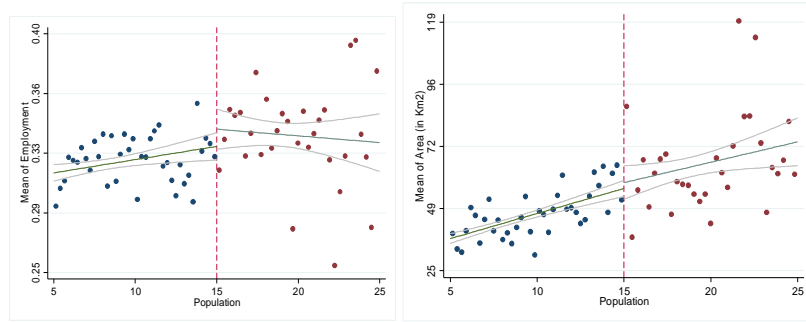


Figure 1: Discontinuity in the predetermined characteristics at the 15,000 cutoff

As a last specification test of our design, we check the continuity of the forcing variable, i.e. population size nearby the cutoff of 15,000 inhabitants performing a McCrary test, by running a kernel local linear regressions of the log of the density separately on both sides of the threshold (McCrary, 2008). If there were any discontinuities at the cutoff point, one might be concerned that mayors are able to manipulate the assignment variable and sort below/above the threshold in order to pass from an electoral system to the other, or in other words if units of observation have a great deal of control over the assignment variable and if there is a perceived benefit to a treatment, one would certainly expect units of observation on one side of the threshold (single ballot municipalities) to be systematically different from those on the other side (double ballot municipalities)⁶. As we can see in Figure 2, the log of the frequency of population size to the right and to the left of the threshold is not statistically significant at conventional levels, considering municipalities with a population size lower than 30,000 inhabitants (it is equal to 0.1895 with a standard deviation of 0.1178).

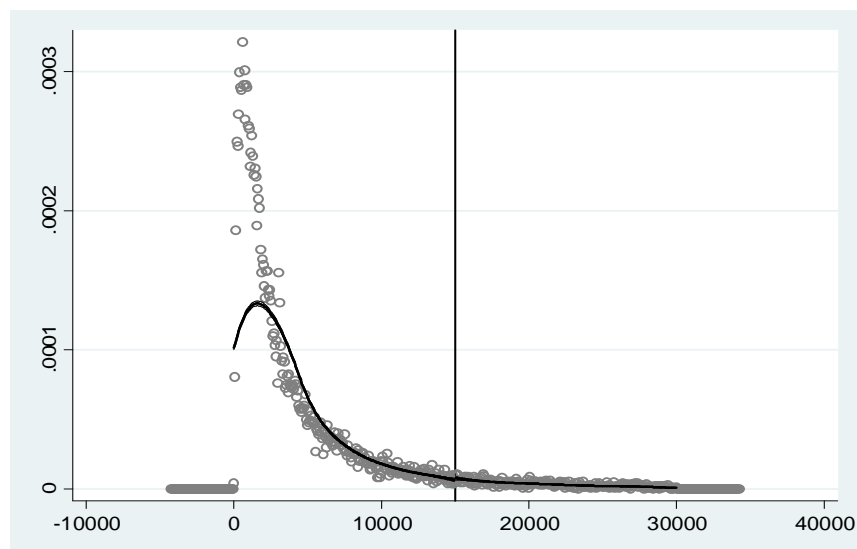


Figure 2: McCrary test. Manipulation of the forcing variable at the 15,000 cutoff

⁶ However, in principle, it is very hard in Italy to manipulate population size at municipal level, since Census is run independently by the National Statistical Office, so that false reporting should be ruled out.

4. Main results

Table 3 reports the main Sharp RDD results where an optimal bandwidth, as proposed by Imbens and Kalyanaraman (2012) (IK hereafter), is adopted. In column (1) we only control for province and year fixed effects, whereas in columns (2) we add a linear interaction term between the treatment status and the normalized forcing variable that is population size. In columns (3) and (4) we control for candidates' features and for municipal characteristics respectively.

In all specifications we use a linear polynomial of the forcing variable and standard errors are robust to heteroscedasticity and clustered at municipal level to take into account potential common shocks affecting municipalities.

Table 3. Effect of Runoff Elections on Average Education. Sharp RDD. LLR Results

| VARIABLES | (1) | (2) | (3) | (4) | (5) |
|--|----------------------|----------------------|-----------------------|------------------------|------------------------|
| | Politicians' Educ | Politicians' Educ | Politicians' Educ | Politicians' Educ | High Skill |
| Runoff | -0.206** (0.0978) | -0.241** (0.0998) | -0.247** (0.0995) | -0.293*** (0.0932) | -0.0260** (0.0128) |
| Politicians' Age | | | -0.0195** (0.0096) | -0.0370*** (0.0097) | -0.0051*** (0.0014) |
| Female Politicians | | | -0.356*** (0.115) | -0.383*** (0.113) | -0.0076 (0.0164) |
| Incumbent | | | -0.0708* (0.0370) | -0.0735** (0.0356) | 0.0133** (0.0052) |
| Voter Turnout | | | | -0.676 (0.456) | -0.160** (0.0631) |
| Employment | | | | -0.115 (0.259) | 0.0695* (0.0385) |
| Area | | | | 0.0016*** (0.0005) | 0.0001 (0.0001) |
| Education Pop | | | | 0.713*** (0.0547) | 0.0136* (0.0077) |
| Constant | 14.41*** (0.343) | 14.36*** (0.353) | 15.64*** (0.583) | 10.74*** (0.740) | 0.213** (0.0987) |
| Population Polynomial Interaction Term | First No | First First | First First | First First | First First |
| Province and year FE | Yes | Yes | Yes | Yes | Yes |
| Optimal Bandwidth | IK | IK | IK | IK | IK |
| Observations | 4,169 | 4,169 | 4,169 | 4,165 | 4,222 |
| R-squared | 0.276 | 0.277 | 0.284 | 0.333 | 0.574 |

Note: Municipalities between 7,550 and 22,450 inhabitants. Sharp RDD estimates. We control for province and year fixed effects and we focus on the period 1994-2017. Estimation method: LLR with a first order polynomial of the forcing variable and interaction term. Standard Errors are robust to heteroscedasticity and are cluster at municipal level (reported inside the brackets). Significance at the 10% level is represented by *, at the 5% level by **, and at 1% level by ***.

Results highlight how municipalities voting under runoff are characterized by a lower quality of local politicians, as measured by the average educational attainment, compared to cities using a single ballot scheme (columns 1-4). In fact, the change in the electoral system leads to a reduction in the level of education of local politicians by 0.293 years (see column 4). The coefficient of interest is statistically significant and stable across specifications. This corresponds to a decrease of 2.09 percent with respect to the average value of 14 years of education in the 7,550-22,450 population window.

One channel through which runoff scheme reduces the educational attainment of local politicians is related to the features of the double ballot system *per se*, since in cities above 15,000 cutoff, candidates running for a mayor position are usually supported by more than one list, whereas in single ballot elections candidates are affiliated to one list only⁷. In fact, in elections held under runoff, where more lists support candidates running for a mayor position, the selection of candidates by political parties might not be related to their intrinsic characteristics, as experience or educational attainment, but to some other features, such as family ties or connections, in order to attract electors to cast their vote in their favor at polls.

Moreover, voters' ideology and affiliation to parties assume a greater importance compared to candidates' skills in large municipalities⁸ and especially in cities above 15,000 voting under runoff. This usually happens in the first round, whereas in the second round voters cannot cast any vote to parties and rely on candidates' ability, such as the educational attainment, to choose their best candidate in the electoral competition.

In column (5) we look at a different measure of politicians' quality, based on the previous occupation. We consider the proportion of politicians in the government apparatus who were engaged in entrepreneurial and professional activities before obtaining a seat in a municipal council, and we build the variable *High Skilled* with a mean of 0.3543 within the 7,550-22,450 population range, and a standard deviation of 0.2392. A higher proportion of politicians whose previous occupation is skill-intensive would be interpreted as an indicator of a higher quality of the political body. We find that elections held under runoff rule lead to a lower proportion of high-skilled local politicians by 2.6 percentage points.

One potential concern with estimates presented in Table 3 might be that the IK optimal bandwidth, a $-/+49\%$ above and below the threshold, is quite wide and includes elections that might be reasonably considered far from close. To address this potential issue we re-estimate specifications reported in Table 3 after narrowing the bandwidth to half the optimal. This alternative method leads to an optimal bandwidth of 3,725 inhabitants above and below the cutoff of 15,000. Results are displayed in Table 4. The impact of runoff is still negative, stable across specifications, statistically significant at the 5 percent level, and similar in terms of magnitude to that presented in the previous table. The educational attainment of local politicians and the proportion of high quality administrators reduce under runoff by 0.26 years and 3.6 percentage points respectively (see columns 4 and 5).

⁷ The heterogeneity of local politicians in terms of average years of schooling, as measured by the standard deviation of the average educational attainment, within the government apparatus is positively affected by runoff elections. The coefficient of the treatment variable is 0.1209 (p-value: 0.008).

⁸ In small municipalities the proportion of *liste civiche*, i.e. parties that do not belong to either left or right wing, is very high. In particular, for cities with a population size lower than 5,000 it is allowed to have *liste civiche* only. In these municipalities voters are more attached to candidates' characteristics rather than ideology or party affiliation.

Table 4. Effect of Runoff Elections on Average Education. Sharp RDD. LLR Results

| | (1) | (2) | (3) | (4) | (5) |
|-----------------------|---------------------|---------------------|---------------------|---------------------|-----------------------|
| VARIABLES | Politicians' Educ | Politicians' Educ | Politicians' Educ | Politicians' Educ | High Skill |
| Runoff | -0.298** (0.126) | -0.296** (0.126) | -0.298** (0.125) | -0.266** (0.118) | -0.0362** (0.0177) |
| Constant | 14.35*** (0.377) | 14.36*** (0.377) | 16.20*** (0.801) | 11.94*** (1.095) | 0.304** (0.146) |
| Province and year FE | Yes | Yes | Yes | Yes | Yes |
| Controls | No | No | Politicians | Demographic | All |
| Population Polynomial | First | First | First | First | First |
| Interaction Term | No | First | First | First | First |
| Optimal Bandwidth | Half IK | Half IK | Half IK | Half IK | Half IK |
| Observations | 1,692 | 1,692 | 1,692 | 1,688 | 1,713 |
| R-squared | 0.321 | 0.321 | 0.328 | 0.362 | 0.593 |

Note: Municipalities between 11,275 and 18,725 inhabitants. Sharp RDD estimates. We control for province and year fixed effects and we focus on the period 1994-2017. Estimation method: LLR with a first order polynomial of the forcing variable and interaction term. Standard Errors are robust to heteroscedasticity and are cluster at municipal level (reported inside the brackets). Significance at the 10% level is represented by *, at the 5% level by **, and at 1% level by ***.

Sharp RDD results are also shown in Figure 3. In particular, we plot the mean of our outcome variables against the 15,000 population threshold, with an half IK optimal bandwidth, by estimating a model that assumes a linear relationship between our dependent variables and the forcing variable. The connected points are the predicted values from a linear regression of our different measures of politicians' quality on population size. In particular, in the left panel we focus on the average educational attainment of all of local politicians, whereas in the right panel the focus is on the proportion of high skilled politicians. Overall, these descriptive graphs confirm results displayed in Table 4.

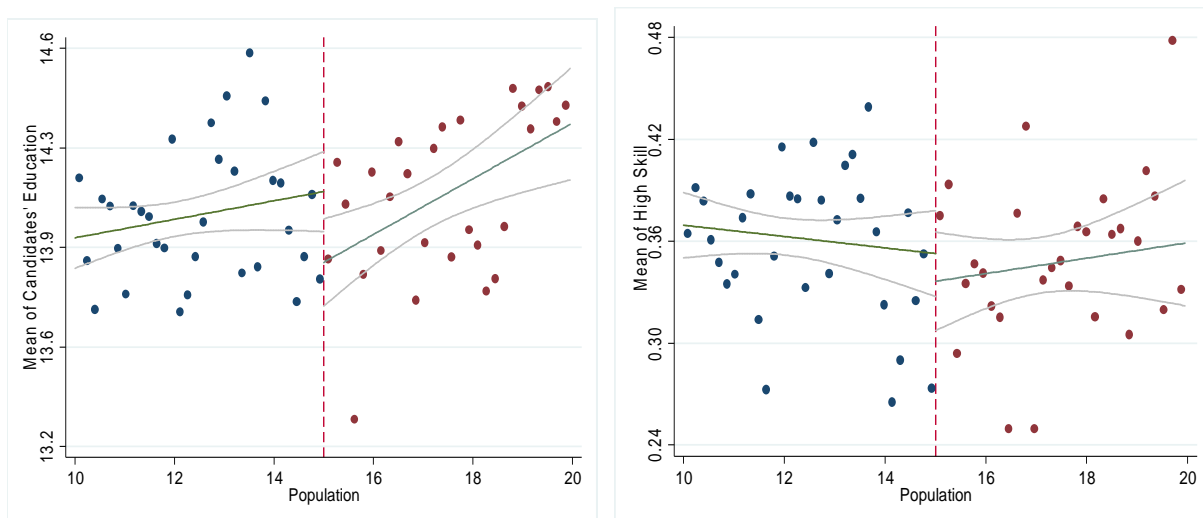


Figure 3: Sharp RDD. Effect of Runoff on Politicians' Quality at the 15,000 cutoff

Another potential channel explaining the negative effect of the electoral system on the educational attainment of politicians is the level of political competition characterizing the electoral

race. Many authors (see for instance, Bordignon *et al.*, 2016; Cox and Neto, 1997, and Chamon *et al.*, 2008) found that under runoff rule more candidates and political parties compete at the electoral race, compared to single ballot elections.

To directly test whether political closeness, as measured by the number of candidates/parties running at elections, affects the quality of local elected politicians, we implement a Fuzzy RDD⁹ with optimal bandwidth (half IK), where the measure of political closeness is instrumented on a dummy variable taking the value of 1 if cities vote under runoff (above 15,000 inhabitants) and 0 otherwise (below 15,000 threshold).

Table 5. Effect of Political Competition on Average Education. Fuzzy RDD. LLR Results

| | (1) | (2) | (3) | (4) |
|---------------------------|----------------------|----------------------|----------------------|----------------------|
| VARIABLES | Politicians' Educ | Politicians' Educ | Politicians' Educ | Politicians' Educ |
| Panel A: 2SLS | | | | |
| #Candidates | -0.237 (0.242) | -0.192 (0.216) | | |
| #Parties | | | -0.0301 (0.0305) | -0.0266 (0.0286) |
| Constant | 16.09*** (1.050) | 13.97*** (0.978) | 15.30*** (0.290) | 14.15*** (0.990) |
| Province and year FE | Yes | Yes | Yes | Yes |
| Controls | No | All | No | All |
| Population Polynomial | First | First | First | First |
| Optimal Bandwidth | Half IK | Half IK | Half IK | Half IK |
| Observations | 1,689 | 1,685 | 1,448 | 1,444 |
| R-squared | 0.113 | 0.184 | 0.127 | 0.199 |
| Panel B: First Stage | | | | |
| VARIABLES | #Candidates | #Candidates | #Parties | #Parties |
| Runoff | 0.5887*** (0.205) | 0.6266*** (0.127) | 4.8048*** (0.281) | 4.8701*** (0.281) |
| First Stage <i>F-Stat</i> | 13.08 | 13.17 | 105.03 | 81.72 |
| (p-value) | (0.000) | (0.000) | (0.000) | (0.000) |
| R squared | 0.152 | 0.189 | 0.623 | 0.627 |

Note: Municipalities between 11,275 and 18,725 inhabitants. Sharp RDD estimates. We control for province and year fixed effects and we focus on the period 1994-2017. Estimation method: LLR with a first order polynomial of the forcing variable. Standard Errors are robust to heteroscedasticity and are cluster at municipal level (reported inside the brackets). Significance at the 10% level is represented by *, at the 5% level by **, and at 1% level by ***.

Results are displayed in Table 5. In columns (1) and (2) we focus on the number of candidates running for a mayor position, whereas in columns (3) and (4) we look at the number of political parties. Furthermore, in odd columns we only control for province and year fixed effects, whereas in even columns we add all of candidates and municipal controls. From the First-stage regression (Panel B) we can notice that both the number of candidates and political parties sharply increase in the neighborhood of the threshold and the instrument is highly correlated with our measures of political competition (*F-stat* much larger than 10). However, from the second stage (Panel A) we

⁹ We adopt this empirical methodology, since the number of candidates running for a mayor position and the number of political parties are not a deterministic function of the threshold where the electoral system changes. In fact, the decision of running might depend on some unobserved candidates' characteristics. Restricting our analysis in the neighborhood of the cutoff allows us to assume that municipalities just below and above the cutoff are comparable in all their baseline characteristics, including those not observed.

find a negative, but never statistically significant effect on the educational attainment of local politicians. This suggests that the lower educational level of politicians in cities above the threshold compared to that in municipalities below the cutoff is not driven by the divergence in the level of political closeness characterizing the two electoral regimes.

5. Runoff and Politicians' Quality: Heterogeneous Effects

The richness of our dataset allows us to investigate whether the change in the electoral system differently affects the educational attainment, as measured by years of schooling, of mayors, councilors, aldermen, male and female local politicians nearby the 15,000 threshold separately. Table 6 reports the results.

In cities just above the cutoff mayors have an educational attainment lower than the level of education of mayors in municipalities just below the threshold by 0.757 years. In other words, runoff decreases the education of mayors by 4.89 percent roughly compared to the average value of 15.48 in our sample, using half IK optimal bandwidth. The coefficient attached to our treatment variable is significant at 10 percent level. We find similar results, even if the significance is at 1 percent level, when our outcome variable is as measured by years of schooling of councilors (we observe a decrease of 2.76 percent passing from a single to a double ballot). Furthermore, the runoff rule negatively affects the educational attainment of male politicians within the local government apparatus. Conversely, no effect is found on female politicians' education and on aldermen education (in this case we did not expect any impact since aldermen are chosen directly by mayors once the local government is composed).

Table 6. Effect of Runoff Elections on Education for Mayors, Councilors and Aldermen. Sharp RDD. LLR Results

| VARIABLES | (1) Mayors Educ | (2) Councilors' Educ | (3) Aldermen Educ | (4) Male Politicians' Educ | (5) Female Politicians' Educ |
|-----------------------|---------------------|-------------------------|----------------------|-------------------------------|---------------------------------|
| Runoff | -0.757* (0.407) | -0.381*** (0.117) | -0.0486 (0.257) | -0.293** (0.126) | -0.0507 (0.249) |
| Constant | 12.09*** (2.997) | 11.97*** (1.028) | 13.36*** (2.173) | 11.56*** (1.292) | 10.56*** (1.759) |
| Controls | Yes | Yes | Yes | Yes | Yes |
| Province and year FE | Yes | Yes | Yes | Yes | Yes |
| Population Polynomial | First | First | First | First | First |
| Interaction term | First | First | First | First | First |
| Optimal Bandwidth | Half IK | Half IK | Half IK | Half IK | Half IK |
| Observations | 1,576 | 1,613 | 1,301 | 1,684 | 1,496 |
| R-squared | 0.180 | 0.403 | 0.213 | 0.344 | 0.219 |

Note: Municipalities between 11,275 and 18,725 inhabitants. Sharp RDD estimates. We control for province and year fixed effects and we focus on the period 1994-2017. Estimation method: LLR with a first order polynomial of the forcing. Standard Errors are robust to heteroscedasticity and are cluster at municipal level (reported inside the brackets). Significance at the 10% level is represented by *, at the 5% level by **, and at 1% level by ***.

In Figure 4 we plot Sharp RDD results presented in Table 6. In particular, in Panel (a) we focus on the average educational attainment of all of mayors (Panel b), councilors (Panel c), aldermen

(Panel d), male politicians (Panel e) and female politicians (Panel f). Again, descriptive graphs confirm results displayed in Table 6 and highlight a sharp decrease in local politicians' education passing from a single to a double ballot.

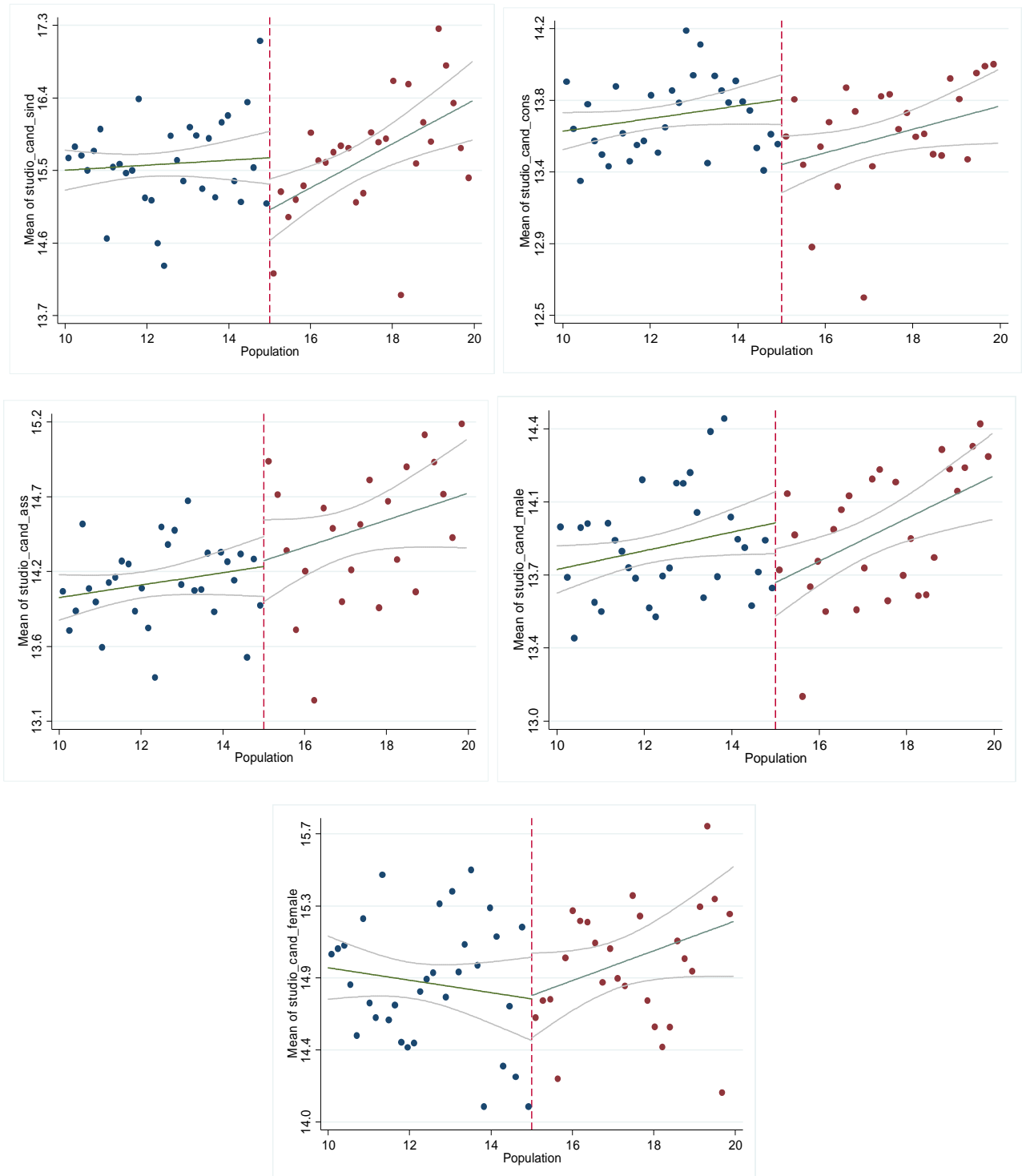


Figure 4: Sharp RDD. Heterogeneous Effect of Runoff on Education at the 15,000 cutoff

The channels through which runoff rule affects the quality of local politicians may work dissimilarly based on the educational attainment of population. To investigate the heterogeneous

effect of double ballot rule on politicians' education we run separate regressions¹⁰ splitting the sample based on the median value (7.680873) of the variable *Population Education*.

Table 7 reports the results. In columns (1) and (2) we focus on municipalities in our sample (half IK optimal bandwidth) with an educational attainment of the resident population above the median, whereas in columns (3) and (4) we look at cities characterized by a low level of education (below the median). Moreover, in odd specifications we control for province and year fixed effects only, while in even columns we add the full set of controls. In all specifications, we use a linear polynomial of the forcing variable along with a linear interaction term between the treatment and the normalized population size.

All in all, runoff scheme negatively affects local politicians' quality, as measured by their average years of schooling, only in those municipalities with a low level of education. In particular, in these cities (below the median) runoff leads to a reduction in politicians' education by 0.3 years, corresponding to a decrease of 2.21 percent compared to the average value of 13.73 in our sample using half IK optimal bandwidth. The coefficient is stable across specifications reported in columns (3) and (4) and is significant at 10 percent level.

Table 7. Runoff Elections and Politicians' Education. Heterogeneity based on Education population LLR Results

| | (1) >Median Education Pop | (2) >Median Education Pop | (3) <=Median Education Pop | (4) <=Median Education Pop |
|-----------------------|---------------------------------|---------------------------------|----------------------------------|----------------------------------|
| VARIABLES | Politicians' Educ | Politicians' Educ | Politicians' Educ | Politicians' Educ |
| Runoff | -0.198 (0.189) | -0.233 (0.188) | -0.333* (0.187) | -0.304* (0.186) |
| Constant | 15.02*** (0.547) | 16.48*** (1.629) | 15.50*** (0.313) | 15.60*** (1.204) |
| Controls | No | All | No | All |
| Province and year FE | Yes | Yes | Yes | Yes |
| Population Polynomial | First | First | First | First |
| Interaction term | First | First | First | First |
| Optimal Bandwidth | Half IK | Half IK | Half IK | Half IK |
| Observations | 893 | 893 | 799 | 795 |
| R-squared | 0.306 | 0.320 | 0.381 | 0.392 |

Note: Municipalities between 11,275 and 18,725 inhabitants. Sharp RDD estimates. We control for province and year fixed effects and we focus on the period 1994-2017. Estimation method: LLR with a first order polynomial of the forcing. Standard Errors are robust to heteroscedasticity and are cluster at municipal level (reported inside the brackets). Significance at the 10% level is represented by *, at the 5% level by **, and at 1% level by ***.

6. Robustness checks

As a first check, in Table 8 we adopt several specifications of our model to assess the sensitivity and magnitude of the runoff system effect on politicians' quality. In particular, in column (1) of Table 8 we use IK optimal bandwidth, like in Table 3, and we add a quadratic function of the normalized forcing variable along with a second order interaction term between *Runoff* and population size, whereas in column (2) we replicate specifications of column (1), but we add the full

¹⁰ The regressions are the same as that reported in column (4) of Table 4, in which we consider the full set of controls.

set of controls. Results are similar to those previously presented. In fact, runoff elections lead to a decrease in the average years of schooling of local politicians by 0.255 years (see column 2).

Table 8. Effect of Runoff Elections on Average Education with several bandwidths and polynomials. Sharp RDD.

| | (1) | (2) | (3) | (4) | (5) |
|-----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| VARIABLES | Politicians' Educ | Politicians' Educ | Politicians' Educ | Politicians' Educ | High Skill |
| Runoff | -0.257* (0.135) | -0.255** (0.129) | -0.377** (0.146) | -0.297** (0.141) | -0.0401* (0.0214) |
| Constant | 14.35*** (0.359) | 10.71*** (0.742) | 14.92*** (0.452) | 11.99*** (1.236) | 0.366** (0.172) |
| Controls | No | Yes | No | Yes | Yes |
| Province and year FE | Yes | Yes | Yes | Yes | Yes |
| Population Polynomial | Second | Second | First | First | First |
| Interaction term | Second | Second | First | First | First |
| Optimal Bandwidth | IK | IK | CCT | CCT | CCT |
| Observations | 4,169 | 4,165 | 1,213 | 1,209 | 1,230 |
| R-squared | 0.277 | 0.333 | 0.343 | 0.381 | 0.616 |

Note: Sharp RDD estimates. We control for province and year fixed effects and we focus on the period 1994-2017. Estimation method: LLR. Standard Errors are robust to heteroscedasticity and are cluster at municipal level (reported inside the brackets). Significance at the 10% level is represented by *, at the 5% level by **, and at 1% level by ***.

In columns (3) and (4) we instead use as optimal bandwidth that proposed by Calonico *et al.* (2014a) (CCT hereafter). We end up with a sample of 2,700 observations above and below the 15,000 cutoff. The effect of runoff rule is still negative and statistically significant at 5 percent level. Finally, in column (5), using CCT optimal bandwidth, we measure our outcome variable as the proportion of high skilled local politicians in the municipal bodies. Again the effect of runoff rule on politicians' quality is negative and significant at conventional levels.

All in all, findings are not sensitive to the inclusion of higher order polynomials of the population size and to the use of different optimal bandwidths. In all specifications we find a negative and statistically significant effect of runoff elections on politicians' quality.

As a second test, we present estimations, using half IK optimal bandwidth and a linear polynomial of the forcing variable with a first order interaction term, where we include among controls different predetermined measures of political competition. Results are reported in Table 9. In columns (1) and (2) we respectively include the difference in vote between the first and second best candidates over the total number of eligible voters at time $t-1$, and the percentage difference in votes between the two leading parties at time $t-1$, that are inverse measures of expected electoral closeness, whereas in columns (3) and (4) we use the number of candidates and of parties competing in the previous electoral race respectively. In all of specifications, the effect of runoff system on the average years of schooling of candidates is still negative and significant at 1 percent level, independently from the political closeness measure used.

Table 9. Effect of Runoff Elections on Average Education with measures of Political Competition. Sharp RDD.

| | (1) | (2) | (3) | (4) |
|-------------------------------------|----------------------|----------------------|----------------------|----------------------|
| VARIABLES | Politicians' Educ | Politicians' Educ | Politicians' Educ | Politicians' Educ |
| Runoff | -0.328*** (0.125) | -0.439*** (0.140) | -0.334*** (0.126) | -0.436*** (0.146) |
| % Candidates Electoral Margin (t-1) | -0.429 (0.320) | | | |
| %Parties Electoral Margin (t-1) | | -0.374 (0.419) | | |
| #Candidates (t-1) | | | 0.0163 (0.0255) | |
| #Parties (t-1) | | | | -0.0013 (0.0142) |
| Constant | 12.93*** (1.218) | 13.90*** (1.356) | 12.85*** (1.219) | 12.82*** (1.478) |
| Controls | All | All | All | All |
| Province and year FE | Yes | Yes | Yes | Yes |
| Population Polynomial | First | First | First | First |
| Interaction term | First | First | First | First |
| Optimal Bandwidth | Half IK | Half IK | Half IK | Half IK |
| Observations | 1,403 | 1,173 | 1,403 | 1,176 |
| R-squared | 0.351 | 0.333 | 0.351 | 0.331 |

Note: Municipalities between 11,275 and 18,725 inhabitants. Sharp RDD estimates. We control for province and year fixed effects and we focus on the period 1994-2017. Estimation method: LLR with a first order polynomial of the forcing. Standard Errors are robust to heteroscedasticity and are cluster at municipal level (reported inside the brackets). Significance at the 10% level is represented by *, at the 5% level by **, and at 1% level by ***.

As a final robustness check, we exploit the panel nature of our dataset and re-estimate our model by means of OLS with fixed effects at municipal level. Controlling for municipal fixed effects, rather than for province dummies, should eliminate any bias deriving from unobserved time invariant municipal characteristics that might affect both population size, and in turn the probability that the municipality will vote under single or double ballot, as well as politicians' quality.

In Table 10, we present the within-group estimates. In all of specifications we control for a linear polynomial of the normalized population size with a first order interaction term between the forcing variable and *Runoff*. From column (1) to (3) we choose half IK optimal bandwidth, and we find that runoff rule leads to a decrease both in the average years of schooling of local politicians (by 0.278 years, see column 2) as well as in the proportion of high skilled administrators (by 6.79 percentage points, see column 3).

From column (4) to (6), when an optimal bandwidth as proposed by Calonico *et al.* (2014a) is chosen, findings highlight that the impact of runoff rule on the average educational attainment of local politicians is always negative and statistically significant at the 5 percent level, even if the magnitude is larger (about 0.4 years).

Table 10. Effect of Runoff Elections on Average Education with municipal FE. Sharp RDD

| VARIABLES | (1) | (2) | (3) | (4) | (5) | (6) |
|--------------------------|----------------------|----------------------|------------------------|----------------------|----------------------|-----------------------|
| | Politicians' Educ | Politicians' Educ | High Skill | Politicians' Educ | Politicians' Educ | High Skill |
| Runoff | -0.265* (0.158) | -0.278* (0.154) | -0.0679*** (0.0243) | -0.416** (0.182) | -0.415** (0.176) | -0.0656** (0.0272) |
| Constant | 15.06*** (0.189) | 6.513 (7.117) | 0.454 (0.922) | 15.22*** (0.228) | 12.20 (11.41) | 0.444 (1.560) |
| Controls | No | All | All | No | All | All |
| Municipal and year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Population Polynomial | First | First | First | First | First | First |
| Interaction term | First | First | First | First | First | First |
| Optimal Bandwidth | Half IK | Half IK | Half IK | CCT | CCT | CCT |
| Observations | 1,692 | 1,688 | 1,713 | 1,300 | 1,296 | 1,317 |
| R-squared | 0.232 | 0.241 | 0.554 | 0.227 | 0.242 | 0.570 |
| Number of Municipalities | 515 | 514 | 515 | 386 | 385 | 385 |

Note: Sharp RDD estimates. We control for municipal and year fixed effects and we focus on the period 1993-2017. Estimation method: LLR with a first order polynomial of the forcing. Standard Errors are robust to heteroscedasticity and are cluster at municipal level (reported inside the brackets). Significance at the 10% level is represented by *, at the 5% level by **, and at 1% level by ***.

7. Concluding remarks

One of the major concerns in all modern democracies is that the level of education might register a dramatic downward shift. In fact, the percentage of newly elected politicians with a degree decreased steadily by 28 percent roughly at national level in Italy, and, this trend is also present at municipal level. Consequently, the selection of high quality politicians is crucial for the well-functioning of any kind of democracies, as widely highlighted by the theoretical literature (Caselli and Morelli, 2004; Messner and Polborn, 2004).

In our paper we empirically investigated the effect of an exogenous change, decided by population threshold, in the electoral system (single ballot vs runoff) on the quality of local politicians, as measured by the average educational attainment. Findings show that in municipalities voting under runoff rule (just above the threshold) the average years of schooling of local politicians is lower by roughly 2 percent with respect to the educational attainment of politicians in cities using a single ballot scheme (just below the threshold). We also find that the change in the electoral system negatively affects the level of education of mayors, councilors and male local politicians, whereas no effect is found on aldermen and female politicians.

The results are similar when we measure the quality of politicians in terms of past employment, when we control for different predetermined measures of political closeness and when we adopt several specifications of our model (different optimal bandwidths and second order polynomials of the forcing variable) to check the sensitivity in terms of sign, significance and magnitude of the runoff effect on our outcome variable.

We speculate that the negative effect is exclusively driven by the different selection process of politicians adopted by parties under the two divergent regimes. In municipalities voting under a single ballot, the candidate who collects a majority of votes (elected mayor) wins the control of government (along with the only list that supported him/her during elections) and appropriates all office rents, because he/she will certainly have a majority within the council. In this case, the only

list supporting the mayor candidate is assumed to be composed by high quality candidates, if the target of winning the electoral race and of getting the whole rent wants to be reached. Whereas in municipalities voting under runoff this may not happen, since the mayor can be with the majority of the council in hand of opposition parties, and in turn all parties obtaining a positive share of votes can aspire to influence government, and hence participate in the division of office rents, depending on the share of votes obtained in the election. This mechanism explaining the willingness of parties not to select high skilled candidates in runoff elections highlight how parties should think carefully both on which candidates they should invest more as well as who is more appropriate in terms of educational attainment to become politician, as political parties' actions are likely to generate possibly unintended differences in terms of politicians' quality, and in turn, in terms of efficiency of the local public administration and of voters' well-being.

References

- Baltrunaite, A, Bello, P, Casarico, A, and Profeta, P. 2014. Gender quotas and the quality of politicians. *Journal of Public Economics* 118, 62–74.
- Barone, G, de Blasio, G. 2011 Electoral rules and voter turnout, *International Review of Law and Economics*, Vol. 36, 25-35.
- Besley, T., 2007. *Principled agents? The political economy of good government*. Oxford University Press, USA.
- Besley, T., Coate, S., 1997. An economic model of representative democracy. *The Quarterly Journal of Economics* 112, 85-114.
- Besley, T., Montalvo, J.G., Reynal-Querol, M., 2011. Do educated leaders matter? *The Economic Journal* 121, 205-227.
- Besley, T., Persson, T., Sturm, D.M., 2007. Political competition, policy and growth: theory and evidence from the US. *The Review of Economic Studies* 77, 1329-1352.
- Bordignon, M, Tabellini, G, and Nannicini, T. 2016. Moderating political extremism: single round vs runoff elections under plurality rule. *American Economic Review*, American Economic Association, vol. 106(8), 2349-2370.
- Bracco, E, Brugnoli, A. 2012. Runoff vs plurality: effects of electoral system on intergovernmental grants. A regression discontinuity analysis. *Lancaster Economics Working Paper* 2.
- Caselli, F., Morelli, M., 2004. Bad politicians. *Journal of Public Economics* 88, 759-782.
- Callander, S. 1999. Electoral competition with entry. Working Papers 1083, California Institute of Technology, Division of the Humanities and Social Sciences.
- Calonico, S., Cattaneo, MD., Titiunik, R., 2014. Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica* 82 (6): 2295–2326.
- Caughey, D, Sekhon. JS. 2011. Elections and the regression discontinuity design: Lessons from close US House races, 1942–2008. *Political Analysis* 19(4), 385–408.
- Chamon, M, Pinho de Mello, JM, and Firpo, S. 2008. Electoral rules, political competition and fiscal spending: regression discontinuity evidence from Brazilian municipalities. *Textos para discussão* 559, Department of Economics PUC-Rio (Brazil).
- Chattopadhyay, R. Duflo, E. 2004. Women as policy-makers: Evidence from a randomized policy experiment in India, *Econometrica*, 72(5), 1409-1443.
- Cox, G, Neto, OA, 1997. Electoral institutions, cleavage structures and the number of parties. *American Journal of Political Science* 41, 149–174.

- Daniele, G, Geys B. 2015. Organised crime, institutions and political quality: empirical evidence from Italian municipalities, *The Economic Journal* 125, 233-255.
- De Paola, M. and V. Scoppa (2011), Political Competition and Politician Quality: Evidence from Italian Municipalities, *Public Choice*, 148(3), 547-559.
- Ferraz, C., Finan, F., 2011. Motivating politicians: the impacts of monetary incentives on quality and performance. NBER Working Paper.
- Gagliarducci, S., Nannicini, T., 2011. Do better paid politicians perform better? Disentangling incentives from selection. *Journal of the European Economic Association* 11, 369-398.
- Galasso, V, Nannicini, T. 2011. Competing on good politicians. *American Political Science Review* 105, 79-99.
- Jones, B.F., Olken, B.A., 2005. Do leaders matter? National leadership and growth since World War II. *The Quarterly Journal of Economics* 120, 835-864.
- Imbens, G., Kalyanaraman, K., 2012. Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies* 79 (3), 933-959.
- Lee, D., 2008. Randomized experiments from non-random selection in U.S. House elections. *Journal of Econometrics* 142, 675–697.
- Lee, DS, Moretti, E, and Butler, MJ. 2004. Do voters affect or elect policies? Evidence from the U.S. House.” *The Quarterly Journal of Economics* 119(3), 807–859.
- Matozzi, M, Merlo, A. 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, 698–714.
- Messner, M., Polborn, M.K., 2004. Paying politicians. *Journal of Public Economics* 88, 2423-2445.
- Milesi-Ferretti, G.M., Perotti, R., and Rostagno, M. 2002. Electoral systems and public spending. *The Quarterly Journal of Economics*, 117 (2), 609–657.
- Osborne, M.J., Slivinski, A., 1996. A model of political competition with citizen-candidates. *The Quarterly Journal of Economics* 111, 65-96.
- Pande, R. 2003. Can mandated political representation increase policy influence for disadvantaged minorities?, *American Economic Review*, 93(4), 1132–1151.
- Persson, T, Tabellini, G. 2005. *The Economic Effects of Constitutions*, Vol. 1, The MIT Press.
- Rizzo, L, Zanardi, A. 2010. Single ballot vs double ballot: does it matter for fiscal policies? Evidence from Italy. Technical Report, mimeo 2010