DISCUSSION PAPER SERIES

No. 10323

MODERATING POLITICAL EXTREMISM: SINGLE ROUND VS RUNOFF ELECTIONS UNDER PLURALITY RULE

Massimo Bordignon, Tommaso Nannicini and Guido Tabellini

PUBLIC ECONOMICS

Centre for Economic Policy Research

ISSN 0265-8003

MODERATING POLITICAL EXTREMISM: SINGLE ROUND VS RUNOFF ELECTIONS UNDER PLURALITY RULE

Massimo Bordignon, Tommaso Nannicini and Guido Tabellini

Discussion Paper No. 10323
December 2014
Submitted 17 December 2014

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801
www.cepr.org

MODERATING POLITICAL EXTREMISM: SINGLE ROUND VS RUNOFF ELECTIONS UNDER PLURALITY RULE[†]

Abstract

We compare single round vs runoff elections under plurality rule, allowing for partly endogenous party formation. Under runoff elections, the number of political candidates is larger, but the influence of extremist voters on equilibrium policy and hence policy volatility is smaller, because the bargaining power of the political extremes is reduced compared to single round elections. The predictions on the number of candidates and on policy volatility are confirmed by evidence from a regression discontinuity design in Italy, where cities above 15,000 inhabitants elect the mayor with a runoff system, while those below hold single round elections.

JEL Classification: c14, d72 and H72
Keywords: electoral rules, policy volatility and regression discontinuity design

Massimo Bordignon massimo.bordignon@unicatt.it
Def, Università Cattolica del Sacro Cuore and CESifo

Tommaso Nannicini tommaso.nannicini@unibocconi.it
IGIER, Università Bocconi and IZA

Guido Tabellini guido.tabellini@unibocconi.it
IGIER, Università Bocconi, CIFAR, CESifo and CEPR

[†] We thank Pierpaolo Battigalli, Carles Boix, Daniel Diermeier, Massimo Morelli, Giovanna Iannantuoni, Francesco de Sinopoli, Ferdinando Colombo, Piero Tedeschi, Per Petterson-Lindbom, the coeditor Debraj Ray, four anonymous referees, and seminar participants at Berkeley, CIFAR, Pompeu Fabra, Brescia, Cattolica, Munich, Warwick, Cesifo Workshop, IGIER workshop in Political Economics, IIPF annual conference, NYU conference in Florence for several helpful comments. We also thank Massimiliano Onorato for excellent research assistance, and Veruska Oppedisano, Paola Quadrio, and Andrea Di Miceli for assistance in collecting the data. Financial support is gratefully acknowledged from the Italian Ministry for Research and Catholic University of Milan for Bordignon, from ERC (grant No. 230088) and Bocconi University for Nannicini, and from the Italian Ministry for Research, CIFAR, ERC (grant No. 230088), and Bocconi University for Tabellini.

1 Introduction

In some electoral systems, citizens vote twice: in a first round they select a subset of candidates, over which they cast a final vote in a second round. The system for electing the French President, where the two candidates who get more votes in the first round are admitted to the second, is possibly the best known example. But variants of this *runoff* system are increasingly used in many other countries, for example in Latin America, in US gubernatorial primary elections, and in many local elections, including Italian municipal elections (see Cox, 1997, and Golder, 2005). How does the runoff system differ from the more common *single round* plurality rule, where candidates are directly elected at the first round? In spite of its obvious relevance, this question remains largely unaddressed, particularly when it comes to studying the economic policies enacted under these two electoral systems.

This paper contrasts runoff vs single round elections under plurality rule, focusing on the policy platforms that get implemented in equilibrium. We analyze a model where parties or candidates with ideological preferences set a one-dimensional policy before the election. The number of parties is partly endogenous. We start out with four candidates. Before the elections, however, candidates choose whether or not to merge into a party, and bargain over rents and the policy platform that would result from merging. We obtain two main results. First, in equilibrium the number of parties is larger under runoff than under single round elections. Second, and more important, if the electorate is polarized then the runoff system moderates the influence of extremist candidates and voters on the equilibrium policy, thereby inducing more centrist policies. The reason is that runoff elections reduce the bargaining power of the extremist candidates, which typically appeal to a smaller electorate. Intuitively, with a single round and under sincere voting, the extremes can threaten to cause the electoral defeat of the nearby moderate candidate if he refuses to strike an alliance. Under runoff this threat is empty, provided that enough extremist voters are willing to vote for the closest moderate, rather than abstain, in the second round. The analysis also clarifies that these results hold under general but not universal features of the political system. In particular, the results are fairly robust to the number of parties and to the assumptions on voter's behavior, but they require the presence of large and sufficiently polarized groups of moderate voters.

Because of the larger influence of the political extremes, equilibrium policies are more polarized under single round than under runoff. Therefore, conditional on the same degree of political turnover, policy volatility is lower in the latter. The model thus yields three empirical predictions: under runoff, as opposed to single round elections, we should observe (1) more political candidates, (2) less alliances between moderate and extremist parties, and (3) less policy volatility. We test these predictions with data on municipal elections in Italy.

Since 1993, Italian mayors are directly elected and have a prominent role in determining policy. Municipalities below 15,000 inhabitants adopt a single round system, while a runoff system is in place above this threshold. The data also suggest that voters are indeed mobile between candidates: a relevant share of the voters supporting the excluded candidates seems to participate again in the second round. Moreover, empirical tests in the spirit of Fujiwara (2011) do not detect evidence of strategic voting. The Italian party system over our sample period (1993–2007) displayed two main moderate blocks, one on the center-left and one on the center-right (both with an average vote share in national elections from 30 to 35 percent, depending on the election), plus an extremist party on the left (from 4 to 8 percent) and one on the right (from 4 to 10 percent). In local elections, the underlying party structure mainly reflected the features of the national party system. Overall, this institutional setup allows us to test the model's predictions with a quasi-experimental strategy, that is, implementing a Regression Discontinuity Design (RDD) around the 15,000 threshold.

We test the implications of our model with respect to both *politics* and *policy*. First, we check whether the number of candidates for mayor is larger under the runoff system, as opposed to the single round. The positive discontinuity at 15,000 is indeed large and statistically significant: under runoff elections the number of candidates for mayor increases by about 29%. Second, to test the prediction that runoff elections moderate political extremism and reduce policy volatility, we focus on one of the main policy tools of municipalities, the business property tax. In 1993, with the introduction of this tax, Italian municipalities were given large discretion in setting the tax rate, whose proceeds could be freely allocated to all municipal functions, such as social assistance, housing, education, and so on.

The intuition for this test is simple. The size of government is influenced by ideology, with left-wing governments generally raising more tax revenues and imposing higher business property taxes (this is indeed confirmed by our data). Hence, on average a change in the identity of the mayor should lead to a sharper policy change where the influence of the extremist parties is stronger, namely under single round elections. The RDD evidence supports this prediction. We measure the volatility of the business property tax rate in two ways: by the intertemporal variance (i.e., across legislative terms for the same municipality) and by the cross-sectional variance (i.e., within population bins in the same year). Both indicators display a discontinuity at 15,000, with less volatility above the threshold, which is both large and statistically significant. The estimated coefficients point to an impact of about 61% of runoff elections on the time series volatility of the tax rate, and an impact of about 71% on the cross-sectional volatility around the population threshold.

Alternative explanations for this (reduced-form) effect on tax volatility are rejected by our data, because the turnover between different mayors is similar in both runoff and single

round elections. Moreover, in a small and selected subsample of municipalities where we can measure the political identity of mayoral candidates, runoff elections have a negative impact on the probability that the leftist political extreme—i.e., the Communist Party—joins the main center-left coalition at the local level, in line with a direct implication of our model. Overall, the empirical evidence supports the hypothesis that the runoff system reduces the influence of the political extremes and induces policy moderation.

Our results have important implications for the design of democratic institutions. Political extremism is widespread in many advanced and developing countries (including Italy) and is often counterproductive. It reduces ex-ante welfare if voters are risk averse, and it induces sharp disagreement that often disrupts decision making in governments or legislatures. In this respect, runoff electoral systems have an advantage over single round systems, as they moderate the influence of extremist groups and reduce the welfare costs associated with (partisan) policy volatility. Care should be taken, however, in extrapolating our results to different institutional contexts, such as primary elections to select parties' candidates.[1]

The existing theoretical literature on these issues is not large. Some informal conjectures have been advanced by institutionally oriented political scientists (Sartori, 1995; Fisichella, 1984). Analytical work has mostly asked whether variants of "Duverger's Law" or "Duverger's Hypothesis" carry over to the runoff system under strategic voting (Messner and Polborn, 2004; Cox, 1997; Callander, 2005; Bouton, 2013; Bouton and Gratton, 2013).[2] Our results on the number of parties under the two systems support both Duverger's Hypothesis and Duverger's Law in the presence of sincere voting and strategic parties, along the lines of Fey (2007) and Morelli (2004). Unlike these papers, however, we focus on which policies are implemented in equilibrium. This is an important question, largely unaddressed in the literature. An exception is Osborne and Slivinsky (1996). In a citizen-candidate model with sincere voting and ideologically motivated candidates, they study the equilibrium configura-

[1]Runoff primary elections have been proposed to alleviate political extremism within the US parties (see Fiorina, 2005). The mechanism at work in this paper would not necessarily support this conclusion, however. In our framework, runoff elections dampen political extremism if there are more moderate than extremist voters, but not otherwise. This condition may not hold in primary elections, where often participation amongst party activists is much higher than in the electorate at large.

[2]This terminology is due to Riker (1982). "Duverger's Law" states that plurality rule leads to a stable two-party configuration, as strategic voters should concentrate their votes on the two most serious candidates, while "Duverger's Hypothesis" suggests that a configuration with several parties/candidates should emerge from proportional representation. Duverger's Law can be rationalized as a result of strategic voting (see Feddersen, 1992, and the literature discussed there), and there is an extensive theoretical literature on strategic behavior in single ballot elections under different electoral rules (see Myerson and Weber, 1993; Fey, 1997). Less is known about the runoff system under strategic behavior. Bouton (2013) and Bouton and Gratton (2013) show that Duverger's Law may hold even under runoff, with strategic voting and appropriate equilibrium selection. Callander (2005) considers strategic parties, and suggests that a larger number of candidates is not a necessary implication of a runoff, as pre-existing parties may strategically locate on policy issues so as to prevent the entrance of other candidates.

tion of candidates and policies in the two systems, concluding that policy platforms are in general more dispersed under single ballot plurality rule than under runoff. But in keeping with the Duverger's tradition, their result is obtained in a long run equilibrium where all possibilities for profitable entry by endogenous candidates are exhausted. We instead discuss this issue from a different perspective, where pre-existing policy oriented parties (or candidates) bargain under the two different electoral systems. Lizzeri and Persico (2005) also study the policy effect of runoff elections. They suggest that a runoff system reduces the number of effective parties. This is desirable in the context of their model, as electoral competition with several parties leads to equilibrium policies that cater only for narrow constituencies.

The empirical evidence on these issues is small. Earlier works, using identification strategies based on conditional independence, have detected either a positive or a zero effect of runoff on the number of political candidates in different contexts (Wright and Riker, 1989; Engstrom and Engstrom, 2008; Cox, 1997). More recent works on Brazil, independent from our study and also based on RDD for identification, have found results comparable to ours: Chamon et al. (2009) show that runoff is indeed characterized by a larger number of effective candidates; Fujiwara (2011) finds support for Duverger's argument of strategic voter behavior under single round elections, but—unlike Chamon et al. (2009)—he discloses an effect of the electoral rule on the number of candidates that is not statistically different from zero in most specifications. Due to institutional limitations, however, RDD identification is more problematic in Brazil, because: (i) the sample size around the threshold of 200,000 voters is smaller; (ii) the running variable is not the number of inhabitants but the number of registered voters, potentially more prone to manipulation. We do not know of any empirical paper that studies the moderating effects of runoff as opposed to single round elections.

The rest of the paper is organized as follows. Section 2 presents the basic model. Sections 3 and 4 study party and policy formation under single round and runoff elections, respectively, deriving the main results (proofs are in the Online Appendix I). Section 5 discusses some extensions (derivations are in the Online Appendix II). Section 6 describes the electoral system of Italian municipalities and tests the model's predictions on the number of candidates and policy volatility (validity tests are in the Online Appendix III). Section 7 concludes.

2 The model

This section outlines a stylized model. We deliberately focus on the strategic behavior of parties, and keep the model simple to illustrate the main incentives at work under different electoral rules. We discuss below and in the Online Appendix II the robustness of our results under different assumptions.

2.1 Voters

The electorate consists of four groups of voters indexed by $J = 1-4$, with policy preferences:

$$c^J = -C(|t^J - q|)$$

where $q \in [0,1]$ denotes policy, t^J is group J's bliss point, and $C(.)$ is an increasing and strictly convex function with $C(0) = 0$. Thus, voters lose utility at an increasing rate if policy is further from their bliss point. The bliss points of each group have a symmetric distribution on the unit interval, with: $t^1 = 0$, $t^2 = \frac{1}{2} - \lambda$, $t^3 = \frac{1}{2} + \lambda$, $t^4 = 1$, and $\frac{1}{2} \geq \lambda > 0$. Groups 1 and 4 will be called "extremist," groups 2 and 3 "moderate." The two extremist groups have a fixed size $\underline{\alpha}$. The size of the two moderate groups is random: group 2 has size $\overline{\alpha} + \eta$, group 3 has size $\overline{\alpha} - \eta$, where $\overline{\alpha}$ is a known parameter with $\overline{\alpha} > \underline{\alpha}$, and η is a random variable with mean and median equal to 0 and a known symmetric distribution over the interval $[-e, e]$, with $e > 0$. Thus, the two moderate groups have expected size $\overline{\alpha}$, but the shock η shifts voters from one moderate group to the other. We normalize total population size to unity, so that $\overline{\alpha} + \underline{\alpha} = \frac{1}{2}$.

The only role of η is to create some uncertainty about which of the two moderate groups is the largest. Throughout we assume:

$$(\overline{\alpha} - \underline{\alpha}) > e, \qquad \underline{\alpha}/2 > e. \tag{A1}$$

The first assumption in (A1) implies that, for any realization of the shock η, any moderate group is always larger than any extremist group. The second assumption implies that, for any realization of the shock η, the size of any moderate group is always smaller than the size of the other moderate group plus one of the extremist groups. We discuss the effects of relaxing these assumptions below. The realization of η becomes known at the election and can be interpreted as a shock to the participation rate or to voters' preferences.

Finally, throughout we assume that voters vote *sincerely* for the party that promises to deliver highest utility; the Online Appendix II discusses strategic voting.

2.2 Candidates

There are four political candidates, $P = 1-4$, who care about being in government but also have ideological policy preferences corresponding to those of voters:

$$u^P = U^P(q, r^P) = -C(|t^P - q|) + EV(r^P) \tag{1}$$

where r^P denotes the rents from being in government, $V(.)$ is an increasing and strictly concave function with $V(0) = 0$, and E denotes the expectation operator with respect to electoral uncertainty. The ideological policy preferences of each candidate are identical to those of the corresponding group of voters: $t^P = t^J$ for $P = J$. Rents only accrue to the party in government, and are endogenously split between the candidates belonging to this party according to the procedure described below, subject to the constraints that total rents cannot exceed $R > 0$, and that each candidate in government gets at least $\bar{r} > 0$, where of course $R > \bar{r}$. These minimal rents $\bar{r}$ are meant to capture the idea that there is some indivisibility in the perks from office, which constrains the negotiations between candidates. The total value of being in government, $R > 0$, is a fixed parameter.[3]

2.3 Policy choice and party formation

Before the election, candidates can merge into parties and present their platforms. We define mergers between candidates as "parties," although they can be thought of as electoral cartels of pre-existing parties. Once elected, the governing party cannot be dissolved.

If a candidate runs alone, he can only promise to voters that he will implement his bliss point: $q^P = t^P$. If a party is formed, it can promise to deliver any policy lying in between the bliss points of its members; thus, a party formed by candidates P and P' can offer any $q^{PP'} \in [t^P, t^{P'}]$. Policies outside this interval cannot be promised by this coalition. This assumption can be justified as reflecting lack of commitment by the candidates, as policies outside of $[t^P, t^{P'}]$ would be ex-post Pareto sub-optimal for both members.[4]

We assume that parties contain at most two members, and they have to be adjacent. Thus, say, candidate 2 can form a party with either 3 or 1, while candidate 1 can only form a party with 2. This simplifying assumption captures a realistic feature. It implies that coalitions can only be formed between ideologically close parties, and that moderate parties can sometimes run together, while opposite extremists cannot form a coalition between them, as voters would not support this coalition.[5] This gives moderate candidates an advantage (see below). Section 5 discusses how to relax some of these restrictions and Section 6 shows that they are not inconsistent with the Italian political system.

[3]We focus on a four candidate model, with two large moderates on both sides of the political spectrum and two extremists, because it fits well our main testing ground (see Section 6) and because it also allows us to make extensive use of the assumed symmetry to simplify the derivation of our results. We discuss below and in the Online Appendix II how to relax this assumption.

[4]Morelli (2004) and Levy (2004) use similar assumptions to explain the role of parties in politics.

[5]See again Morelli (2004) for a similar modeling choice, and Axelrod (1970) for a justification of this assumption. There are counterexamples in the real world of opposite extremes striking an electoral deal, but they are usually short lived. Moreover, they are never observed in our testing ground (see Section 6).

Candidates bargain on how to split the rents from office and on the policy q to be implemented if they win. Bargaining takes place before knowing the realization of η that determines the relative size of groups 2 and 3, and agreements cannot be renegotiated once the election result is known. Bargaining takes place in two stages. In the first stage parties are formed; in the second stage rent allocation and policy are determined within each party.

Specifically, in the first stage one of the candidates is selected at random and proposes to form a party to one of the adjacent candidates (or does not make any proposal). If the proposal is rejected (or no proposal is made), another candidate at random is selected to make a proposal (or none). If the proposal is accepted, the party is formed and cannot be broken until stage two of the game. The first stage game continues by selecting at random one of the candidates not included in any party and who have not made any proposal yet, to propose a merger to one of his remaining adjacent candidates (if any are left). The process continues until every candidate has either merged into a party or has had a chance to make a proposal. Note that there is no discounting of utilities when a proposal is rejected. The first stage ends with a partition of the set of all candidates into parties (possibly consisting of a single candidate). Possible partitions are either two parties, $\{1, 2\}$, $\{3, 4\}$; four parties, $\{1\}$, $\{2\}$, $\{3\}$, $\{4\}$; or three parties, $\{1, 2\}$, $\{3\}$, $\{4\}$, or $\{1\}$, $\{2, 3\}$, $\{4\}$, or $\{1\}$, $\{2\}$, $\{3, 4\}$.

In the second stage, each party sets a policy platform and determines an allocation of rents within the party (conditional on winning the election). The policy and rent allocation are determined by Nash bargaining between party members, taking as given the equilibrium policy platforms set by all remaining parties. The disagreement point for the Nash bargaining corresponds to a unilateral breakup of the party (given that all the other parties remain in place unaffected and cannot be renegotiated).

The sequential nature of the two stages (over party formation and over policy and rents) can be interpreted as reflecting lack of commitment. When parties are formed, candidates cannot commit to the substantive decisions that the party will have to make during the electoral competition. These decisions are the outcome of subsequent bargaining, under the threat of a unilateral party breakup. This sequential protocol is not implausible, and has the advantage of making the equilibrium outcome independent of the specific order with which candidates are selected to speak in the party formation stage.

An equilibrium is thus naturally defined as a partition of candidates into parties and a set of policy platforms and rent allocations, such that: (i) within each party, policy and rents correspond to the Nash bargaining solution between the candidates, taking as given the equilibrium policies and the composition of the opposing parties; (ii) the party system is such that each candidate finds it optimal to belong to his party, given the order of moves and anticipating the subsequent equilibrium outcomes.

2.4 Electoral rules

The rest of the paper contrasts two electoral rules. Under a single round rule, the candidate or party that wins the relative majority in the single election forms the government. Under a closed runoff rule, voters cast two sequential votes. First, they vote on whoever stands for election. The two parties or candidates that obtain more votes are then allowed to compete again in a second round. Whoever wins the second round forms the government. We discuss additional specific assumptions about information revelation and renegotiation between the two rounds of election in context, when illustrating the runoff system in detail.

3 Single round elections

We now derive equilibrium policies and party formation under single round elections. Suppose first that $\lambda > 1/4$ and consider the first stage of the game. With $\lambda > 1/4$ the centrist party $\{2, 3\}$ is not feasible, because each moderate candidate is closer to the extremist than to the other moderate, and thus at least one group of moderate voters would vote for the extremist candidate rather than for the centrist party. Moreover, any candidate running alone (say candidate 1 or 2) does not have any chance of victory if he runs against a moderate-extremist party (say, of candidates $\{3, 4\}$ together). The reason is that the party $\{3, 4\}$ always gets the support of all voters in groups 3 and 4 for any policy $q \in [t^3, t^4]$, and by (A1) this is the largest group of voters in a three-party equilibrium. Hence, a two-party system with extremists and moderates joined together is the only equilibrium outcome of the first stage of the game, irrespective of the sequence with which proposals are made.

Next, consider the second stage of the game. Equilibrium policies and rent allocation are determined by Nash bargaining within each party, where the disagreement point corresponds to the breakup of the party and the consequent victory of the opponent. Given that disutility from the policy is increasing in the distance from the bliss point ($C(.)$ is convex), moderate candidates have more bargaining power than the extremists. The reason is that extremists have more to lose from a party breakup that delivers the victory to the opponents. Hence in equilibrium moderate candidates get a larger share of the rents and equilibrium policies are closer to their bliss point. Nevertheless, the moderates have to accommodate at least part of the requests of the extremists, since they too stand to lose from a party breakup. Online Appendix I computes the Nash bargaining solution and formally proves (superscript e and m denote extremist and moderate candidates, respectively):

Proposition 1 *If $\lambda > 1/4$, the unique equilibrium is a two-party system, where the moderate-extremist parties* $(\{1, 2\}, \{3, 4\})$ *compete in the elections and have equal chances of winning.*

The equilibrium policy platforms q satisfy: $|t^e - q| > |t^m - q| > 0$. And equilibrium rents (conditional on winning the election) are shared according to: $r^m > r^e > \bar{r}$.

Note that, if all candidates run alone, the extremists do not have a chance. By (A1), the moderate groups are always larger than the extremist groups, for any shock to the participation rate η. Hence, in a four candidates equilibrium, the two moderates win with probability 1/2 each. This means that the moderate candidates 2 and 3 would be better off in the four candidates outcome than in the two-party equilibrium. In both situations, they would win with the same probability, 1/2, but they would not have to share rents nor compromise on policy in case of victory. But the two moderate candidates are caught in a prisoner's dilemma. If the last stage of the party formation game is reached with four candidates, then the last moderate to speak (as proposer or respondent) would find it optimal to merge with the extremist, since this would guarantee victory at the elections. Anticipating this outcome, the first moderate to speak (as proposer or respondent) would also merge with the extremist. Hence in equilibrium a two-party system always emerges. This in turn gives some bargaining power to the extremist candidates. Even if they have no chances of winning on their own, they become an essential player in the coalition.

Next, suppose that $\lambda < 1/4$. In this case the centrist party wins the election with certainty. By the symmetry of the model, Nash bargaining within the centrist party splits utilities symmetrically, so that $q = 1/2$ and $r^P = R/2$ for $P = 2, 3$. Given risk aversion in both rents and policy, it is easy to verify that both moderates are better off in forming the centrist party than with any other outcome, including a four-party system (in which case each moderate would win with probability 1/2). We thus have:

Proposition 2 *If $1/4 \geq \lambda$, then the unique equilibrium under single round elections is a three-party system with a centrist party, $(\{1\}, \{2, 3\}, \{4\})$. The centrist party wins the election with certainty, and sets $q = 1/2$ and $r^P = R/2$ for $P = 2, 3$.*

Summarizing, if the electorate is sufficiently polarized ($\lambda > 1/4$), the single round penalizes the moderate candidates and voters. A centrist party cannot emerge, because the electorate is too polarized and would not support it. The moderate candidates and voters would prefer a situation where all candidates run alone, because this would maximize their possibility of victory and minimize the loss in case of a defeat. But this party structure cannot be supported, and in equilibrium we reach a two-party system where moderate and extremist candidates join forces. This in turn gives extremist candidates and voters a chance to influence policy outcomes. If instead the electorate is not too polarized, $1/4 \geq \lambda$, then a single round system would induce the emergence of a centrist party. Extremist candidates and voters lose the elections, and moderate policies are implemented.

4 Runoff elections

We now consider a closed runoff system. The two candidates or parties that gain more votes in the first round are admitted to the second round, which in turn determines who is elected to office. To preserve comparability with single round elections, we start with exactly the same bargaining rules used in the previous section. Thus, first parties are formed (with random selection of proposers), and then the policy and the rent allocation are set by Nash bargaining within each party, with the disagreement point given by the party breakup, and taking as given the equilibrium policy set by the opponents. In particular, candidates can merge into parties and policies are determined only before the first round of election. Once a party structure is determined, it cannot be changed in any direction in between the two rounds. We also retain assumptions (A1), together with the assumption of sincere voting. We relax all of these assumptions in the next section.

As with single round, the equilibrium depends on how polarized is the electorate. If voters are polarized ($\lambda > 1/4$), then there is no policy in the interval $[t^2, t^3]$ that would command the support of all moderate voters. Hence, the centrist party $\{2, 3\}$ would lose the election with certainty and it is not an option. In this case, assumptions (A1) play an important role, because they determine who wins admission to the second round. In particular, a moderate candidate running alone always makes it to the second round, irrespective of whether the other moderate candidate has merged with the extremist or not. Furthermore, at the final ballot, a moderate running alone would attract all the closest extremist voters, winning the runoff election with probability 1/2. Anticipating this outcome, and given that the rents of the extremist candidates cannot fall below $\bar{r} > 0$ if a new party is formed, both moderates prefer to run alone. Hence (see the Online Appendix I for a formal proof):

Proposition 3 *Suppose that (A1) hold and that $\lambda > 1/4$. Then the unique equilibrium under runoff elections is a four-party system where all candidates run alone, and each moderate candidate wins with probability 1/2 on a policy platform coinciding with his bliss point.*

Intuitively, under runoff, voters are forced to converge to moderate platforms, because in the second round extremist candidates are eliminated from the electoral arena.

Next, consider $\lambda < 1/4$. Here if the centrist party is formed, it wins for sure on the symmetric policy $q = 1/2$. This outcome (with equal rent splitting, $r^P = R/2$) is strictly preferred to the expected outcome under the four-party equilibrium, by the same argument given in the previous subsection. Hence, if $\lambda < 1/4$, then Proposition 2 described above holds under runoff elections too. As a result, the electoral rule matters only if the political system is sufficiently polarized (i.e., $\lambda > 1/4$).

5 Extensions

This section discusses four extensions. The first three are only relevant under the runoff system: the possibility that some extremist voters are attached to their parties and do not vote for the moderate candidates in the second round; the possibility of victory at the first round if a candidate gains more than 50% of the votes; and the possibility of endorsement by the excluded parties in between the first and second round. The fourth extension discusses how to relax some of the restrictions on party formation and on the number of groups in society, emphasizing the relevant assumptions for the main qualitative results.

The Online Appendix II discusses strategic voting, and shows that this would add ambiguity to the baseline model's predictions. If strategic voters are few, the previous results do not change. And even if strategic voters are many and act as a bloc, there are equilibria in which the contrast between single round vs runoff elections described above under sincere voting continues to hold or is even stronger. Nevertheless, other equilibria are possible if many voters are strategic and if they are unevenly distributed across groups. In some of these equilibria, strategic voting blurs the sharp distinction between the two electoral rules, inducing policy moderation under single round elections, or vice versa enhancing the bargaining power of extremists under runoff elections.

5.1 Runoff elections with attached voters

Extremist voters are often very ideological and may not support a moderate party. This section investigates what happens in this case. Suppose that inside each extremist group a constant fraction $0 < \delta < 1$ of voters is ideologically "attached" to a candidate. These attached individuals vote only if "their" candidate participates on its own or as a member of a party; otherwise they abstain. This assumption plays no role under the single round, since all candidates always participate in the election, either on their own or inside a party. Hence we only consider runoff elections.

We assume that the fraction δ of attached voters is not too large, otherwise there is no relevant difference between single round and runoff elections:

$$2e/\underline{\alpha} > \delta. \tag{A2}$$

Under this assumption, merging with extremists presents a trade-off for the moderate candidates: a merger increases their chances of final victory, because it draws the support of the extremist attached voters; but if they win, they get less rents and possibly worse policies. In the single round, moderates face a similar trade-off. But it is much steeper, because the

probability of victory increases by 1/2 as a result of merging. Under runoff elections with attached voters, instead, the fall in the probability of victory is less drastic, and moderate candidates may or may not choose to run alone, depending on parameter values and on expectations about the behavior of the opponents.

Specifically, consider all possible party configurations before any voting has taken place. In the symmetric case in which no new party is formed and four candidates initially run for elections, the two moderates gain access to the second round and each moderate wins with probability 1/2. In the other symmetric case of a two-party system, each moderate-extremist coalition wins again with probability 1/2. In the asymmetric party system, instead, the Online Appendix II proves:

Lemma 1 *The probability that the moderate candidate (say 2) wins in the second round if he runs alone, given that his opponents (3 and 4) have merged, is* $1/2 - h$, *where* $h \equiv (\Pr(\eta \leq \delta\underline{\alpha}/2) - 1/2)$ *and where* $1/2 > h > 0$ *if (A2) holds.*

Thus, parameter h measures the handicap of running alone under runoff elections, given that the opponents have merged. Assumption (A2) implies that the moderate candidate has a strictly positive chance of winning in the second round if it runs alone, even if his opponents have merged. If (A2) were violated, then runoff elections would not offer any advantage to the moderate candidates, and the equilibrium would be identical to the single round. Intuitively, if the share of their attached voters is larger than any possible realization of the electoral shock, the extremist candidates retain all their bargaining power and the electoral system does not make any difference. More generally, the handicap h increases with the fraction of attached voters, δ, and the size of extremist groups, $\underline{\alpha}$, while it decreases with the range of electoral uncertainty, e.

The Online Appendix II proves that the equilibrium depends on the size of h. Suppose $\lambda > 1/4$ so that a centrist party cannot be formed. Then if h is large, the unique equilibrium is a two-party system, as in the single round, since moderates always prefer to merge with extremists, who then retain some bargaining power. If h is small, on the other hand, the unique equilibrium is a four-party system, as in the previous section; here the bargaining power of the extremists is entirely wiped out, and the runoff system induces that four-party equilibrium which was unreachable under a single round because of the polarization of the electorate. Importantly, the Online Appendix II shows that even when they form, the coalitions between moderates and extremists do so on a strictly more moderate policy platform compared to the single round case, and the more so the smaller is the handicap of running alone, h. Intuitively, the bargaining power of moderates has increased, because a runoff system gives them the option of running alone without being sure losers, so that the disagreement point is less fearsome for them. If instead $\lambda < 1/4$, the equilibrium configuration

is again a three-party system, with a centrist party that runs on the platform $q = 1/2$ and wins for sure, as this dominates any other possible outcome for the moderates.

5.2 Victory at the first round

In the Italian electoral system discussed below, like in most other runoff systems, a candidate who wins more than 50% of the votes in the first round is elected without going through a second round. How does this modified rule affect the results presented so far? To answer, note that this issue only matters in the asymmetric and off equilibrium case, where say candidates 1 and 2 have merged into a single party while the remaining candidates 3 and 4 are running alone. The reason is that in a two-party system whoever gains a plurality also wins more than 50% of the votes; and in a four-party system assumption (A1) implies that no single candidate can ever reach 50% as a vote share.

Consider then a three-party system consisting of say $\{1, 2\}$, $\{3\}$, and $\{4\}$, and suppose for simplicity that there are no attached voters. Suppose further that the shock η can be decomposed into two shocks, $\eta = \varepsilon_1 + \varepsilon_2$, each corresponding to one of the two rounds. Specifically, in the first round the size of group 2 voters is $\bar{\alpha} + \varepsilon_1$, while group 3 voters are $\bar{\alpha} - \varepsilon_1$. In the second round, the size of group 2 voters is $\bar{\alpha} + \varepsilon_1 + \varepsilon_2$, while group 3 voters are $\bar{\alpha} - \varepsilon_1 - \varepsilon_2$. Thus, ε_1 is a permanent shock. The random variables ε_1 and ε_2 are independently and identically distributed and are symmetric around a mean of 0 over the interval $\left[-\frac{e}{2}, \frac{e}{2}\right]$, for consistency with (A1). The final probability of victory—at either the first or second round—for $\{1, 2\}$ is thus the probability of winning more than 50% of the votes at the first round, $\Pr(\varepsilon_1 > 0) = 1/2$, plus the probability of getting less than 50% at the first round and yet winning at the second round, $Pr(\varepsilon_1 \leq 0, \varepsilon_1 + \varepsilon_2 > 0)$. This modified electoral rule is thus equivalent to the simpler version of runoff elections with attached voters discussed in the previous subsection, for the special case of $h = Pr(\varepsilon_1 \leq 0, \varepsilon_1 + \varepsilon_2 > 0)$. Note that, by symmetry, we also have $h = Pr(\varepsilon_1 > 0, \varepsilon_1 + \varepsilon_2 \leq 0)$. Thus, under this modified runoff system, the handicap of running alone for 3 corresponds to the risk that 3 would have won the second round ($\varepsilon_1 + \varepsilon_2 \leq 0$) and yet loses the election because $\{1, 2\}$ together gain more than 50% at the first round ($\varepsilon_1 > 0$). The Online Appendix II shows that $1/4 > h > 0$, and for the special case of a uniform distribution for both ε_1 and ε_2 we have $h = 1/8$.

5.3 Runoff elections with endorsements

Here we continue to assume that a fraction δ of extremist voters are attached and that (A1) and (A2) hold, but we allow some renegotiation to take place in between the two rounds of voting. Specifically, the excluded candidates can endorse one of the two candidates admitted

to the second round, if the latter approves. This is a common practice in many runoff systems, including municipal elections in Italy. The consequence of an endorsement is to mobilize the support of attached extremist voters, who vote for the neighboring moderate candidate in the second round only if there is an explicit endorsement by the extremist politician; otherwise they abstain. We assume that the policy cannot be renegotiated in between the two rounds; this is in line with the interpretation that the policy is dictated by the identity (ideology) of the candidate, which cannot be changed after the first round.[6] As a result of endorsing and in case of victory, the extremist candidate gets a share of the rents. For simplicity, we discuss first the case where this share is fixed at $\bar{r} > 0$; then, we consider what happens if instead rents were fully negotiable at the endorsement stage by Nash bargaining between the moderate candidate and the endorsing extremist. Before the first round of election, both policy and rents are always negotiable, as in the previous sections.

To rule out multiple equilibria, we assume that in a four-party system endorsements occur sequentially: after the outcome of the first ballot is observed, one of the two moderates is drawn at random and decides whether to accept the endorsement of his extremist neighbor; next, the other moderate decides whether to accept the endorsement of the other extremist. We only consider the case $\lambda > 1/4$, since otherwise a centrist party is always formed before the first round, as with no endorsements.

Clearly, an excluded extremist politician is always eager to endorse: by endorsing he has nothing to lose, but he gains a share of rents in the event of a victory. Furthermore, by endorsing, the extremist makes it more likely that the closer moderate candidate wins, which improves the policy outcome. The issue is whether moderate candidates seek an endorsement. They face a trade-off: an endorsement brings in the votes of the attached extremists, but reduces rents. To formally model this extension, suppose again that the shock η can be decomposed into $\eta = \varepsilon_1 + \varepsilon_2$, as in the previous subsection. Thus the first ballot reveals some relevant information about the chances of victory of one or the other moderate party in the second ballot. Suppose further that the random variables ε_1 and ε_2 are independently and identically distributed, with a uniform distribution over the interval $[-e/2,\ e/2]$. This specification is entirely consistent with that assumed for η in the previous sections.

To describe the equilibrium, we work backwards, from a situation in which the two moderate candidates have passed the first round of election (endorsements can only arise if moderates have not already merged with extremists). We then ask what this implies for merger decisions before the first round takes place.

[6] In Italian municipal elections, endorsement after the first round of voting is highly regulated by the law. The endorsing party has the right to share the winner's majoritarian prize in the city council, but he has to sign the program presented by the endorsed candidate before the first round. See Section 6 for details on the Italian institutional framework.

As shown in the Online Appendix II, whether the gain in the probability of winning is worth the dilution of rents or not depends on the realization of ε_1 relative to a threshold $\check{\varepsilon} \lesseqgtr 0$; this threshold in turn is an increasing function of the size of attached voters, $\delta\underline{\alpha}$, and a decreasing function of the rents that must be left to the extremists in case of an endorsement, $\bar{r}$. If ε_1 is below the threshold, then the probability of victory for 2 is so low that he prefers to be endorsed even if this dilutes rents. While if ε_1 is high enough, he is so confident of winning that he prefers no endorsement. And symmetrically for the other moderate, so that depending on the realization of ε_1 there may be equilibria where both moderates accept the endorsement of the extremists, both refuse, or only one accepts (see Online Appendix II).

Next, consider what happens before the first round, when moderate candidates bargain with extremists over party formation. Here, moderates lose any incentive to merge with the extremists before the first round of election. By (A1), they know that they will always make it to the second round. They also know that, after the first round, they will always be able to get the endorsement of the extremists if they wish to do so. But waiting until after the first round gives the moderates an additional option: if the shock ε_1 is sufficiently favorable, then they can run alone in the second round as well, without having to share the rents from office. This option of waiting has no costs, since the extremists are always willing to endorse at the minimal rents. Putting it differently, if rents are not negotiable ex post, the moderate's bargaining position improves, as he can commit to buy the extremist votes ex post at the minimal costs. Summing up:

Proposition 4 *Suppose that endorsements are allowed after the first round of election and rents are not negotiable at the endorsement stage. Then the unique equilibrium outcome at the first electoral round is a four-party system where all candidates run alone and each moderate candidate passes the first post with probability 1/2 on a policy platform that coincides with his bliss point. After the first round of election, endorsements by the extremists take place on the basis of the realization of the shock ε_1 as described in the Online Appendix II.*

What happens if instead rents are negotiable also at the endorsement stage, and not only ex ante? Specifically, suppose that rents are determined as the Nash bargaining outcome between the moderate and extremist, once endorsements on both sides have been decided, and where the disagreement point corresponds to the (unilateral) breakup of the coalition, as in the previous section. Here moderate candidates lose the benefit of commitment. The advantage of postponing the agreement for the moderate is still there, but now there are realizations of the shocks so adverse that the moderate candidate is in a weak bargaining position and has to pay extra rents to the extremists in order to preserve his endorsement. To avoid this risk, the moderate candidate might then prefer to strike an agreement ex ante with

the extremist, even if this means compromising on policies as well as giving up some rents. This is more likely to happen if minimum rents $\overline{r}$ are small and if the fraction of attached voters δ is large. It generally remains true, however, that the possibility of subsequent endorsements improves the bargaining position of the moderate candidate before the first round, making it more likely that moderate policies are enacted in equilibrium.[7]

5.4 Alternative party systems

This subsection discusses alternative assumptions about the distribution of voters' preferences and the party system: adding a centrist candidate; changing the relative size of moderate vs extremist groups; and relaxing some restrictions on party formation. For simplicity, throughout we consider the simplest version of the model, with no attached voters and no endorsements.

A centrist candidate. Suppose that a fifth group of voters is added to those described in Section 2. The new group is located at 1/2, has a corresponding centrist candidate $P = c$, and has size $\alpha^c = 1 - 2 \cdot (\bar{\alpha} + \underline{\alpha}) > 0$. If indifferent between two parties, the centrist voters split their vote in half. All remaining assumptions continue to hold. In particular, the shock η redistributes votes between the two moderate candidates, 2 and 3, and the two moderates are closer to the extremists than to the center ($\lambda > 1/4$). The protocol also remains as before, although here we allow for parties of up to three candidates to be formed.[8]

We make two assumptions about the size of the centrist group. First, it is not larger than the extremists: $\alpha^c \leq \underline{\alpha}$. Under this assumption, in the equilibrium with single round the extremists and the moderates always merge. The centrist candidate may also be included in one of these coalitions, if its size is large enough, but the important point is that extremists are never left out. Intuitively, the incentives for the moderates to merge remain as described in Proposition 1, and the extremist is a more attractive partner than the centrist because it is not smaller, it is ideologically closer and has less bargaining power.[9]

Second, we assume that the centrist party is not so small as to be irrelevant. In particular, suppose that the centrist group is sufficiently large that the moderates prefers to merge with the centrist candidate rather than to run alone, given that the other moderate is

[7]Details are available upon request.

[8]The disagreement point is still well defined even with three candidates inside the same party, because of the assumption that under disagreement the party is dissolved. This assumption gives the centrist party considerable bargaining power: if, at the policy formation stage, candidate c chooses to leave his party, say $\{1, 2, c\}$, then the remaining candidates 1 and 2 are forced to dissolve the party and can no longer merge into the smaller party $\{1, 2\}$.

[9]The centrist candidate has more bargaining power than the extremist because, by being closer to the opponents' equilibrium policy, it has less to fear from disagreement resulting in electoral defeats. Hence, even a smaller extremist would be a more attractive partner than the centrist candidate.

running alone. The Online Appendix II states this condition formally and proves that in the equilibrium with runoff the centrist candidate always merges with one of the moderates. The extremist candidates are always left out under runoff elections, however, and do not merge with the moderates. Intuitively, as in Proposition 3, the moderates do not need the support of extremist voters to pass the first round, and can count on getting their vote anyway in the second round. This is not the case with the centrist group, however, since in the second round centrist voters are indifferent and split their vote in half between the two moderates. Thus under runoff the moderate candidates have an incentive to merge with the centrist (which one does so depends on the order of moves in the party formation stage).

Combining these two results, we obtain that the moderating effect of runoff elections is likely to survive even with a centrist candidate. Under the stated assumptions, extremist candidates are always included in the equilibrium parties with the moderates under single round elections but not under runoff elections; and the centrist candidate is always included in a party with one of the moderates under runoff elections. These two features of the equilibrium imply a larger number of parties, and in general more moderate policies, under runoff compared to single round elections. Online Appendix II provides a formal comparison.[10]

The relative size of moderates vs extremists. Consider the same four group version of the model discussed in Section 2, but suppose now that moderates have size $\underline{\alpha}$ and extremists size $\overline{\alpha}$, with $\underline{\alpha} < \overline{\alpha}$, exactly the reverse of what we assumed in Section 2. The shock $\eta = \varepsilon_1 + \varepsilon_2$ changes the relative size of the two larger groups, now the extremists, in the same symmetric way described in Section 2. The size of the two moderate groups remains fixed at $\underline{\alpha}$. Everything else is kept unchanged, including the distribution of the shock and the bargaining protocol. It is easy to see that now runoff elections *increase* policy volatility, for the same reasons discussed above. Under runoff elections, the larger extremist candidates do not need the support of moderate voters to pass the first round, and with attached voters they retain more bargaining power.

Nevertheless, there remains a reason why runoff elections can induce policy moderation even if the moderates are smaller than the extremists. Moderates have an option that is precluded to the extremists: if $\lambda \leq 1/4$, they can merge into a centrist party. As shown in a previous version of the paper, if the distribution of the shock η is large enough relative to the size of $(\overline{\alpha} - \underline{\alpha})$, the new centrist party can pass the first round with positive probability. If this happens the centrist party wins the election, because in the second round it attracts the support of the excluded extremist. This can induce the two moderates two merge into a

[10] As shown in the Online Appendix II, at least one of the major parties runs on a more moderate platform under runoff than under single round elections. For a subtle reason, however, we cannot rule out the possibility that for some functional forms party $\{1, 2, c\}$ under single round elections runs on a more moderate policy than party $\{2, c\}$ under runoff elections, even though only the former includes the extremist candidate.

centrist party, rather than with the extremists.

Finally, what happens if we retain the assumption that moderates are larger than extremists in expected value, but not for all realization of the shock ε_1? Specifically, take the same model as in Section 5.3 (neglecting endorsements), with the shock $\eta = \varepsilon_1 + \varepsilon_2$ redistributing votes between the two moderates, but change the first condition in (A1) and assume instead that $(\overline{\alpha} - \underline{\alpha}) < \frac{e}{2}$. Under this inequality, all four candidates now have a positive probability of passing the first round under runoff elections. The equilibrium under single round elections is not affected by this change, since by the second condition in (A1) the incentives to merge remain unaltered. The equilibrium under runoff elections, instead, could be different. Since all four candidates can make it to the second round, the position of the moderate candidates is now weakened. How much so depends on the asymmetry in expected size between moderates and extremists, relative to electoral uncertainty. If the extremist candidate is relatively unlikely to pass the first round in a four party system, then runoff continue to display more policy moderation than single round elections, irrespective of the number of parties under runoff.[11] If however electoral uncertainty is large (or equivalently moderates and extremists are sufficiently similar in expected size), now the result on policy moderation could be reversed, and for some parameter values runoff could entail less policy moderation than single round elections.

Ruling out parties resulting from the merger of three candidates. Section 2 assumes that parties cannot include more than two adjacent candidates. The Online Appendix II shows that this restriction can be re-interpreted as stating that polarization (as captured by λ) is sufficiently high, or that candidates care sufficiently about policies relative to rents. Here is the intuition. If $\lambda > 1/4$ and if party $\{1, 2, 3\}$ was formed, it would have to run on a policy sufficiently close to the bliss point of candidate 3, t^3; otherwise all moderate voters in group 3 would be lost to extremist candidate 4. Of course, this constraint benefits candidate 3, but hurts candidates 1 and 2. If candidates care sufficiently about policy relative to rents and if λ is sufficiently high, then either candidate 1 or candidate 2 cannot be compensated enough for this unpleasant policy choice through a more favorable rent allocation, and party $\{1, 2, 3\}$ is not formed in equilibrium.[12]

Discussion. The gist of these extensions is that two features of voters' preferences are needed for the main result on the moderating effect of runoff elections. First, the electorate is sufficiently polarized, in the sense that the two moderate groups are sufficiently far apart

[11] If $(\overline{\alpha} - \underline{\alpha}) < \frac{e}{2}$, it is possible that the equilibrium under runoff has two parties rather than four, because now the moderates have a stronger incentive to merge so as to avoid electoral defeat either at the first or second round.

[12] The grand coalition consisting of all candidates can also be ruled out, provided that candidates are not very risk averse and that minimal rents $\bar{r}$ are sufficiently high (results available upon request).

(λ is sufficiently large), and that any centrist candidate would command the support of a relatively small group of voters. Without this assumption, a centrist party would emerge, or would lead to the exclusion of extremists as attractive partners, and there would be no difference between single round and runoff elections. Second, there are sufficiently more moderate than extremist voters. Runoff elections help the larger parties, since they exclude the smaller ones from the final electoral competition. Hence, if the extremists are larger than the moderates, runoff election increase their influence (or might do so if moderates and extremists are similar in size), with the caveat noted above, that small moderates (but not small opposite extremists) have the option of merging. This property of the distribution of voters preferences—large but polarized moderate groups of voters—is plausible, and as shown in the next section it holds in the Italian political system that we study.

6 Evidence from Italian municipal elections

Our theoretical model delivers three testable predictions: (1) runoff elections increase the number of political candidates with respect to single round elections; (2) extremist parties are less likely to merge with moderate ones under runoff; (3) policy volatility is lower—and policy moderation higher—under runoff. In this section, we use an RDD identification strategy to test predictions (1) and (3), and we provide suggestive evidence on prediction (2). We exploit a reform in municipal elections in Italy, which introduced single round vs runoff elections for municipalities of different population size.

6.1 Electoral rules for Italian municipalities

Until 1993, municipal governments in Italy were ruled by a pure parliamentary system. Citizens voted for party lists under proportional representation to elect the legislative body (i.e., the city council); the council then appointed the mayor and the executive office. Since 1993, instead, the mayor has been directly elected under plurality rule, with a single round for municipalities below 15,000 inhabitants, and with a runoff system above (see Law 81/1993).

Specifically, below this population threshold, each party (or coalition) presents one candidate for mayor and a list of candidates for the city council. Voters cast a single vote for the mayor and his supporting list (they can also express preference votes over the candidates for councillor within the same list). The mayoral candidate who gets more votes becomes mayor and his list gains 2/3 of all seats in the council. The remaining 1/3 of the seats are divided among the losing lists in proportion of their vote shares.[13]

[13]There is a minimum level that a list must obtain in order to gain seats, equal to 4% of the votes.

Above the 15,000 threshold, parties (or coalitions) present lists of candidates for the council, and declare their support to a specific candidate for mayor. Each mayoral candidate can be supported by more than one list. There are two rounds of voting. At the first round, voters cast two votes, one for a mayoral candidate and one for a party list, and the two votes may be disjoint (i.e., voters are allowed to vote for, say, mayor A and a list supporting mayor B), although both choices are expressed in the same ballot. Most voters, however, leave just one mark on the ballot, expressing a single choice either for a party list (in which case, the preference is also extended to the mayoral candidate supported by that list) or for a mayoral candidate (in which case, the preference is expressed for the candidate only). If a candidate for mayor gets more than 50% of the votes in the first round, he is elected. Otherwise, the two best candidates run against each other in a second round (taking place two weeks after the first). In this second round, the vote is only over the mayor, not the party lists. In between the two rounds, lists supporting the excluded mayoral candidates are allowed to endorse one of the remaining two candidates (if the endorsed candidate agrees). Like in the single round system, the rules for the allocation of council seats entail a majority premium for the lists supporting the winning candidate for mayor. Thus, this electoral rule is very similar to the runoff system with endorsements described in our model.

As discussed in Section 6.3, our identification strategy is valid only if there are no other (non-electoral) policies or institutions that vary at or around the threshold of 15,000 inhabitants. The closest policy thresholds based on population size are at 10,000 (where the mayor's wage, the size of the council, and the size of the executive office sharply increase) and at 30,000 inhabitants (where the mayor's wage and the size of the council sharply increase). Both thresholds are outside of our sample (see below).[14]

The 15,000 threshold entails a change in the electoral system for electing both the mayor and the city council. Thus, strictly speaking, our test concerns the consequences of both changes. Nevertheless, there are many reasons to believe that the only relevant difference is the method for electing the mayor. One of the main features and effects of the 1993 reform was the strengthening of the political power of mayors, both formally and effectively. Since 1993, Italian mayors can appoint and dismiss the executive officers at will; they also have the prerogative of appointing the city manager and shaping all municipal policies (see Law 81/1993). It is true that, if the city council approves a vote of no confidence, then the mayor is forced to step down. But this is a very rare event in Italian local politics. As a matter of fact, within the universe of mayoral elections from 1993 to 2007, only in 1.11% the mayor was removed because the council approved a vote of no confidence, and only in 1.69% because the council resigned (therefore ending the term under both circumstances).

[14]For a summary of Italian institutions varying with population, see Gagliarducci and Nannicini (2013).

It is true that this is equilibrium behavior, but it is important to stress that, whenever the mayor steps down, the legislature automatically comes to an abrupt end and new elections for both the mayor and the council are held. The direct election also gives the mayor sufficient leverage to sidestep a tiring bargaining with political parties over every single issue; since 1993 the mayor is indeed the crucial player of municipal politics in Italy.[15] This is also reflected by the fact that *all* mayoral candidates usually receive more absolute votes than their lists. Finally, the electoral rules for the council below and above the 15,000 threshold are not very different: in both cases, the system is proportional with open lists and majority premium. The only difference is that below the 15,000 threshold, but not above, the mayor is constrained to receive the support of only one list (what we call "alliance restriction" henceforth), and voters do not have the option of casting a "separate vote" for a mayoral candidate and a list supporting another candidate. Note, however, that there are no different constraints on the number of mayoral candidates. In Section 6.4, we present empirical tests supporting the claim that the above institutional variations have second-order effects with respect to the change from single round to runoff elections.

6.2 Data sources and variables

Because we do not want our estimates to be affected by observations far away from the 15,000 population threshold, and to avoid overlap with other policies, we restrict the sample to Italian municipalities between 10,000 and 20,000 inhabitants (about 10% of all Italian municipalities), for elections that took place after the 1993 reform. The complete sample is thus made up of 2,027 mayoral terms, referred to 661 towns. Both below and above the 15,000 threshold, mayoral terms lasted for four years from 1993 to 2000, and five years both before and afterwards. As explained below, in some regressions we also consider the years preceding the reform (from 1985 onwards) to implement falsification exercises.

The data refer to three kinds of variables. First, we have data on population (both from the 1991 and the 2001 Census) and other general features of the municipality, such as per capita income, geographic location, and demographic features (again, from both the 1991 and 2001 Census). The source for these data is ANCI (*Associazione Nazionale Comuni Italiani*). Second, we collected political variables at the municipal level, such as the number of candidates for mayor, vote shares, voter turnout, number of council lists, and party alliances. All these variables vary over time. Their source is the Statistical Office of the Italian Ministry of Internal Affairs. Third, we have data on the municipal tax rate on business property, provided by the Italian Ministry of Internal Affairs. This tax instrument was introduced in

[15] See Di Virgilio (2005) for evidence and discussion on the institutional features of Italian local politics.

1993, at about the same time as the electoral reform. Property taxes are the main source of local tax revenues, covering on average more than 50% of the overall municipal tax revenues. Municipal governments are free to allocate tax proceeds to a variety of alternative uses, such as social assistance, local schools, and public infrastructures. We focus on the business property tax because of its salience in the political debate at the municipal level. The partisan conflict over the appropriate level of business taxation is traditionally sharp, with left-wing candidates pushing for a higher tax rate compared to right-wing ones. In a small subsample of municipalities where we are able to identify the political orientation of the mayor, there is a strong partisan effect on the business real estate tax: on average, left-wing governments set a larger tax rate by 0.209 percentage points (+3.7% over the right-wing average tax rate of 5.665), and the difference is statistically significant at the 5% level.[16]

6.3 Empirical strategy

Formally, under the standard assumption of continuity of potential outcomes at the population threshold $P_c = 15,000$, we can identify the local average treatment effect around P_c as: $E[Y_i(1) - Y_i(0)|P_i = P_c] = \lim_{P_i \downarrow P_c} Y_i - \lim_{P_i \uparrow P_c} Y_i$, where $Y_i(1)$ is the potential outcome under runoff elections for municipality i, $Y_i(0)$ the potential outcome under single round elections for the same municipality, P_i population size (as of the last available Census), Y_i the observed outcome, and where we omit time subscripts to simplify notation (see Hahn, Todd, and Van der Klaauw, 2001). This is a local effect because it captures the causal impact of the runoff system only for towns around the threshold P_c.

The identifying assumption of continuity of potential outcomes requires that: (i) no other institutions change in a neighborhood of 15,000; (ii) municipalities did not sort around the 15,000 threshold according to their unobservable characteristics after the introduction of the new electoral law. As discussed, the first condition is met in the Italian context. We empirically check for the second condition below.

Various methods can be used to estimate the discontinuity at P_c, that is, to consistently estimate the limit of two regression functions on either side of the threshold. We apply both a spline polynomial approximation and local linear regression (see Imbens and Lemieux, 2008). The first method uses the whole sample of municipalities between 10,000 and 20,000

[16] In a multivariate regression controlling for population, margin of victory, region and time fixed effects, the impact of left-wing governments on the tax rate remains quantitatively similar and statistically different from zero at the 5% level. This is consistent with anecdotal evidence. Consider the electoral platform of *Rifondazione Comunista*, a small left-wing extremist party (approximately between 4 and 8 percent in national elections). For the municipal elections of 2004 the party platform read: "On the real estate tax, an articulated policy is needed, with the aim to reduce the rate on the first residential home for low and medium income households and increase instead the rate on second homes and business real estates."

inhabitants and chooses a flexible functional form to fit the relationship between Y_i and P_i on either side of P_c. Specifically, we estimate the model:

$$Y_i = \sum_{k=0}^{p} (\delta_k P_i^{*k}) + D_i \sum_{k=0}^{p} (\gamma_k P_i^{*k}) + \varepsilon_i, \tag{2}$$

where D_i is a treatment dummy equal to one if $P_i \geq P_c$, and the normalized variable $P_i^* = P_i - P_c$ allows us to interpret γ_0 as the jump between the two regression functions at P_c. The local average treatment effect is consistently estimated by $\hat{\gamma}_0$. Usually, a third-grade polynomial ($p = 3$) is used in the empirical literature, but we assess the robustness of the results to other functional form specifications (namely, $p = 2$ and $p = 4$).

The second method fits linear regression functions to the observations distributed within a distance h on either side of the threshold. Specifically, we restrict the sample to towns in the interval $P_i \in [P_c - h, P_c + h]$ and estimate the model:

$$Y_i = \delta_0 + \delta_1 P_i^* + D_i(\gamma_0 + \gamma_1 P_i^*) + \varepsilon_i. \tag{3}$$

Again, $\hat{\gamma}_0$ identifies the local average treatment effect. We present the robustness of the results to multiple bandwidths around P_c (namely, $h = 1,000$, $h/2$, and $2h$).

Finally, to also exploit the (limited) time variation in our data, we estimate the following diff-in-diff model:

$$Y_{it} = \alpha_i + \beta_t + \gamma_0 D_{it} + x_{it}'\rho + \epsilon_{it}, \tag{4}$$

where α_i and β_t are city and year-of-election fixed effects, respectively, while x_{it} is a vector of time-varying covariates. In this case, the identifying variation is coming from municipalities that crossed the threshold P_c between the 1991 and the 2001 Census, and the underlying assumption is that they were on a common trend with respect to the others. This assumption is less compelling than the RDD continuity condition, but we will test its plausibility with a falsification exercise on pre-1993 political outcomes.

6.4 Preliminary analysis

Before moving to the results, we discuss the plausibility of the model's assumptions in the context of Italian politics, as well as the validity of the RDD identifying assumptions.

Party system and political polarization. Our theoretical predictions on the differential impact of the runoff system on the number of candidates and policy moderation are derived under the assumptions of sufficient polarization in the electorate ($\lambda > 1/4$ in the model and the absence of a large centrist party) and of moderate parties being larger than

the extremists. These assumptions fit very well the Italian environment. Political analysts agree that the party system that emerged from the crisis of the so-called "First Republic" in the early 1990s is strongly polarized. Over our sample period, Italy displayed two main moderate blocks, one on the center-left and one on the center-right (both with an average vote share from 30 to 35 percent in national elections), plus an extremist party on the left (*Rifondazione Comunista*, from 4 to 8 percent) and one on the right (*Lega Nord*, from 4 to 10 percent). Centrist parties tried to break this structure several times, but they never got more than 6 percent as a combined vote share. The two main moderate parties never reached political agreements between them at either the national or local level.

Although there existed geographical heterogeneity and some of these parties changed their official name and symbol over time, the structure of the national party system remained fairly stable and was reflected in local politics, where other parties were rare and less influential. This is also in line with our model, as national parties can be seen as pre-existing candidates who must decide whether to merge at the local level or not. To shed more light on the stability of these party preferences just above and just below 15,000, Table A1 in the Online Appendix III estimates our RDD specifications using as outcomes the vote shares that the main parties obtained in each municipality during the 2001 national election. We use the electoral results in the proportional tier of the mixed-member electoral system in place in Italy from 1994 to 2006, because voting is more likely to be sincere there.[17] Specifically, we implement both a spline polynomial approximation as in equation (2), with polynomials of three different orders (third, second, and fourth, respectively), and local linear regression as in equation (3), with a bandwidth $h = 1,000$, as well as with half and double bandwidth. We detect no discontinuity at 15,000 in the vote shares of the moderate, extremist, and centrist parties in national elections. This means that voters' preferences are stable just below and above the threshold, and in line with the model's assumptions.

Non-attached voters. Another crucial assumption of the theory is that at least some voters are not "attached," that is, they vote for a second-best candidate in the second round if their preferred candidate did not pass the first. If all voters were attached (i.e., $\delta = 1$ in the model), then runoff and single round would yield the same equilibria. To check that this assumption is not violated by the data, we compare the votes cast in the first and second round for each runoff election that had two rounds of voting. In Figure A5 in the Online Appendix III, we plot the drop in turnout between the first and second round (on the vertical axis) against the total votes received in the first round by all the excluded candidates (on the horizontal axis); both variables are measured as a fraction of eligible voters. If the drop

[17] The mixed system in place for national elections during our sample period was 75% majoritarian and 25% proportional; see D'Alimonte (2001) for institutional details. We use data from the 2001 national election; results are qualitatively identical with the 1994 and 1996 election (available upon request).

in participation coincided with the votes for the excluded candidates, all observations should lie along the 45° line. This is obviously not the case: most of them lie well below the line, meaning that in most elections the drop in participation between the two rounds is much smaller than the votes received by the excluded candidates. Thus, the figure suggests that a large fraction of those who voted for losers in the first round vote again in the second.[18]

Strategic voting. The baseline predictions of our model are derived under the assumption of sincere voting. As discussed in the Online Appendix II, the widespread presence of strategic voters would add some ambiguity to the model's predictions. We thus propose a simple test for strategic voting in the spirit of Fujiwara (2011). The intuition is simple. Fujiwara (2011) finds a strong (negative) effect of runoff on the combined vote share of the first two candidates. This effect may be due to two different reasons: (i) voters strategically favor the first two candidates under single round elections (even if there is an equal number of candidates under single round and runoff), in line with Duverger's Law; (ii) there are more candidates under runoff and therefore the vote share of the first two is mechanically lower. Fujiwara (2011) finds support for the first interpretation in Brazil.[19]

We estimate the effect of runoff on the vote share of the top two candidates (and separately of the first, second, and third candidate), obviously in electoral races where we observe more than two candidates.[20] In particular, in Table A2 in the Online Appendix III, we implement our RDD specifications in two subsamples both above and below 15,000: electoral races with exactly three, and with more than three, candidates. If strategic voting were at work, we should observe that below the threshold voters converge on the first two candidates—at the expense of the third candidate or of the others—when there are more than two candidates. We instead find that this is not the case in Italian municipalities.

This evidence is not a conclusive test of strategic voting, because the strategic behavior of political candidates might be driven by the "anticipation" of strategic voting by the voters of some municipalities. In other words, municipalities with three or more candidates under both

[18]Under the assumptions that those who vote in the second round also participate in the first round, that those who vote for the top two candidates in the first round also participate in the second round, and that there are no endorsements, we can compute the fraction of attached voters (the parameter δ) as the ratio between the drop in participation and the votes to the excluded candidates. The median value of this ratio is about 50%. Of course, a violation in one of the above assumptions would result in an upward or downward bias in the estimate. Online Appendix Figure A5 also reveals that voting for losers in the first round is substantial, ranging from about 5% to more than 50%, with a median value around 30%. But the size of votes for losers is unrelated to the drop in turnout, which remains roughly constant at about 15% of eligible voters. This further suggests that the drop in turnout is not driven by disappointed voters.

[19]More precisely, he concludes that the first interpretation has to be preferred because the effect of runoff on the number of candidates is not precisely estimated to be different from zero (although it is positive in all of his specifications). He interprets his findings as evidence of strategic voting. But he also acknowledges that, because of sample size limitations at the Brazilian threshold of 200,000 registered voters, he cannot completely rule out strategic entry or exit by candidates too.

[20]We do not observe the individual vote shares of candidates ranked below the third position.

single round and runoff are selected samples, whose comparison should be interpreted with caution. Yet, in contrast with the Brazilian evidence, the hypothesis of widespread strategic behavior by voters finds no support in our data. This is why we interpret this evidence as supportive of our baseline assumptions in the theoretical model.

Manipulative sorting. Finally, as a preliminary check on the validity of our RDD strategy, we test for manipulative sorting around the 15,000 threshold in response to the electoral reform in 1993. In particular, in the Online Appendix Figure A5, we test if the *difference* between the density in the 1991 Census (before the reform) and the density in the 2001 Census (after the reform) shows a discontinuity at the 15,000 threshold, in the spirit of McCrary (2008). Such a discontinuity would imply that some municipalities reacted to the electoral reform by manipulating their population size, therefore violating the identifying assumption of our RDD exercise. The figure performs this test by using the density difference as outcome and fitting a 3rd-order polynomial in population size on either size of the threshold. There is no evidence of manipulative sorting between the 1991 and the 2001 Census, as the point estimate of the discontinuity is -0.007 (standard error, 0.027).

To further check against the possibility of manipulative sorting, we perform balance tests of both time-invariant and pre-treatment city characteristics. The time-invariant characteristics are geographic location, area size, and altitude from sea level. The pre-treatment characteristics come from the 1991 Census and refer to the age structure, educational attainments, employment, and house facilities. Online Appendix Table A3 uses the time-invariant variables as outcomes and estimates equation (2) with polynomials of different order and equation (3) with different bandwidths. Online Appendix Table A4 does the same with the pre-treatment variables from the 1991 Census. None of these variables displays a significant discontinuity at the threshold, and this further supports the validity of our setup.

6.5 Estimation results on political outcomes

The theory predicts that the number of candidates is larger under runoff elections than under single round. Is this consistent with the evidence? We have data on both the number of candidates for mayor and the number of (party) lists for the city council. The main outcome of interest is the number of mayoral candidates, for two reasons. First, this is what the theory has predictions about. Second, the number of lists may reflect both different electoral rules and different restrictions above and below 15,000: as already mentioned, below the threshold there has to be a one-to-one correspondence between lists and mayoral candidates, whereas above 15,000 mayoral candidates can be supported by more than one list. Nevertheless, comparing the number of lists is also relevant, particularly because it allows an intertemporal comparison: before 1993 the mayor was not directly elected and we only have data on lists.

Here, we use data on all 2,027 mayoral terms pooled together, because the outcome of interest (the number of candidates) is time-varying. To accommodate for possible serial correlation, we cluster the standard errors at the city level. Treatment assignment depends on population size as measured by the last available Census, that is, either 1991 or 2001. On average, in municipalities between 15,000 and 20,000, 5.1 candidates run for mayor, as opposed to 3.6 in municipalities between 10,000 and 15,000. The electoral lists supporting the candidates for mayor are 6.9 above 15,000 and 3.7 below.

Clearly, these differences in the number of candidates and lists might be confounded by other factors associated with population size. To identify the causal effect of the electoral rule separately from the effect of city size, we thus implement our RDD strategy as discussed in Section 6.3. In Table 1, we report the main estimates of the impact of runoff elections on political outcomes. We implement both a spline polynomial approximation as in equation (2), with polynomials of three different orders, and local linear regression as in equation (3), with three different bandwidths. In panel A we report the baseline results, while in panel B we also add city characteristics as control variables (namely, macro-region dummies, area size, altitude, per-capita transfers, per-capita income, labor force participation, elderly index, family size, mayor's duration in office, and a dummy identifying second-term mayors). As long as these additional covariates are balanced around the population threshold, their inclusion should not affect the estimates, but increases accuracy.

The results in Table 1 show a positive and statistically robust effect of runoff on the number of mayoral candidates. Just above the threshold, we observe approximately one more candidate and two more lists. If we look at the baseline estimate of 1.103 in column 1, runoff elections produce a 29% increase in the number of candidates relative to single round elections just below the threshold. The impact on the number of lists is even greater (+51%), but, as said, it is confounded by the alliance restriction.

To assess the relevance of this institutional restriction, in Table 1 we estimate the effect of the electoral rule on the average number of lists supporting each mayoral candidate (captured by the ratio "lists/candidates"). We also estimate the distinct effects of the electoral rule on the number of lists supporting the winning candidate vs the losing candidates. The estimated effect on the ratio between lists and candidates is not statistically different from zero, meaning that the alliance restriction is not binding. This is confirmed by the fact that the effect on the number of lists supporting the losing candidates is statistically significant, but there is no significant discontinuity in the number of lists supporting the elected mayor. As a matter of fact, this last outcome variable can only be affected by the restriction on feasible alliances, as the winning candidate is one by definition, both above and below the threshold, while the result on the number of lists supporting the losing candidates can be driven by the number of

candidates itself.[21] The (zero) result on the number of lists supporting the winning candidate (i.e., the future mayor) is also reassuring about the fact that the confounding institutional variations at 15,000 do not affect the future bargaining within the municipal government coalition in ways that are not captured by our model.

Figure 1 provides a visual illustration of the results on political outcomes (first four graphs). There, we report both the scatterplot of each outcome (averaged over 250-inhabitant intervals) and the spline third-order polynomial (with the 95% confidence interval). The discontinuities of political outcomes at the threshold are clearly visible both from the scatterplots and from the estimated polynomials, with the exception of the number of lists supporting the winning candidate, for which we have no significant results as expected. Clearly, the RDD setup allows for identification only in a neighborhood of 15,000, but the positive association between runoff and the number of candidates persists far away from the threshold.

In Table 2, we run a falsification test on the only political outcome available for the pre-treatment period. If the sorting before 1993 (if any) were associated with potential outcomes, a discontinuity in the pre-treatment number of party lists should show up in the data. As before 1993 a parliamentary system was in place, we can only run our falsification test on the number of electoral lists. Table 2 reports the RDD estimates for all mayoral terms elected from 1985 to 1992, and for municipalities between 10,000 and 20,000 inhabitants. No significant discontinuity is detected. Before the 1993 electoral reform, the number of lists was exactly equal just below and just above the 15,000 threshold. This provides strong evidence in favor of the robustness of the baseline results.

To further assess the sensitivity of our results, the Online Appendix Figure A7 summarizes a set of 1,000 placebo estimates at false thresholds for the main outcomes. Specifically, to evaluate the possibility that our results arise from random chance rather than a causal relationship, we implement estimations at false population thresholds below and above the 15,000 threshold (namely, any point from 13,501 to 14,000 and from 15,501 to 16,000 in

[21] This argument can be illustrated by means of a simple example with two scenarios. Under the *first*, candidates A (supported by list A1) and B (supported by B1) run in single round elections; A wins. In the counterfactual case of runoff, candidates A (supported by A1 and A2) and B (supported by B1 and B2) run; A wins. Here, the runoff system has no (causal) impact on the number of candidates (2 in both systems) but increases the average number of lists per candidate (2 as opposed to 1). The confounding channel that drives this effect is that more lists are supporting the same number of candidates under runoff (because there is no restriction by law). This can be also estimated by looking at the number of lists supporting the winning candidate: 2 lists under runoff (A1 and A2) vs 1 under single round (A1), meaning that the estimated effect of runoff on the number of mayor's lists should be positive. Consider, instead, a *second* scenario. Candidates A (supported by A1) and B (supported by B1) run in single round; A wins. Candidates A (supported by A1), B (supported by B1), and C (supported by C1) run in runoff; A wins. Here, runoff has a causal impact on the number of candidates (3 as opposed to 2) but no effect on the number of mayor's lists (which is always 1). Our estimates clearly indicate that this second scenario, and not the first, is confirmed by the data. This means that the (confounding) change in the alliance restriction is not driving our results.

order to stay away from the true threshold). At these false thresholds, we expect to find no systematic evidence of treatment effects similar to our baseline results. For each outcome, the figure reports the cumulative distribution function of the 1,000 placebo point estimates (using a specification with spline 3rd-order polynomial), normalized with respect to the baseline point estimates from Table 1. This means, for instance, that a normalized coefficient of 100 stands for a placebo point estimate equal to the true baseline estimate at 15,000. Thus, most normalized coefficients should be close to zero, and we should observe only a few normalized coefficients outside the interval [-100, +100]—in fact no more than 5% in each tail. Indeed, only 1.6% of the placebo estimates are larger than the baseline result for the number of candidates in absolute value (but they have the opposite sign), and none of the placebo estimates exceed the baseline result for the number of lists and the number of opposition lists. All cumulative distribution functions are steeper around zero, where the false estimates tend to concentrate. By contrast, and again as expected, there are no robust results for the number of lists supporting the mayor.[22]

Finally, in Table 3, we implement diff-in-diff estimations on political outcomes as in equation (4). As discussed, the identifying variation comes from municipalities crossing the population threshold from the 1991 to the 2001 Census, under the restriction that movements from above to below and vice versa have symmetric effects. Again, the empirical evidence is in line with the model's predictions, as point estimates for all political outcomes are quantitatively similar to the RDD results.[23]

Overall, we conclude that the evidence on political outcomes strongly supports the theoretical prediction concerning the number of candidates in single round vs runoff elections.

6.6 Estimation results on policy volatility

In this section, we test the main prediction of the theory, on policy moderation. Ideally, we would like to test whether extremist parties are more often included in the municipal government coalition, and exert more policy influence, under single round elections. Unfortunately,

[22]The same holds for the lists/candidates ratio (available upon request).

[23]In Appendix Table A5, we remove the symmetry restriction and separately look at the effect of moving from below to above 15,000 (33 municipalities) vs moving from above to below (9 municipalities) in a cross-section of municipalities for which political outcomes are available both in the 1990s and in the 2000s. The two effects are very similar and again in line with the theoretical predictions: municipalities that moved to the runoff system in the 2000s experienced an increase in the number of candidates by 27%; those that moved to the single round system experienced a drop by 34%. These (symmetric) results do not support Callander's (2005) claim that the effect of runoff should depend on initial conditions. Furthermore, Online Appendix Table A5 allows us to evaluate the diff-in-diff assumption of common trend, as in the last row we estimate whether municipalities that crossed the threshold are associated with different pre-treatment levels of political competition (i.e., number of party lists) in the 1980s. This falsification exercise supports the identifying assumption that population variations were sufficiently exogenous.

we cannot do that because of data limitations (although we say something about the point below). In several municipalities and particularly in small ones, candidates for mayor are supported by electoral lists that use local labels—such as the name of the city or electoral slogans. This is how the lists are named in the official data, and this is why we cannot identify the exact political identity of each electoral list in most cases.

Instead, we directly test another prediction of the theory on policy outcomes: whether the average policy volatility is lower in municipalities above 15,000 inhabitants, where the runoff system should moderate the influence of extremist voters. Indeed, a change in the partisan identity of the local government should be associated with a smaller policy change in those municipalities where the extremist parties are excluded from government or less influential. This test, of course, assumes that political turnover is the same above and below the threshold; something that we test and cannot reject (see below).

Policy volatility. We measure policy volatility in two ways. First, we consider the *intertemporal* variation in the business property tax rate. To do this, we measure the unconditional variance of the tax rate across legislative terms in the same municipality. Thus, for each municipality, we average the yearly tax rates over the mayoral term, excluding election years to avoid the overlapping of different mayors over the same calendar year and possible electoral cycle effects. Let τ_t^i denote this average tax rate for municipality i and the mayoral term initiated in year t. We then compute the unconditional variance of these average tax rate across mayoral terms for each municipality, $y^i = Var(\tau_t^i)$, obtaining one observation (i.e., one measure of volatility) per municipality.[24]

Next, we consider the *cross-sectional* variation in the business property tax, within bins of municipalities of similar population size ("similar" meaning within intervals of 100 inhabitants). Specifically, we first compute the same average tax rate τ_t^i defined above, for each municipality i and each mayoral term t. For each term t and each bin b we then compute the unconditional variance of τ_t^i across municipalities of the same bin, $y_t^b = Var(\tau_t^i)$. Finally, for each bean b we compute the simple average of these variances across mayoral terms, and obtain a cross sectional variance for each bin, $y^b = E(y_t^b)$.[25]

[24]Municipalities that crossed the threshold from the 1991 to the 2001 Census are included twice (once for each electoral system), while the others are included once. For policy volatility, we do not repeat the diff-in-diff analysis, because the time interval before and after the 2001 Census entails too few mayoral terms to reliably compute different tax volatility measures for each subperiod (i.e., before and after 2001).

[25]The average frequency of municipalities within each bin is around 27, with the minimum value equal to 4 and the maximum to 56. In the two bins just below and just above the 15,000 threshold, the average frequency is around 25 municipalities per bin. All of the following results are qualitatively similar with bin sizes of 10 inhabitants (about 5 municipalities in each bin) and of 50 inhabitants (about 15 municipalities), and they are available upon request. At the price of reducing the outcome variation, we prefer a size of 100 inhabitants because in this case the unconditional variance is more precisely estimated within each bin.

The RDD results are reported in Table 4, for both indicators of volatility. The intertemporal variance of the business property tax shows a sharp and negative discontinuity when moving from just below to just above the 15,000 threshold. Point estimates are consistently negative and statistically significant at standard levels, although they are more volatile than with political outcomes. The baseline estimate of -0.455 in column 1 corresponds to a decrease of about 61% in the variance of the tax rate just above the threshold. Similar results hold for the cross-sectional variance. Here, all estimates are by weighted least squares (with weights based on the frequency of municipalities in each bin) to account for heteroskedasticity and to accommodate for the different accuracy in the estimation of the variance in bins of different numerosity. The baseline estimate of -0.659 in column 1 indicates that, in a neighborhood of the threshold, the runoff system decreases the variance of the property tax by about 71%, compared to single round elections. Point estimates are stable when comparing specifications without and with covariates (panel A vs panel B).[26]

A graphical representation of the results on policy volatility is provided in Figure 1 (last two graphs), where the negative discontinuities at the threshold are evident both in the scatterplots and in the estimated polynomials. These effects appear to be more local—that is, less persistent far away from the threshold—compared to those on political outcomes, but we cannot assign any causal interpretation to the association between population size and policy volatility once we move away from the institutional cutoff at 15,000.

Online Appendix Figure A7 (again, last two graphs for the policy volatility measures) implements placebo estimations at false thresholds. Results on both the time and cross-sectional volatility of the tax rate are very robust, as only 2.7% (3.5%) of the false estimates are larger than the baseline one for the cross-sectional (time) variance in absolute value.

Overall, the evidence provided above is strongly consistent with the prediction of the theory that runoff elections induce smaller policy volatility, compared to the single round.

Potential channels. The above results are reduced-form effects. There remains the concern that the lower tax volatility under runoff elections could be driven by other channels, rather than policy moderation. In particular, the electoral system could affect the level of political turnover, by influencing the probability of government crises (through a vote of no confidence by the council) or the probability of political swings between left and right administrations. In the estimates with covariates (panels B in Table 4), we already control for this channel by including two proxies of political turnover (namely, the duration in office of the elected mayor and whether he reaches a second term or not). Nevertheless, we can

[26]There is instead some sensitivity of the point estimates to the functional form of the polynomial and to the estimation method. This might also reflect measurement error in the unconditional variance of the tax rate in relatively small samples. On average, there are only 4 mayoral terms from which the intertemporal variance is computed, and the cross-sectional variance is computed from bins of heterogeneous size.

directly test whether political turnover is affected by the electoral system. Table 5 reports the RDD estimates on the two observable outcomes associated with political turnover: the average duration in office (measured in days) and the fraction of mayors in their second term. None of these outcomes shows a significant discontinuity at the 15,000 threshold, and the point estimates display no consistent pattern. This rules out the most plausible alternative explanation of our (reduced-form) results.

Finally, to provide some direct evidence on the political extremism channel, we estimate the effect of runoff elections on the probability that the leftist political extreme (i.e., the Communist Party, *Rifondazione Comunista*) joins the main center-left coalition at the local level.[27] The Italian Ministry of Internal Affairs provides details on the party lists supporting different candidates for mayor in the first round. We manually coded these data to create a dummy variable (*Communist Party alone*) that equals one in elections where the Communist Party ran either alone or allied with other smaller leftist parties (e.g., *La Rete*, *Verdi*, *Pdci*), but not with the more moderate and larger center-left party of the time (e.g., *DS*, *Ulivo*). Here, as discussed above, we face a key problem, because in several municipalities mayoral candidates are supported by electoral lists with local names. These lists may correspond to national political parties or not; we simply do not know. After dropping the municipalities with this missing information, we are left with a (self-selected) sample that is only half the original sample (i.e., 1,045 observations, of which 670 are below the threshold). Another limitation is that in some municipalities where we observe a center-left coalition but we do not observe the Communist Party running alone, it could be either because this extremist party joined the main coalition, or because it was not organized in those municipalities. Both instances are coded as zero in our dummy variable of interest. Measurement error due to the self-declared nature of the data could also be an issue, although we do not expect it to bias the results in a predetermined direction.

Table 6 reports RDD estimations where the dependent variable is the dummy *Communist Party alone* (which equals one in about 11% of the elections in the small sample). Point estimates are large and positive, as expected, and they are statistically significant at standard levels with most estimation methods. On average, the probability that the Communist Party runs alone in the runoff system more than double compared to the single round.

On the whole, the quasi-experimental and descriptive evidence discussed in this section supports the conclusion that runoff systems indeed induce policy moderation, because they dampen the influence of extremist parties or exclude them from government coalitions.

[27]The same exercise cannot be replicated for the center-right coalition, where the extremist parties are either too small at the local level (e.g., *Msi*, *La Destra*), or geographically concentrated in some areas of the country and focused on separatist issues (e.g., *Lega Nord*).

7 Concluding remarks

Political extremism is often regarded as harmful, because it enhances policy uncertainty and it hinders the effective functioning of democracies (e.g., Bingham Powell, 1982). Knowing which political institutions can alleviate the adverse consequence of political extremism is therefore important. This is particularly true for young democracies, where often extremism is rampant and democratic constitutions have to be designed from scratch.

This paper has compared single round vs runoff elections from this perspective. With a polarized electorate and large moderate parties, the runoff system reduces the influence of political extremes. This happens because the runoff allows moderate parties to pursue their own policy platform without being forced to compromise with the neighboring extreme. This also implies that the number of political candidates is larger under runoff than single round elections. The evidence from Italian local elections is consistent with these predictions. In particular, municipalities just above 15,000 inhabitants (which rely on runoff elections) have a larger number of candidates and less volatile tax rates, compared to municipalities just below 15,000 inhabitants (which have single round elections).

While our results clearly speak in favor of runoff electoral systems, one should also be careful not to overstress this implication. An important limitation of our work is that we do not study the long run effects on the number of parties and on their ideological positions. One may argue that the runoff system, while reducing the influence of political extremism on policies, allows extremist parties to survive by giving them visibility at the first round. In normal times this is of no concern; but in times of crisis this can make a difference, allowing extremists to gain consensus. Under a single round system, instead, extremists would join some bigger party and in the long run they would disappear inside this party. Conversely, we do not consider the possibility that single round elections create an excessively high barrier to entry of new parties, and this may restrict the menu choices of voters in the long run - something that could be an issue in case of large shocks to voters' preferences. We leave these important issues for future research.

References

[1] Axelrod, R.M., 1970. *Conflict of Interest: A Theory of Divergent Goals with Applications to Policy*, Markham, Chicago.

[2] Battigalli, P. , 1996. "Strategic Independence and Perfect Bayesian Equilibrium," *Journal of Economic Theory*, 70(1), 201–234.

[3] Bingham Powell, G. Jr., 1982. *Contemporary Democracies. Participation, Stability, and Violence*, Harvard University Press, Cambridge MA.

[4] Bouton, L., 2013. "A Theory of Strategic Voting in Runoff Elections," *American Economic Review*, 103(4), 1248–1288.

[5] Bouton, L., Gratton, G., 2013. "Majority Runoff Elections: Strategic Voting and Duverger's Hypothesis," mimeo, Boston University.

[6] Callander, S., 2005. "Duverger's Hypothesis, the Run-off Rule, and Electoral Competition," *Political Analysis*, 13, 209–232.

[7] Castanheira, M., 2003. "Why Vote for Losers?," *Journal of the European Economic Association*, 1(5), 1207–1238.

[8] Chamon, M., de Mello, J.M.P., Firpo, S., 2009. "Electoral Rules, Political Competition and Fiscal Expenditures: Regression Discontinuity Evidence from Brazilian Municipalities," IZA DP 4658.

[9] Cox, G., 1997. *Making Votes Count*, Cambridge University Press, Cambridge UK.

[10] D'Alimonte, R., 2001. "Mixed Electoral Rules, Partisan Realignment, and Party System Change in Italy',' in M.S. Shugart and M.P. Wattenberg (eds.), *Mixed-Member Electoral Systems: The Best of Both Worlds?*, Oxford University Press.

[11] Di Virgilio, A., 2005. "Il sindaco elettivo: un decennio di esperienze in Italia," in Caciagli, M., Di Virgilio, A. (eds.), *Eleggere il sindaco. La nuova democrazia locale in Italia e in Europa*, UTET, Torino.

[12] Engstrom, R.L., Engstrom, R.N., 2008. "The Majority Vote Rule and Runoff Primaries in the United States," *Electoral Studies*, 27(3), 407–16.

[13] Feddersen, T., 1992. "A voting model implying Duverger's Law and Positive Turnout," *American Journal of Political Science*, 36(4), 938–962.

[14] Fey, M., 1997. "Stability and Coordination in Duverger's Law: Formal Model of Preelection Polls and Strategic Voting," *American Political Science Review*, 91, 135–147.

[15] Fey, M., 2007. "Duverger's Law without Strategic Voters," mimeo, University of Rochester.

[16] Fiorina, M.P., 2005. *Culture War? The Myth of a Polarized America*, Longman.

[17] Fisichella, D., 1984. "The Double Ballot as a Weapon against Anti-System Parties," in Lijphart, A., Grofman, B. (eds.), *Choosing an Electoral System: Issues and Alternatives*, Praeger, New York.

[18] Fujiwara, T., 2011. "A Regression Discontinuity Test of Strategic Voting and Duverger's Law," *Quarterly Journal of Political Science*, 6, 197–233.

[19] Gagliarducci, S., Nannicini, T., 2013. "Do Better Paid Politicians Perform Better? Disentangling Incentives from Selection," *Journal of the European Economic Association*, 11, 369–398.

[20] Golder, M., 2005. "Democratic Electoral System around the World, 1946–2000," *Electoral Studies*, 24, 103–21.

[21] Hahn, J., Todd, P., Van der Klaauw, W., 2001. "Identification and Estimation of Treatment Effects with Regression Discontinuity Design," *Econometrica*, 69, 201–209.

[22] Imbens, G. W., Lemieux, T., 2008. "Regression Discontinuity Designs: A Guide to Practice," *Journal of Econometrics*, 142(2), 615–635.

[23] Kreps, D., Wilson, R., 1992. "Sequential Equilibria," *Econometrica*, 50, 863–94.

[24] Levy, G. , 2004. "A Model of Political Parties," *Journal of Economic Theory*, 115(2), 250–277.

[25] Lizzeri, A., Persico, N., 2005. "A Drawback of Electoral Competition," *Journal of the European Economic Association*, 3, 1318–48.

[26] McCrary, J., 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test," *Journal of Econometrics*, 142, 698–714.

[27] Messner, M., Polborn, M., 2004. "Robust Political Equilibria under Plurality and Runoff Rule," mimeo.

[28] Morelli, M., 2004. "Party Formation and Policy Outcomes under Different Electoral Systems," *Review of Economic Studies*, 71, 829–853.

[29] Myerson, R., Weber, R., 1993. "A theory of Voting Equilibria," *American Political Science Review*, 87 (1), 102–114.

[30] Osborne, M.J., Slivinsky, A., 1996. "A model of Political Competition with Citizen-Candidates," *Quaterly Journal of Economics* , 111, 65–96.

[31] Riker, W.H., 1982. "The two Party System and Duverger's Law: An Essay on the History of Political Science," *American Political Science Review*, 76, 753–766.

[32] Sartori, G., 1994. *Comparative Constitutional Engineering*, New York University Press, New York NY.

[33] Wright, S.G., Riker, W.H., 1989. "Plurality and Runoff Systems and Numbers of Candidates," *Public Choice*, 60, 155–175.

Tables and figures

Table 1 – Impact of runoff on political outcomes, RDD estimates

	Spline 3rd	Spline 2nd	Spline 4th	LLR (h)	LLR $(h/2)$	LLR $(2h)$
			A. Estimations without covariates			
No. of candidates	1.103***	1.098**	1.532***	1.300***	1.731**	1.335***
	(0.382)	(0.487)	(0.302)	(0.408)	(0.676)	(0.293)
No. of lists	2.184***	1.497**	2.163***	1.736***	1.739**	2.291***
	(0.526)	(0.624)	(0.413)	(0.540)	(0.794)	(0.436)
Lists/candidates	0.020	-0.145	-0.014	-0.147	-0.118	0.063
	(0.201)	(0.239)	(0.151)	(0.208)	(0.305)	(0.170)
Opposition lists	1.657***	1.050**	1.676***	1.426***	1.147**	1.643***
	(0.374)	(0.448)	(0.297)	(0.384)	(0.569)	(0.299)
Mayor's lists	-0.016	-0.110	-0.031	-0.209	-0.016	0.044
	(0.226)	(0.252)	(0.184)	(0.231)	(0.307)	(0.196)
Obs.	2,027	2,027	2,027	364	175	761
			B. Estimations with covariates			
No. of candidates	1.220***	1.147**	1.598***	1.331***	1.779**	1.418***
	(0.375)	(0.472)	(0.297)	(0.396)	(0.677)	(0.287)
No. of lists	2.260***	1.463**	2.254***	1.695***	2.025***	2.406***
	(0.517)	(0.610)	(0.405)	(0.557)	(0.735)	(0.423)
Lists/candidates	0.031	-0.159	0.001	-0.154	-0.116	0.076
	(0.186)	(0.220)	(0.140)	(0.184)	(0.290)	(0.153)
Opposition lists	1.717***	1.055**	1.746***	1.388***	1.358**	1.733***
	(0.384)	(0.462)	(0.301)	(0.392)	(0.560)	(0.302)
Mayor's lists	-0.014	-0.111	-0.024	-0.202	0.058	0.055
	(0.216)	(0.240)	(0.177)	(0.226)	(0.326)	(0.187)
Obs.	2,027	2,027	2,027	364	175	761

Notes. Election years between 1993 and 2007; municipalities between 10,000 and 20,000. Dependent variables: *No. of candidates* running for mayor in the first round; *No. of lists* supporting mayoral candidates in the first round; *Lists/candidates* ratio; *Opposition lists* supporting the losing candidates; *Mayor's lists* supporting the winning candidate. Estimation methods: spline polynomial approximation as in equation (2), with 3rd, 2nd, and 4th polynomial, respectively; local linear regression as in equation (3), with bandwidth $h = 1,000$, $h/2$, and $2h$, respectively. Estimations in Panel B also include the following covariates: macro-region dummies, area size, altitude, transfers, income, participation rate, elderly index, family size, mayor's duration in office (in days), mayor's second-term dummy. Robust standard errors clustered at the city level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 2 – Falsification tests on pre-treatment political outcomes, RDD estimates

	Spline 3rd	Spline 2nd	Spline 4th	LLR (h)	LLR $(h/2)$	LLR $(2h)$
			A. Estimations without covariates			
No. of lists	-0.178	-0.231	-0.121	-0.033	-0.128	-0.336
	(0.449)	(0.544)	(0.349)	(0.506)	(0.668)	(0.365)
Obs.	783	783	783	137	67	284
			B. Estimations with covariates			
No. of lists	-0.034	-0.202	-0.124	0.069	0.070	-0.244
	(0.348)	(0.419)	(0.290)	(0.351)	(0.502)	(0.292)
Obs.	783	783	783	137	67	284

Notes. Election years between 1985 and 1992; municipalities between 10,000 and 20,000. Dependent variable: *No. of lists*, i.e., party lists competing under proportional representation in this pre-treatment period (1985–1992). Estimation methods: spline polynomial approximation as in equation (2), with 3rd, 2nd, and 4th polynomial, respectively; local linear regression as in equation (3), with bandwidth $h = 1,000$, $h/2$, and $2h$, respectively. Estimations in Panel B also include the following covariates: macro-region dummies, area size, altitude, transfers, income, participation rate, elderly index, family size, mayor's duration in office (in days), mayor's second-term dummy. Robust standard errors clustered at the city level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 3 – Impact of runoff on political outcomes, diff-in-diff estimates

	A. Estimations without covariates	B. Estimations with covariates
No. of candidates	1.186***	1.159***
	(0.300)	(0.300)
No. of lists	2.303***	2.259***
	(0.394)	(0.392)
Lists/candidates	0.284*	0.282*
	(0.170)	(0.170)
Opposition lists	1.787***	1.746***
	(0.308)	(0.308)
Mayor's lists	0.143	0.152
	(0.181)	(0.181)
Obs.	2,027	2,027

Notes. Election years between 1993 and 2007; municipalities between 10,000 and 20,000. Dependent variables: *No. of candidates* running for mayor in the first round; *No. of lists* supporting mayoral candidates in the first round; *Lists/candidates* ratio; *Opposition lists* supporting the losing candidates; *Mayor's lists* supporting the winning candidate. Estimation methods: diff-in-diff specifications with municipality and year-of-election fixed effects, as in equation (4). Estimations in column B also include the following (time-varying) covariates: transfers, income, participation rate, elderly index, family size. Robust standard errors are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 4 – Impact of runoff on policy volatility, RDD estimates

	Spline 3rd	Spline 2nd	Spline 4th	LLR (h)	LLR $(h/2)$	LLR $(2h)$
			A. Estimations without covariates			
Time variance	-0.455**	-0.647***	-0.238*	-0.651**	-0.697*	-0.378**
of business property tax	(0.182)	(0.240)	(0.140)	(0.255)	(0.389)	(0.160)
Obs.	575	575	575	118	59	236
Cross-sectional variance	-0.659**	-0.937***	-0.313	-0.694**	-0.364	-0.443**
of business property tax	(0.258)	(0.294)	(0.201)	(0.256)	(0.590)	(0.203)
Obs.	92	92	92	19	9	37
			B. Estimations with covariates			
Time variance	-0.450***	-0.614***	-0.237*	-0.563***	-0.167	-0.377***
of business property tax	(0.170)	(0.224)	(0.132)	(0.211)	(0.167)	(0.140)
Obs.	575	575	575	118	59	236
Cross-sectional variance	-0.627**	-0.856***	-0.352*	-0.736**	-0.832	-0.371*
of business property tax	(0.276)	(0.306)	(0.199)	(0.274)	(0.278)	(0.184)
Obs.	92	92	92	19	9	37

Notes. Election years between 1993 and 2007; municipalities between 10,000 and 20,000. Dependent variables: *Time variance* (i.e., variance across terms averaged over the entire sample period) and *Cross-sectional variance* (i.e., variance across municipalities averaged over bins of 100 inhabitants) of the business property tax rate. Estimation methods: spline polynomial approximation as in equation (2), with 3rd, 2nd, and 4th polynomial, respectively; local linear regression as in equation (3), with bandwidth $h = 1,000$, $h/2$, and $2h$, respectively. When the dependent variable is the cross-sectional variance, estimates are by weighted least squares, with weights given by (the inverse of) the numerosity of each bin. Estimations in Panel B also include the following covariates: macro-region dummies, area size, altitude, transfers, income, participation rate, elderly index, family size, mayor's duration in office (in days), mayor's second-term dummy. Robust standard errors clustered at the city level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 5 – Impact of runoff on political turnover, RDD estimates

	Spline 3rd	Spline 2nd	Spline 4th	LLR (h)	LLR $(h/2)$	LLR $(2h)$
			A. Estimations without covariates			
Office duration	46.205	53.998	-9.315	26.305	37.703	13.074
	(84.679)	(109.911)	(61.303)	(99.559)	(154.764)	(65.940)
Second term	-0.052	0.033	0.016	0.015	-0.011	-0.018
	(0.070)	(0.088)	(0.050)	(0.080)	(0.122)	(0.055)
Obs.	2,027	2,027	2,027	364	175	761
			B. Estimations with covariates			
Office duration	67.896	44.074	-14.651	21.255	62.105	5.052
	(73.369)	(92.901)	(54.787)	(79.217)	(124.915)	(58.560)
Second term	-0.050	0.031	0.012	0.007	-0.014	-0.019
	(0.069)	(0.088)	(0.049)	(0.080)	(0.126)	(0.054)
Obs.	2,027	2,027	2,027	364	175	761

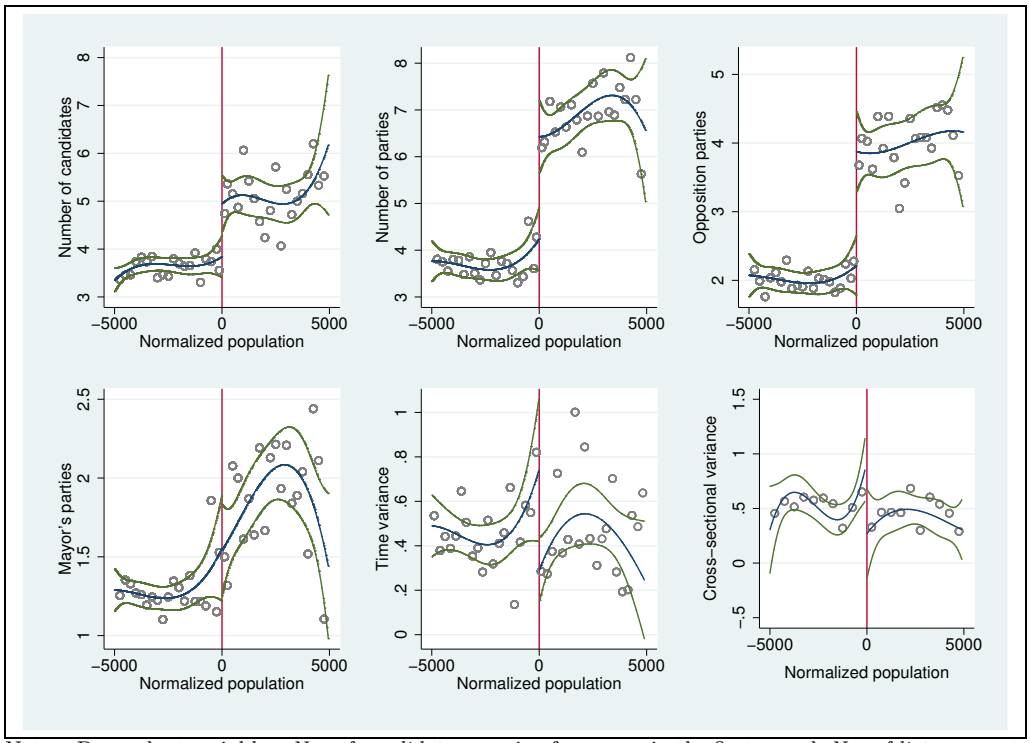
Notes. Election years between 1993 and 2007; municipalities between 10,000 and 20,000. Dependent variables: *Office duration* of mayors, measured in days; fraction of mayors in their *Second term*. Estimation methods: spline polynomial approximation as in equation (2), with 3rd, 2nd, and 4th polynomial, respectively; local linear regression as in equation (3), with bandwidth $h = 1,000$, $h/2$, and $2h$, respectively. Estimations in Panel B also include the following covariates: macro-region dummies, area size, altitude, transfers, income, participation rate, elderly index, family size. Robust standard errors clustered at the city level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Table 6 – Impact of runoff on Communist Party's alliances, RDD estimates

	Spline 3rd	Spline 2nd	Spline 4th	LLR (h)	LLR $(h/2)$	LLR $(2h)$
			A. Estimations without covariates			
Communist Party alone	0.172**	0.258**	0.069	0.221**	0.230*	0.131*
	(0.086)	(0.101)	(0.068)	(0.091)	(0.136)	(0.072)
Obs.	1,045	1,045	1,045	198	96	404
			B. Estimations with covariates			
Communist Party alone	0.167**	0.244***	0.081	0.216**	0.200*	0.126*
	(0.081)	(0.094)	(0.063)	(0.083)	(0.118)	(0.066)
Obs.	1,045	1,045	1,045	198	96	404

Notes. Election years between 1993 and 2007; municipalities between 10,000 and 20,000. Dependent variable: the dummy *Communist Party alone* is equal to one if the Communist Party presented its own list (or some electoral alliance with smaller leftist parties) in the first round of the municipal election, and zero otherwise. Estimation methods: spline polynomial approximation as in equation (2), with 3rd, 2nd, and 4th polynomial, respectively; local linear regression as in equation (3), with bandwidth $h = 1,000$, $h/2$, and $2h$, respectively. Estimations in Panel B also include the following covariates: macro-region dummies, area size, altitude, transfers, income, participation rate, elderly index, family size, mayor's duration in office (in days), mayor's second-term dummy. Robust standard errors clustered at the city level are in parentheses. Significance at the 10% level is represented by *, at the 5% level by **, and at the 1% level by ***.

Figure 1 – Impact of runoff on political outcomes and policy volatility



Notes. Dependent variables: *No. of candidates* running for mayor in the first round; *No. of lists* supporting mayoral candidates in the first round; *Opposition lists* supporting the losing candidates; *Mayor's lists* supporting the winning candidate; *Time variance* (i.e., variance across terms averaged over the entire sample period) and *Cross-sectional variance* (i.e., variance across municipalities averaged over bins of 100 inhabitants) of the business property tax rate. The central line is a spline 3rd-order polynomial in the normalized population size (i.e., population minus 15,000); the lateral lines represent the 95% confidence interval of the polynomial. Scatter points are averaged over 250-inhabitant intervals. Municipalities between 10,000 and 20,000 only.