Single Round vs Runoff Elections under Plurality Rule: A Theoretical Analysis*

Massimo Bordignon† Tommaso Nannicini‡ Guido Tabellini§

October 2016

Abstract

We compare single round vs runoff elections under plurality rule, allowing for partly endogenous party formation. With large and sufficiently polarized groups of moderate voters, 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. These results are robust to several extentions.

*We thank Pierpaolo Battigalli, Carles Boix, Daniel Diermeier, Massimo Morelli, Giovanna Iannantuoni, Francesco de Sinopoli, Ferdinando Colombo, Piero Tedeschi, Per Petterson-Lindbom, 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. 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.

†Def, Università Cattolica del Sacro Cuore; CESifo. E-mail: massimo.bordignon@unicatt.it.

‡IGIER, Bocconi University; IZA. E-mail: tommaso.nannicini@unibocconi.it.

§IGIER, Bocconi University; CIFAR; CEPR; CESifo. E-mail: guido.tabellini@unibocconi.it.

1 Introduction

An important difference between electoral rules is whether citizens vote once or twice. In the majority of elections, there is only a single round. An increasing number of countries, however, is turning to a dual ballot (or *runoff*) mechanisms to select a winner. Under the runoff, citizens vote twice. First they vote over several candidates, and then they vote again in a second round among the candidates who received more votes in the first round. According to Borman and Golder (2013), the majority of presidential elections around the world now uses this system, with the French system for electing the President of the Republic probably being the best known example. With variants, the runoff is also used in many countries to select party candidates in gubernatiornal primaries (such as in the US), heads of executive in regional and state elections and mayors of large cities (Italy, Brazil). The runoff is not only limited to presidential elections. For instance, Italy just approved a new proportional electoral rule for the Lower House that assigns an absoulte majority of seats in a runoff between the two parties that received more votes in the first round.

Despite its increasing popularity, not much is known about the effect of the runoff system, particularly on policy choices. Traditionally, political scientists group the runoff system together with proportional systems, since it does not provide incentives to candidates to gain plurality in the first round. According to the seminal work by Duverger (1954), this implies a larger number of serious candidates running in the first round than under a first pass the post system; Riker (1982) termed this prediction Duverger's "hypothesis", in contrast with Duverger's "law" that instead states that a first pass the post system should have only two serious candidates running. But how the prediction on the number of candidates translates in a prediction in terms of policy choices is not obvious. and has not been extensively studied.[1]

In this paper, we contrast single ballot versus dual ballot electoral systems under plurality rule, focusing on the formation of parties and on bargaining among policies between pre-existing candidates. We consider a model with sincere voting where citizen candidates set a one-dimensional policy before the elections. 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, in line with Duverger's hypothesis.

[1] 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 (Myerson and Weber, 1993; Fey, 1997). As discussed below, less is known about the runoff.

Second, and more important, if the electorate is polarized and yet extremist groups are not the majority, the runoff system reduces the influence of extremist candidates and voters on the equilibrium policy, thereby inducing more centrist policies. Thus, conditional on turnover, the runoff system reduces harmful policy volatility and is preferable in terms of ex-ante social welfare. The positive predictions of this model are in line with the empirical evidence in Bordignon et al. (2016) - see below.

The reason for these results 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 when the second vote is cast enough extremist voters are willing to vote for the closest moderate, rather than abstain. Our analysis also clarifies that these results hold under general but not universal features of the political and electoral systems. In particular, the results are fairly robust to the details of the electoral system, the possibility of endorsement between the two rounds, the number of parties or candidates, the distribution of voters' preferences, and assumptions on voters' behavior. However, the results require the presence of large and sufficiently polarized groups of moderate voters.

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, 2015; Solow, 2015). Our results on the number of parties under the two rules 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 configuration 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 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 also small. The most relevant paper is Bordignon et al. (2016). There we contrast two systems for electing Italian mayors: single round plurality rule in cities with less than 15 000 inhabitants, and runoff under plurality rule above that population threshold. Estimating by regression discontinuity design (RDD), we find that the runoff system increases the number of candidates and reduces policy volatility, in line with the predictions of this paper. Other relevant empirical references are discussed in Bordignon et al. (2016).

Earlier work, using identification strategies based on conditional independence, 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). Two recent studies have applied RDD to elections in Brazil. Chamon et al. (2009) show that runoff has a larger number of effective candidates; Fujiwara (2011) finds support for Duverger's argument of strategic voter behavior under single round elections, and—unlike Chamon et al. (2009)—he estimates an effect of the electoral rule on the number of candidates that is not statistically different from zero in most specifications.

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 (all proofs are in Online Appendix I). Section 5 discusses a number of extensions (all proofs are in Online Appendix II). Section 6 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 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:

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

The first assumption in (A1) implies that, for any realization of the shock η, any moderate group is always larger than any extreme 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 extreme groups. We discuss the effects of relaxing these assumptions below. The realization of η becomes known at the elections 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. We discuss strategic voting in section 5.

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:

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

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 expectations 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.

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.[2]

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[3]. This gives moderate candidates an advantage (see below). Section 5 discusses how to relax some of these restrictions.

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

[2] Morelli (2002) and Levy (2004) use similar assumptions to explain the role of parties in politics.

[3] See again Morelli (2002) 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.

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). The possible partitions are either two parties, $\{1,2\}, \{3,4\}$; or 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 break up of the party (given that all 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 break up. 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 in detail the runoff system.

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 break up 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 break up 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 break up. 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. In a four candidates situation, each moderate would gain by a unilateral deviation that led him to form a party with his extremist neighbor, since this would guarantee victory at the elections. 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 ($\lambda > \frac{1}{4}$), 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 rent allocation is set by Nash bargaining within each party, with the disagreement point given by the party break up, 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 ballot. Once a party structure is determined, it cannot be changed in any direction in between the two ballots. We also retain assumptions (A1), together with the assumption of sincere voting. We relax all these assumptions in the next section.

As with a single round, the equilibrium depends on how polarized is the electorate. If voters are very polarized (if $\lambda > 1/4$), 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 or has not merged with his extremist neighbor. 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 $\bar{r} > 0$ whenever a new party is formed, both moderates prefer to run alone. Hence (see 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.*

This result is very intuitive. Under the runoff system, 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. Thus, the electoral rule matters only if the political system is sufficiently polarized.

5 Extensions

This section discusses several extensions to the basic model. 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 explores the robustness of our results to strategic voting. The last one discusses how to relax some of the restrictions on party formation and number of groups in society, emphasizing the relevant assumptions for the main qualitative results.

5.1 Runoff elections with attached voters

Extremists 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 last 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, Online Appendix II proves:

Lemma 1 *The probability that the moderate candidate (say 2) wins in the final 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.

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, 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 most runoff systems, a candidate who wins more than 50% of the vote share 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 vote share; and in a four party system, assumption A1 implies that no single candidate can ever reach 50% of the vote.

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 permament 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 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 $h = (Pr(\varepsilon_1 + \varepsilon_2 \leq 0 \ \& \ \varepsilon_1 > 0)$. Thus, under this modified runoff system, the handicap of running alone for 3 correponds 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 ($\varepsilon_1 > 0$). 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-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. 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 our interpretation that the policy is dictated by the identity (ideology) of the candidate, which cannot be changed after the first round.[4] 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

[4] There might also be institutional reasons. For instance, in the Italian municipal elections, endorsement after the first round of voting is highly regulated by the electoral law. The endorsing party has the right to share the winner's majoritarian prize in the council, but he has to sign the program presented by the endorsed candidate before the first round.

extremist. Before the first ballot, 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 ballot (endorsements can only arise if moderates have not already merged with extremists). We then ask what this implies for merger decisions before the first ballot takes place.

As shown in 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 the moderate candidate 2 is so low that he prefers to be endorsed even if this dilutes his rents. While if ε_1 is high enough, he is so confident of winning that he prefers no endorsement. And symmetrically for the moderate 3, 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 elections. 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 voting and rents are not negotiable at the endorsement stage. Then the unique equilibrium outcome at the first electoral ballot 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 elections, endorsements by the extremists take place on the basis of the realization of the shock ε_1 as described in the online Appendix.*

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) break up 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.

5.4 Strategic voters

Our results so far rely on the assumption of sincere voting. Do they survive if voters instead vote strategically?[5]. Suppose that a fraction $0 \leq s \leq 1$ of voters in each group

[5]Estimations about the relative importance of strategic voting vary largely in the literature. Degan and Merlo (2006) estimate that only 3% of individual voting profiles are inconsistent with sincere voting in US

J behaves strategically, while the remaining ones vote sincerely. Strategic voters take into account the probability of victory of each candidate, and may thus vote for a less preferred candidate who is more likely to win or pass the post. This depends on the beliefs about the voting behavior of all other voters. We study a Nash equilibrium where each strategic voter maximizes expected utility, given correct beliefs about the equilibrium behavior of all the others.[6] Strategic voting may affect our previous results because candidates, by correctly anticipating the voting equilibrium, might be induced to change their choices concerning merger with other candidates and/or proposed policy platforms. We continue to assume $\lambda > 1/4$.

Strategic voting in single round elections. Here there are several equilibria, some of which replicate our previous results with sincere voting, while others produce different results. In particular, it is possible to prove that, even if all voters are strategic ($s = 1$), there is a two party equilibrium in which extremist candidates exert even more influence on policy than under sincere voting.

Specifically, suppose that the voting stage is reached with four parties: $\{1\}, \{2\}, \{3\}, \{4\}$. With strategic voting and symmetry, equilibrium implies that only two parties (one on each side of 1/2) have a positive probability of victory, and that for both, this probability is 1/2. But which parties (whether extremists or moderates) depends on voters beliefs. Suppose that voters coordinate on the following sunspot equilibrium with symmetric beliefs: with equal probabilities, either all votes converge on the extremist parties on each side ($\{1\}$ and $\{4\}$), or they converge on the moderate parties on each side ($\{2\}$, and $\{3\}$). In the first case, it is optimal for all voters in groups 1 and 2 to vote for candidate 1, in the second to vote for candidate 2, and symmetrically for voters in groups 3 and 4. Then, in a four party system, each candidate wins with probability 1/4.

Suppose instead that the voting stage is reached with three parties, say $\{1\}, \{2\}, \{3, 4\}$. In line with the previous assumption, suppose that here too voters in groups 1 and 2 coordinate on a sunspot equilibrium with the same symmetric beliefs as above, namely with equal probabilities either all votes converge on party $\{1\}$ or they converge on party $\{2\}$. Again, voters in groups 1 and 2 find it optimal to validate these beliefs, so that, if

elections. Sinclair (2005) estimates a bigger fraction in the UK, but still of limited empirical relevance. Kawai and Watanabe (2012) use heterogeneity of Japanese municipalities in electoral districts to estimate both the share of potential strategic voters and the share of *misaligned voting* (voters that effectively cast a vote for a candidate different from the most preferred). While the former are up to 85% of all voters, the latter is only 1,4% - 4,2% of votes. Spenkuch (2013), exploiting the simultaneous presence of both a list vote and a candidate vote in German national elections, reaches similar conclusions. In our companion paper on Italian municipalties, we also show that widespread strategic voting is not supported by our data.

[6] This is the standard definition of a voting equilibrium with strategic voters (Myerson and Weber, 1993). For an alternative approach, see Myatt (2007). See also Cox (1997) and Bouton (2012) for a runoff model with strategic voters.

this three parties equilibrium is reached, party $\{3,4\}$ wins with probability 1/2, while $\{1\}$ and $\{2\}$ win each with probability 1/4. The same outcome occurs (in reverse) in the party system $\{1,2\},\{3\},\{4\}$.

Finally, suppose that the voting stage is reached with parties $\{1\},\{2,3\},\{4\}$. Here, given $\lambda > 1/4$, a plausible set of beliefs is that voters on both sides of 1/2 coordinate on the extremist candidates, so that parties $\{1\}$ and $\{4\}$ each win with probability 1/2.

Repeating the steps in the proof of Proposition 1 on the bargaining game between candidates, it can then be verified that a similar equilibrium still holds. Namely, under these beliefs, the equilibrium is a two-party system, where rents are split in half inside each coalition, and the policy platforms are set at the mid point between the bliss points of moderates and extremists on each side of 1/2 (i.e., $q = (t^e + t^m)/2$, where e and m denote the extremist and moderate candidate respectively).[7] Note that, with these voters' beliefs, the extremist candidates have more bargaining power and hence more influence than in the equilibrium with sincere voting described in Proposition 1. The reason is that here the extremist candidates have a chance of winning the election on their own (in fact they have the same chance as the moderate candidates). Both candidates continue to have an incentive to merge (since the sunspot creates uncertainty about who has a chance of victory if running alone); but the symmetry in the sunspot realizations enhances the bargaining power of the extremist relative to sincere voting[8].

This is not the only possibility, however. Suppose that a fraction $s > 1 - \frac{2e}{\alpha}$ of voters is strategic. Then there is also another equilibrium where, irrespective of the number of parties, a strategic extremist voter would vote for the moderate candidate because she expects all other strategic voters to do the same. Realizing this, each moderate candidate prefers to run alone or to merge with the extremist on a policy platform more moderate than in the equilibrium with sincere voting, depending on the size of s. Indeed, given these beliefs, the equilibrium under single round elections is perfectly analogous to the runoff equilibrium with attached voters described in Online Appendix I, except that we need to replace δ (the fraction of attached voters) with $1 - s$ (the fraction of sincere voters) in the definition of h in Lemma 1. Intuitively, here the extremist strategic voters in single round

[7] The proof follows the same steps as that of Proposition 1. The first order condition that pins down the equilibrium policy platform and rent allocations is the same as (4) in **Online Appendix I.** But here equation (3) is replaced by:

$$u^P(q, r^P) - \bar{u}^P = \frac{1}{2}[V(r^P) - C(|q - t^P|)] - \frac{1}{4}[V(R) - C(t^2)]$$

for $P = 1, 2$.

[8] Of course, with different and non symmetric sunspot uncertainty, either the moderate or the extremist candidate could have more bargaining power.

elections behave like the non-attached voters under runoff elections with sincere voting. The moderate candidates thus know that they can capture some of the votes of the extremists even if running alone, and this reduces the extremists' bargaining power (or induces the moderates to run alone if s is large enough).

Strategic voting in runoff elections. Here strategic voting only bites in the first round, since in the second round with only two candidates strategic voters always find it optimal to vote sincerely. This immediately implies that the equilibrium with sincere voting in Proposition 3 remains an equilibrium even under strategic voting. To see this, note that, even if all voters are strategic, there is always a voting equilibrium in the first round where the two moderates pass the post with probability 1. Given this outcome and the absence of strategic voting in the second round, the proof of Proposition 3 immediately follows. In particular, the beliefs described above under single round elections are not compatible with equilibrium under runoff elections, if voters within each group can coordinate amongst themselves and act as a bloc (i.e. if they are bloc-strategic voters). Specifically, consider a four party system and suppose that (at the first ballot) all extremist voters vote for their own candidate. Then it cannot be optimal for the moderate voters as a bloc to also vote for the extremists, since by voting for their own moderate candidate, the two moderates pass the first round even without the support of the extremists. Hence, the sunspot beliefs described above are not consistent with any equilibrium under runoff elections.

Here too, however, other equilibria are possible, for some special configuration of parameters and if the fraction of strategic voters is not the same in all groups. Specifically, consider the model with attached voters, and suppose that the first round voting stage is reached with three candidates, say $\{1\}, \{2\}, \{3,4\}$. Here, provided that the attached voters are many, the strategic voters of groups (3,4) may find it optimal to converge part of their votes on candidate 1, so that this candidate rather than 2 reaches the final ballot with certainty. The reason is that, with many attached voters and more attached voters in group 2 than in group 1, party {3,4} wins for sure against candidate 1 in the second round.[9] For this first round outcome to be incentive compatible, however, strategic voters in group 1 must accept it without shifting their vote towards candidate 2; this may happen if the fraction of strategic voters in group 1 is sufficiently smaller than in groups 3 and 4.[10] Anticipating this result at the first round, candidate 2 is then induced to seek an agreement with 1 even at the price of an extremist policy platform. This example is rather special, of course, but it reverts the previous results, that runoff elections weaken the bargaining power of extremists

[9] A sufficient condition for this to happen is that $\delta\overline{\alpha} > 2e$. This behavior by voters in groups 3 and 4 is known as "push over" in the relevant literature; see Bouton and Gratton (2013).

[10] This can only happen if, given that all voters in groups 3 and 4 are strategic, the share of strategic voters in group 1 does not exceed $\frac{1}{6} + \frac{4}{3}e$.

and induce policy moderation.

Summing up, strategic voting adds considerable ambiguity to the predictions of our model. If strategic voters are few, nothing changes with respect to our previous results. 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, 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.[11]

5.5 Alternative party systems

Our previous results were derived in a model with symmetric distribution of voters'preferences and specific assumptions on the number of candidates and party formation. In this subsection, we ask whether they are robust to a change in these assumptions. Specifically, we consider the case of a centrist candidate, a change in the relative size of moderate vs extremist groups, and we relax 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 $\alpha^c = 1 - (\bar{\alpha} + \underline{\alpha})/2 > 0$. 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 bargaining protocol among candidates also remains as before, although here we allow for parties of up to three candidates to be formed.[12]

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 moderates candidates always merge. The centrist candidate may also

[11] Not all these equilibria would survive suitable refinements of the equilibrium notion. For instance, Bouton and Gratton (2015) are able to rule out "push over" behavior in runoff elections by imposing strict perfection on equilibria.

[12] 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. Observe that 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 $\{1, 2\}$.

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.[13]

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 running alone. Online Appendix II states this condition formally and proves that in the equilibrium with runoff elections 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 offers a formal comparison.[14]

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*

[13]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.

[14]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.

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 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 sub-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 $(\overline{\alpha} - \underline{\alpha}) < \frac{\epsilon}{2}$. 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.[15] 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

[15] If $(\overline{\alpha} - \underline{\alpha}) < \frac{\epsilon}{2}$, 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.

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.[16]

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 (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. Indeed, as shown in Bordignon et al. (2016), it characterized Italian national and municipal politics for the last twenty years.

6 Concluding remarks

This paper has contrasted single round vs runoff elections. Our conclusion is that with a polarized electorate and large moderate parties, the runoff system reduces the influence of political extremes. This is welfare improving, because political extremism increases policy uncertainty and can disrupt decision making in governments or legislatures. The intuitive reason behind our results is that the runoff allows moderate parties to pursue their own policy platforms 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. While our basic setup is very simple, we have also shown that these results are robust to several extensions, including alternative assumptions on the details of the electoral rules, the number of parties and on voters' behaviour. What they require is the presence

[16] 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.

of large and sufficiently polarized groups of moderate voters. As shown in Bordignon et al. (2016), the evidence from Italian municipal elections supports these predictions. Municipalities where the mayor is elected under the runoff (those just above 15,000 inhabitants) have a larger number of candidates and less policy volatility, compared to municipalities that rely on single round plurality rule (those just below 15,000 inhabitants).

7 References

Bordignon, Massimo, Tommaso Nannicini, and Guido Tabellini. 2016. "Moderating Political Extremism: Single Round versus Runoff Elections under Plurality Rule: Dataset." *American Economic Review.*

Bormann, Nils-Christian, and Matt Golder. 2013. "Democratic Electoral Systems around the World, 1946–2011." *Electoral Studies* 32 (2): 360–69.

Bouton, Laurent, and Gabriele Gratton. 2015. "Majority Runoff Elections: Strategic Voting and Duverger's Hypothesis." *Theoretical Economics* 10 (2): 283–314.

Bouton, Laurent. 2013. "A Theory of Strategic Voting in Runoff Elections." *American Economic Review* 103 (4): 1248–88.

Callander, Steven. 2005. "Duverger's Hypothesis, the Run-Off Rule, and Electoral Competition." *Political Analysis* 13 (3): 209–32.

Chamon, Marcos, João M. P. de Mello, and Sergio Firpo. 2009. "Electoral Rules, Political Competition and Fiscal Expenditures: Regression Discontinuity Evidence from Brazilian Municipalities." *IZA DP 4658.*

Cox, Gary W. 1997. *Making Votes Count: Strategic Coordination in the World's Electoral Systems.* Cambridge, UK: Cambridge University Press.

Duverger, Maurice. 1954. *Political Parties.* John Wiley & Sons.

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

Fey, Mark. 2007. "Duverger's Law without Strategic Voters," *mimeo*, University of Rochester.

Fisichella, Domenico. 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.

Fujiwara, Thomas. 2011. "A Regression Discontinuity Test of Strategic Voting and Duverger's Law." *Quarterly Journal of Political Science* 6: 197–233.

Lizzeri, Alessandro and Nicola Persico. 2005. "A Drawback of Electoral Competition," *Journal of the European Economic Association*, 3, 1318–48.

Messner, Matthias, and Mattias Polborn. 2004. "Robust Political Equilibria under Plurality and Runoff Rule," *mimeo.*

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

Osborne, Martin J., and Al Slivinsky. 1996. "A Model of Political Competition with Citizen-Candidates." *Quaterly Journal of Economics* 111 (1): 65–96.

Riker, William H. 1982. "The Two-Party System and Duverger's Law: An Essay on the History of Political Science." *American Political Science Review* 76 (4): 753–66.

Sartori, Giovanni. 1994. *Comparative Constitutional Engineering.* New York University Press, New York NY.

Solow, Benjamin. 2015. "Aggregate Uncertainty in Runoff Elections." Unpublished.

Wright, Stephen G., and William H. Riker. 1989. "Plurality and Runoff Systems and Numbers of Candidates," *Public Choice*, 60, 155–175.

Online Appendix I

Main proofs

Proof of Proposition 1

i) Suppose that (3,4) have merged and have agreed to the policy platform $q^{34} \in [t^3, t^4]$. If 1 and 2 run alone, they lose the election with certainty and get the utility:

$$\bar{u}^P = -C(|q^{34} - t^P|), \qquad P = 1, 2 \tag{2}$$

Let $U^P(q, r^P)$ denote candidate P utility (for $P = 1, 2$) if 1 and 2 merge into a single party and agree to the policy q and rent allocation r^P. By (2) and eq. (1) in the paper, we have:

$$U^P(q, r^P) - \bar{u}^P = \frac{1}{2}[V(r^P) - C(|q - t^P|) + C(|q^{34} - t^P|)] \tag{3}$$

Note first of all that, for any $q \in [t^1, t^2]$ and any $r^P \geq \bar{r}$, and since $\lambda > 1/4$, the RHS of (3) is strictly positive. Hence both 1 and 2 are always strictly better off by forming the party $\{1, 2\}$ than by running alone against $\{3, 4\}$. This implies that, if (3,4) have merged (or are expected to merge) into a single party, candidates 1 and 2 will also merge, irrespective of the sequence of proposals.

To determine q and r^P we solve for the Nash bargaining equilibrium. Thus, we solve $Max_{q,r^1,r^2}(U^1(q, r^1) - \bar{u}^1)(U^2(q, r^2) - \bar{u}^2)$ subject to the constraints that $R \geq r^1 + r^2$, $r^P \geq \bar{r}$ for $P = 1, 2$ and $q \in [t^1, t^2]$.

At an interior optimum the first order conditions imply:

$$\frac{U^1(q, r^1) - \bar{u}^1}{U^2(q, r^2) - \bar{u}^2} = \frac{V_r(r^1)}{V_r(r^2)} = \frac{C_q(|q - t^1|)}{C_q(|t^2 - q|)} \tag{4}$$

Given strict convexity of C and since $q^{34} > t^2 > t^1$, it is easy to verify that $U^1(q, r^1) - \bar{u}^1 > U^2(q, r^2) - \bar{u}^2$ at the the symmetric outcome, $r^P = R/2$ and $q = (t^1 + t^2)/2$ - intuitively, the extremist has more to lose from disagreement since his bliss point is further away from q^{34}. Furthermore, at the symmetric outcome, both ratios on the RHS of (4) are equal to 1. Hence this cannot be an equilibrium. The Nash bargaining equilibrium must entail $r^2 > r^1$ and $q > (t^1 + t^2)/2$ so that all three ratios in (4) exceed unity (note that the left-most side of (4) is decreasing in q, decreasing in r^2 and increasing in r^1). Finally, with enough concavity in $V(.)$ and enough convexity in $C(.)$, the solution to the Nash bargaining equilibrium must be an interior optimum. By symmetry, a similar conclusion applies to Nash bargaining between 3 and 4 (with the appropriate changes), given that 1 and 2 have merged.

ii) Next suppose that 3 and 4 have not merged (or are expected not to merge). In this case, if 1 and 2 also run alone, the two moderate candidates win with probability 1/2 each on a policy plartfrom corresponding to their respective bliss point. In this case, the expected utility of 1 and 2 respectively is:

$$\bar{u}^1 = -\frac{1}{2}C(|t^2 - t^1|) - \frac{1}{2}C(|t^3 - t^1|) \tag{5}$$

$$\bar{u}^2 = \frac{1}{2}V(R) - \frac{1}{2}C(|t^3 - t^2|) \tag{6}$$

Again, let $U^P(q, r^P)$ denote candidate P utility (for $P = 1, 2$) if 1 and 2 merge into a single party and agree to the policy q and rent allocation r^P. Now party $\{1, 2\}$ wins with certainty on any feasible policy platform q, given that 3 and 4 are running alone. We thus have:

$$U^1(q, r^1) - \bar{u}^1 = V(r^1) - C(|q - t^1|) + \frac{1}{2}[C(|t^2 - t^1|) + C(|t^3 - t^1|)] \tag{7}$$

$$U^2(q, r^2) - \bar{u}^2 = V(r^2) - C(|t^2 - q|) - \frac{1}{2}V(R) + \frac{1}{2}C(|t^3 - t^2|)] \tag{8}$$

At an interior optimum, the Nash bargaining outcome between 1 and 2 must still satisfy (4) above. Repeating the same logic as above, evaluate the left-most expression of (4) at the symmetric outcome, $r^P = R/2$ and $q = (t^1 + t^2)/2$. Here too, at the symmetric outcome we have $U^1(q, r^1) - \bar{u}^1 > U^2(q, r^2) - \bar{u}^2$. Hence, by the same argument as above, the Nash bargaining equilibrium must again favor the moderate candidate, so that $r^2 > r^1$ and $q > (t^1 + t^2)/2$ even if 3 and 4 have not merged.

Finally, consider the first stage of party formation. Suppose that 3 and 4 have not merged (or are expected not to merge). We want to show that both 1 and 2 are better off merging into a single party, given that once they have done so the policy and rent allocation will be set according to the Nash bargaining outcome just described. Note that the RHS of (7) is positive since $|q - t^1| < |t^2 - t^1|$ and $C(.)$ is convex. Hence not surprisingly player 1 is better off by merging than running alone. Consider the RHS of (8) evaluated at the Nash bargaining outcome. Since in the Nash bargaining outcome $r^2 > R/2$ (and the function $V(.)$ is concave), we have $V(r^2) - \frac{1}{2}V(R) > 0$. Moreover, we also have $\frac{1}{2}C(|t^3 - t^2|) = \frac{1}{2}C(2\lambda) > C(\lambda) > C(|t^2 - q|) = C(\frac{1}{2} - \lambda - q)$ where the first inequality follows from convexity of $C(.)$ and the second inequality follows from $\lambda > \frac{t^2}{2} = \frac{1}{4} - \frac{\lambda}{2} > t^2 - q$ (since $\lambda > 1/4$ and $q > (t^1 + t^2)/2 = \frac{t^2}{2}$). Hence, the RHS of (8) is also strictly positive at the Nash bargaining outcome, and the moderate candidate too is better off merging (given the anticipated Nash bargaining outcome) rather than running alone, if 3 and 4 have not (or will not merge).

Combining (i) and (ii), we conclude that forming a coalition of the moderate and extremist candidate is a dominant strategy in the first stage, irrespective of the behavior of the opponents. Hence in equilibrium both $\{1,2\}$ and $\{3,4\}$ will form, and the Nash bargaining outcome is as described in part (i) of the proof. QED

Proof of Proposition 3

Suppose that 3 and 4 are running alone (or are expected to do so). If 1 and 2 also run alone, then 2 and 3 win with probability 1/2 each, on a policy platform corresponding to their respective bliss points. Hence, the expected utility of 2 in this case is given by (6) above. If 1 and 2 merge, then their probability of victory remains 1/2 irrespective of the policy q, since, given $\lambda > 1/4$ and sincere voting, in the second round the moderate candidate 3 is able to capture all extremist voters in group 4. Hence, the expected utility of candidate 2, given that he has merged with 1 and that 3 and 4 are running alone, is:

$$U^2(r^2, q) = \frac{1}{2}V(r^2) - \frac{1}{2}C(|t^2 - q|) - \frac{1}{2}C(|t^3 - t^2|) \tag{9}$$

Note that $r^2 \leq R - \bar{r}$ (where $\bar{r} > 0$ denotes the minimal rents that must be given to candidate 1 when party $\{1,2\}$ is formed). Comparing (9) and (6), we see that candidate 2 is strictly better off running alone than under the merger, for any q (even for $q = t^2$). Hence, even if candidate 1 would be better off under a merger, there is nothing that he can offer to the candidate 2 to convince him to merge.

Next, suppose that 3 and 4 have merged (or are expected to merge) and run on a policy platform q^{34}. The probability of final victory for candidate 2 in the final ballot is 1/2, irrespective of whether he has merged with 1 or not, since in any case he can collect the votes of extremist voters close to him. Hence, the expected utility of 2 if he runs alone is given by:

$$\bar{u}^2 = \frac{1}{2}V(R) - \frac{1}{2}C(|q^{34} - t^2|)$$

and his expected utility if he merges with 1 on a policy platform q is:

$$U^2(r^2, q) = \frac{1}{2}V(r^2) - \frac{1}{2}C(|t^2 - q|) - \frac{1}{2}C(|q^{34} - t^2|)$$

Comparing these two expressions and repeating the same argument as above, we see that candidate 2 is always better off running alone than merging with 1, for any policy q.

Given the model's symmetry, the only equilibrium of the runoff thus has both moderate candidates running alone on a policy that coincides with their respective bliss points. QED

Online Appendix II

Extensions

Runoff system with attached voters

Proof of Lemma 1

Suppose that candidates 3 and 4 have merged, while candidate 2 runs alone. Consider the second round of voting. Given the behavior of the attached extremists in group 1, candidate 2 wins if:

$$(1-\delta)\underline{\alpha}+\overline{\alpha}+\eta > \underline{\alpha}+\overline{\alpha}-\eta \tag{10}$$

or more succinctly if:

$$\eta > \delta\underline{\alpha}/2$$

Since η is distributed over the interval $[-e, e]$, this event has probability :

$$1-\Pr(\eta \leq \delta\underline{\alpha}/2) = 1/2 - h$$

and $1/2 > h > 0$, where the first inequality follows from $\delta\underline{\alpha}/2 > 0$ and the second inequality is implied by (A2). QED

We now describe the equilibrium.

Proposition 5 *Suppose that* (*A*1), (*A*2) *hold and that* $\lambda > 1/4$. *Define*

$$\bar{h} = \frac{V(R)-V(R-\bar{r})}{2[V(R-\bar{r})+C(|t^3-t^2|)}$$

(i) If $h < \bar{h}$, *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 that coincides with his bliss point and grabs all the rents if he wins.*

(ii) If $h > \bar{h}$, *then the unique equilibrium under runoff elections is a two party system where moderates and extremists merge on both sides and each party wins with probability 1/2. In this case, the equilibrium policies under runoff are always closer to the moderate candidates' bliss points, and moderate candidates get a larger share of rents if elected, than in the equilibrium under single round elections. Moreover, the smaller is* h, *the closer are the equilibrium policies under runoff election to the moderate candidates' bliss points, and the larger are the rents that go to the moderates if elected.*

(iii) If $h = \bar{h}$, the equilibrium might either be a four party system or a two party system. In both cases, the equilibrium policies will coincide with the bliss points of the moderates.

Proof

We repeat the steps in the proof of Proposition 3, but now taking into account the attached voters. Throughout we assume $\lambda > 1/4$ and that $(A1)$, $(A2)$ hold.

Suppose that 3 and 4 have not merged (or are expected not to merge). In this case, if 1 and 2 also run alone, the two moderate candidates win with probability 1/2 each on a policy platform corresponding to their respective bliss point, and their expected utility are still given by (5) and (6) respectively.

If 1 and 2 merge into a single party, they win with probability $1/2 + h$,and their expected utility can then be written as:

$$U^P(q, r^P; h) = (\frac{1}{2} + h)[V(r^p) - C(|q - t^P|)] - (\frac{1}{2} - h)C(|t^3 - t^P|), \quad P = 1, 2 \tag{11}$$

where $U^P(.)$ is now expressed also as a function of h. Consider candidate 2, and evaluate (11) at his most favorable policy and rent allocation, namely $q = t^2$ and $r^2 = R - \bar{r}$. He is indifferent between merging with 1 on these terms or running alone if:

$$U^2(t^2, R - \bar{r}; h) - \bar{u}^2 = 0 \tag{12}$$

Now solve (12) for h, and denote the solution by $\bar{h}$. Using (6) and (11) we get:

$$\bar{h} = \frac{V(R) - V(R - \bar{r})}{2[V(R - \bar{r}) + C(|t^3 - t^2|)}$$

For $h < \bar{h}$, candidate 2 prefers to run alone, and given the indivisibility of rents below $\bar{r}$, there is nothing that candidate 1 can do to induce him to merge. For $h > \bar{h}$, instead, the electoral advantage of merging is sufficiently large that candidate 2 is willing to merge with 1 for at least some feasible policy q and rent allocation, given that 3 and 4 run alone. Repeating the same procedure for candidate 1, it is easy to verify that 1 is always willing to merge with 2 on the terms most favorable for the latter (intuitively, he stands to gain the minimal rents and a higher probability of a policy closer to his bliss point). By symmetry, the same results holds for candidate 3, given that 1 and 2 run alone.

Now suppose that 3 and 4 have merged (or are expected to do so) on a policy platform of q^{34}. If 1 and 2 also merge, they win with probability 1/2, and their expected utility is given by

$$U^P(q, r^P) = \frac{1}{2}[V(r^P) - C(|q - t^P|) - C(|q^{34} - t^P|)] \tag{13}$$

If instead they run alone, then candidate 2 wins with probability $(1/2 - h)$ while candidate 1 has no chances. Hence their expected utilities are respectively:

$$\bar{u}^1(h) = -(\frac{1}{2} - h)C(|t^2 - t^1|) - (\frac{1}{2} + h)C(|q^{34} - t^1|) \tag{14}$$

$$\bar{u}^2(h) = (\frac{1}{2} - h)V(R) - (\frac{1}{2} + h)C(|q^{34} - t^2|) \tag{15}$$

where $\bar{u}^P(h)$ has been expressed as a function of h. Combining these expressions, we get:

$$U^1(q, r^1) - \bar{u}^1(h) = \frac{1}{2}[V(r^1) - C(|q - t^1|)] + (\frac{1}{2} - h)C(|t^2 - t^1|) + hC(|q^{34} - t^1|) \tag{16}$$

$$U^2(q, r^2) - \bar{u}^2(h) = \frac{1}{2}[V(r^2) - C(|q - t^2|)] - (\frac{1}{2} - h)V(R) + hC(|q^{34} - t^2|) \tag{17}$$

Again evaluate (17) at the policy and rent allocation most favorable for candidate 2, namely $q = t^2$ and $r^2 = R - \bar{r}$, and then solve $U^2(t^2, R - \bar{r}) - \bar{u}^2(h) = 0$ for h. Denoting the solution by $\underline{h}$, we get:

$$\underline{h} = \frac{V(R) - V(R - \bar{r})}{2[V(R) + C(|q^{34} - t^2|)}$$

Again, for $h > \underline{h}$, candidate 2 prefers to merge for at least some feasible policy and rent allocation, while he cannot be induced to merge if $h < \underline{h}$. Repeating the same procedure for candidate 1, again it can be verified that 1 is always willing to merge with 2 even on the terms most favorable to 2. By symmetry, similar results hold for 3 and 4, given that 1 and 2 have merged.

Note that $1/2 > \bar{h} > \underline{h} > 0$, where the inequality $\bar{h} > \underline{h}$ follows from $q^{34} \geq t^3$ and $\bar{r} > 0$. Hence, combining these two results, we conclude that if $h < \underline{h}$ then the equilibrium is unique and consists of a four party system, where each moderate candidate wins with probability 1/2 on a policy platform that coincides with his bliss point. The reason is that, in stage 1 of the game when deciding on party formation, if $h < \underline{h}$ then it is a dominant strategy for the moderate candidate to say no to any merger proposal made by the extremists.

Conversely, if $h > \bar{h}$ then the equilibrium is unique and consists of a two party system where moderate and extremist candidates have merged on a policy platform and rent allocation that coincides with the Nash bargaining outcome (to be derived below), and both parties win with probability 1/2. The reason is that, in stage 1 of the game and if $h > \bar{h}$, it is a dominant strategy for both the moderate and the extremist to merge, irrespective of what the opponents do.

What happens if $\bar{h} > h > \underline{h}$? Note that both moderate candidates are better off in the four party equilibrium than in the two party equilibrium, since they have larger expected rents and (weakly) more favorable policies, and the probability of victory is 1/2 in both cases. But then, given that party formation occurs in sequence, the four party system is the unique equilibrium even in this range parameters. Specifically, the moderate candidate who speaks first will say no to any merger proposal received by the extremist, since he knows that, for $\bar{h} > h$, this will also induce the other moderate to reject any subsequent merger proposal by the other extremist. Hence here too the unique equilibrium is a four party system. Only in the knife-edge case $h = \bar{h}$, where the moderates are indifferent between a two-party and a four party system, there can be multiple equilibria, depending on the moderates' beliefs about what the moderate opponent will do.

Finally, we want to compare the Nash bargaining outcome under runoff elections with that under single round elections in the two party system. This can be achieved comparing (16) and (17) with (3). Specifically, holding fixed all equilibrium variables (q, r^P, q^{34}), define

$$G(h) \equiv \frac{U^1(q, r^1) - \bar{u}^1(h)}{U^2(q, r^2) - \bar{u}^2(h)} \tag{18}$$

where $U^P(q, r^P) - \bar{u}^P(h)$ are given by (16) and (17) respectively, corresponding to the expressions under runoff elections. The function $G(h)$ has the following properties. First, for $h = 1/2$, it reduces to the same expression under single round elections. This can be verified comparing (16) and (17) with (3). This in turn implies that, for $h = 1/2$, the Nash bargaining outcome is identical under the two electoral rules. Second, and holding (q, r^P, q^{34}), fixed, the function $G(h)$ is strictly decreasing in h. This can be verified from (16), (17) and the definition of $G(h)$.[17] This in turn implies that, for any $h < 1/2$, the moderate candidate 2 has more bargaining power under runoff elections than under single round elections, and hence the Nash equilibrium outcome characterized by (4) gives a policy closer to his bliss point and a rent allocation more favorable to him. Moreover, and by the same argument, the smaller is h, the closer is the policy to candidate 2's bliss point, and the larger is the share of rents that goes to this candidate. QED

[17] Specifically, after some algebra, the sign of $G_h(h)$ is the same as that of the following expression:

$$-[C(q^{34}) - C(t^2)]\left[V(R) - V(r^2) + C(t^2 - q)\right] - [V(r^1) + C(t^2) - C(q)][V(R) + C(q^{34} - t^2)]$$

It is easy to verify that the sign of all square brackets is positive, as $q^{34} > t^3 > t^2 > q$, and $V(R) > V(r^2)$.

Victory at the first round

Consider a three-party system consisting of say $\{1,2\}$, $\{3\}$, and $\{4\}$. Let both ε_1 and ε_2 be distributed with density $f(.)$ and cumulative distribution $F(.)$ over the interval $[-e/2, e/2]$. As stated in the text, $f(.)$ is symmetric around 0 and ε_1 and ε_2 are independently distributed. The probability that $\{1,2\}$ wins is: $\Pr(\varepsilon_1 > 0)+Pr(\varepsilon_1 \leq 0, \varepsilon_1+\varepsilon_2 > 0) = 1/2+ \int_{-\frac{e}{2}}^{0}[1-F(-\varepsilon_1)]f(\varepsilon_1)d\varepsilon_1$, where we have used the fact that $Pr(\varepsilon_1+\varepsilon_2 > 0) = 1-F(-\varepsilon_1)$. The handicap of running alone for candidate 3 is thus

$$h = \int_{-\frac{e}{2}}^{0}[1-F(-\varepsilon_1)]f(\varepsilon_1)d\varepsilon_1 = 1/2 - \int_{-\frac{e}{2}}^{0} F(-\varepsilon_1)f(\varepsilon_1)d\varepsilon_1.$$

Note that:

(i) $\int_{-\frac{e}{2}}^{e/2} F(-\varepsilon_1)f(\varepsilon_1)d\varepsilon_1 = \int_{-\frac{e}{2}}^{0} F(-\varepsilon_1)f(\varepsilon_1)d\varepsilon_1 + \int_{0}^{e/2} F(-\varepsilon_1)f(\varepsilon_1)d\varepsilon_1 = 1/2$, where the last equality follows from the assumption that ε_1 and ε_2 are independently and symmetrically distributed around 0.

(ii) $\int_{-\frac{e}{2}}^{0} F(-\varepsilon_1)f(\varepsilon_1)d\varepsilon_1 > \int_{0}^{e/2} F(-\varepsilon_1)f(\varepsilon_1)d\varepsilon_1 > 0$, since $F(.)$ is increasing and $f(.)$ is symmetric around zero.

Combining (i) and (ii), we have that $1/2 > \int_{-\frac{e}{2}}^{0} F(-\varepsilon_1)f(\varepsilon_1)d\varepsilon_1 > 1/4$, implying that $1/4 > h > 0$.

In the special case in which ε_1 and ε_2 are both uniformly distributed over $[-e/2, e/2]$ with density $1/e$, we have:

$$h = \int_{-\frac{e}{2}}^{0}[1-F(-\varepsilon_1)]f(\varepsilon_1)d\varepsilon_1 = \frac{1}{e}\int_{-\frac{e}{2}}^{0}(\frac{1}{2}+\frac{\varepsilon_1}{e})d\varepsilon_1 = 1/8$$

since $\frac{1}{e}\int_{-\frac{e}{2}}^{0}(\frac{1}{2}+\frac{\varepsilon_1}{e})d\varepsilon_1 = \frac{1}{4}+\frac{1}{2e^2}((0)^2-(\frac{e^2}{4})) = \frac{1}{4}-\frac{1}{8} = \frac{1}{8}$.

Equilibrium with endorsements

In this section we discuss what happens when extremists are allowed to endorse the moderates after the first round of voting (if the latter accept). Recall the assumption that $\eta = \varepsilon_1 + \varepsilon_2$, where ε_1 and ε_2 are independently and identically distributed, with a uniform distribution over the interval $[-e/2, e/2]$. Exploiting the properties of uniform distributions, we obtain that η is distributed over the interval $[-e, e]$, it has zero mean, and a symmetric cumulative distribution given by

$$G(z) = \frac{1}{2} + \frac{z}{e} - \frac{z^2}{2e^2} \text{ for } e \geq z \geq 0$$
$$G(z) = \frac{1}{2} + \frac{z}{e} + \frac{z^2}{2e^2} \text{ for } -e \leq z \leq 0$$

We start with the case in which rents are not contractable at the endorsement stage, and in case of victory the endorsing extremist gets rents $\bar{r}$ while the endorsed moderates retains rents $R - \bar{r}$. Suppose that both moderate candidates have passed the first round and that no coalition has formed before the first round. Define

$$\breve{\varepsilon} \equiv \frac{\delta\underline{\alpha}\left[V(R) + C(|t^2 - t^1|)\right]}{2\left[V(R) - V(R - \bar{r})\right]} - \frac{e}{2} \gtrless 0$$

We have:

Lemma 2 *Irrespective of what candidate 3 does, candidate 2 prefers to be endorsed by candidate 1 if* $\varepsilon_1 < \breve{\varepsilon} - \frac{\delta\underline{\alpha}}{2}$*, and he prefers no endorsement if* $\varepsilon_1 > \breve{\varepsilon}$*. In between, if* $\breve{\varepsilon} - \frac{\delta\underline{\alpha}}{2} \leq \varepsilon_1 \leq \breve{\varepsilon}$*, then 2 prefers to seek the endorsement of the extremist if 3 has also been endorsed, while 2 prefers no endorsement if 3 has not been endorsed. Candidate 3 behaves symmetrically (in the opposite direction), depending on whether* $-\varepsilon_1$ *is below or above these same thresholds.*

Proof

Suppose that both 2 and 3 have been endorsed by their extremist neighbors. By our previous assumptions, candidate 2 wins if $\varepsilon_1 + \varepsilon_2 > 0$. When decisions over endorsements are made, the realization of ε_1 is known, but ε_2 is not. Hence the probability that candidate 2 wins is

$$\Pr(\varepsilon_2 > -\varepsilon_1) = \frac{1}{2} + \frac{\varepsilon_1}{e} \tag{19}$$

where the RHS follows from the assumptions on the distribution of the two electoral shocks. Candidate 2's expected utility is:

$$(\frac{1}{2} + \frac{\varepsilon_1}{e})V(R - \bar{r}) - (\frac{1}{2} - \frac{\varepsilon_1}{e})C(|t^3 - t^2|) \tag{20}$$

Suppose instead that 2 refuses the endorsement of 1, while 3 is endorsed by 4. Now 2 loses the support of $\delta\underline{\alpha}$ voters, the attached extremists in group 1, while 3 carries all voters in group 4. Hence, repeating the analysis in (10), the probability that 2 wins is:

$$\Pr(\varepsilon_2 > \frac{\delta\underline{\alpha}}{2} - \varepsilon_1) = \frac{1}{2} + \frac{\varepsilon_1}{e} - \frac{\delta\underline{\alpha}}{2e} \tag{21}$$

if $\varepsilon_1 \geq \frac{\delta\underline{\alpha}}{2} - \frac{e}{2}$, and it is 0 if $\varepsilon_1 < \frac{\delta\underline{\alpha}}{2} - \frac{e}{2}$. Candidate 2's expected utility is then:

$$(\frac{1}{2} + \frac{\varepsilon_1}{e} - \frac{\delta\underline{\alpha}}{2e})V(R) - (\frac{1}{2} - \frac{\varepsilon_1}{e} + \frac{\delta\underline{\alpha}}{2e})C(|t^3 - t^2|)$$

provided that the first expression in brackets is strictly positive and the second expression in brackets is strictly less than 1, which again occurs if $\varepsilon_1 \geq \frac{\delta\underline{\alpha}}{2} - \frac{e}{2}$. If instead $\varepsilon_1 < -\frac{e}{2} + \frac{\delta\underline{\alpha}}{2}$, then the probability that 2 wins is 0 and his expected utility is $-C(|t^3 - t^2|)$.[18]

Equalizing the two expected utilities, candidate 2 is indifferent between these two alternatives if:

$$\varepsilon_1 = \check{\varepsilon} \equiv \frac{\delta\underline{\alpha}\,[V(R) + C(|t^3 - t^2|)]}{2\,[V(R) - V(R - \bar{r})]} - \frac{e}{2} \tag{22}$$

Note that $\frac{e}{2} > \check{\varepsilon} > -\frac{e}{2}$, where the first inequality follows from (A2) and the second by inspection of the equation above. If $\varepsilon_1 > \check{\varepsilon}$ then candidate 2 strictly prefers no endorsement, given that 3 has been endorsed. While if $\varepsilon_1 < \check{\varepsilon}$ then candidate 2 strictly prefers to be endorsed, given that 3 has been endorsed.

Next, suppose that no moderate candidate has been endorsed by the extremist. By symmetry, the probability that 2 wins if he is not endorsed is still described by (19), but, as 2 does not have to share rents with 1 if elected, his expected utility is now

$$(\frac{1}{2} + \frac{\varepsilon_1}{e})V(R) - (\frac{1}{2} - \frac{\varepsilon_1}{e})C(|t^3 - t^2|) \tag{23}$$

If instead candidate 2 accepts to be endorsed and 3 refuses, the probability that 2 wins is:

$$\Pr(\varepsilon_2 > -\frac{\delta\underline{\alpha}}{2} - \varepsilon_1) = \frac{1}{2} + \frac{\varepsilon_1}{e} + \frac{\delta\underline{\alpha}}{2e} \tag{24}$$

if $\varepsilon_1 \leq \frac{e}{2} - \frac{\delta\underline{\alpha}}{2}$ and it is 1 if $\varepsilon_1 > \frac{e}{2} - \frac{\delta\underline{\alpha}}{2}$.[19] In this case, candidate 2's expected utility is:

$$(\frac{1}{2} + \frac{\varepsilon_1}{e} + \frac{\delta\underline{\alpha}}{2e})V(R - \bar{r}) - (\frac{1}{2} - \frac{\varepsilon_1}{e} - \frac{\delta\underline{\alpha}}{2e})C(|t^3 - t^2|)$$

provided that the first expression in brackets is strictly less than 1 and the second expression in brackets is strictly positive, which occurs if $\varepsilon_1 \leq \frac{e}{2} - \frac{\delta\underline{\alpha}}{2}$. If instead $\varepsilon_1 > \frac{e}{2} - \frac{\delta\underline{\alpha}}{2}$, then the

[18] By (A2), the first expression in brackets is always strictly less than 1 and the second expression in brackets is always positive.

[19] By (A2), $\Pr(\varepsilon_2 > \frac{\delta\underline{\alpha}}{2} - \varepsilon_1) < 1$ and $\Pr(\varepsilon_2 > -\frac{\delta\underline{\alpha}}{2} - \varepsilon_1) > 0$ for any $\varepsilon_1 \in [-e/2, e/2]$.

probability that 2 wins is 1 and his expected utility cannot exceed $V(R-\bar{r})$.[20]

Candidate 2 is then indifferent between these two options if $\varepsilon_1 = \breve{\varepsilon} - \frac{\delta\alpha}{2}$. If $\varepsilon_1 > \breve{\varepsilon} - \frac{\delta\alpha}{2}$ then candidate 2 strictly prefers no endorsement, given that 3 has not been endorsed. While if $\varepsilon_1 < \breve{\varepsilon} - \frac{\delta\alpha}{2}$ then candidate 2 strictly prefers to be endorsed, given that 3 has not been endorsed.

By symmetry, 3 has similar preferences, but in the opposite direction and with respect to the symmetric thresholds $-\breve{\varepsilon} + \frac{\delta\alpha}{2}$ and $-\breve{\varepsilon}$ (eg. 3 prefers no endorsement, given that 2 has not been endorsed, if $\varepsilon_1 < -\breve{\varepsilon} + \frac{\delta\alpha}{2}$, and 3 prefers no endorsement, given that 2 has been endorsed, if $\varepsilon_1 < -\breve{\varepsilon}$). QED

Invoking Lemma 2, we now describe the equilibrium continuation if the two moderate candidates have passed the first round and compete over the second round. Equilibrium endorsements depend on whether the thresholds in Lemma 2 are positive or negative. These thresholds are positive for high values of δ (the fraction of attached voters) and low values of $\bar{r}$ (the minimal rents that have to be left to the extremists). This in turn increases the willingness of the moderates to accept endorsements. This provides the intuition for the proposition to follow. Specifically, under (A1-A2), we have:

Proposition 6 *(i) Suppose $\breve{\varepsilon} - \frac{\delta\alpha}{2} > 0$. Then, the equilibrium is unique and at least one moderate candidate accepts the endorsement of the ideologically closer extremist. Specifically, if $\varepsilon_1 > \breve{\varepsilon}$, 3 accepts the endorsement while 2 does not. Symmetrically, if $\varepsilon_1 < -\breve{\varepsilon}$, 2 accepts the endorsement while 3 does not. For all other realizations of ε_1, both 2 and 3 accept the endorsements.*

(ii) Suppose that $\breve{\varepsilon} < 0$. Then, the equilibrium is unique and at most one of the two moderate candidates accepts the endorsement of his extremist neighbor. Specifically, if $\varepsilon_1 < \breve{\varepsilon} - \frac{\delta\alpha}{2}$, 2 accepts the endorsement while 3 does not. Symmetrically, if $\varepsilon_1 > -\breve{\varepsilon} + \frac{\delta\alpha}{2}$, 3 accepts the endorsement while 2 does not. For all other realizations of ε_1, neither 2 nor 3 accept the endorsements.

(iii)Suppose $\breve{\varepsilon} \geq 0 \geq \breve{\varepsilon} - \frac{\delta\alpha}{2}$. Here, there are two cases to consider. If $\frac{\delta\alpha}{2} \geq 2\breve{\varepsilon}$, then the equilibrium is identical to the one described under point (ii). If $\frac{\delta\alpha}{2} < 2\breve{\varepsilon}$, then the equilibrium is unique and depending on the realization of ε_1, both moderates are endorsed by the extremists, none are, or one moderate only is endorsed by the closer extremist.

Proof of Proposition 6

Suppose first that $\breve{\varepsilon} - \frac{\delta\alpha}{2} > 0$. This then implies $\breve{\varepsilon} > 0 > -\breve{\varepsilon} + \frac{\delta\alpha}{2}$.The equilibrium is illustrated in Figure A1. If $\varepsilon_1 > \breve{\varepsilon}$ by Lemma 2, 2 does not accept the endorsement of 1

[20] Assumption (A2) implies that the first expression in brackets is always positive and the second one is always less than 1.

whatever 3 does; and as $\varepsilon_1 > \breve{\varepsilon}$ implies $\varepsilon_1 > 0 > -\breve{\varepsilon} + \frac{\delta\alpha}{2}$, 3 accepts the endorsement of 4 even if 2 is not endorsed. By symmetry, if $\varepsilon_1 < -\breve{\varepsilon}$, 3 does not accept to be endorsed, while 2 is endorsed. If $\varepsilon_1 \in [-\breve{\varepsilon} + \frac{\delta\alpha}{2}, \breve{\varepsilon} - \frac{\delta\alpha}{2}]$, then both moderates find it optimal to seek the endorsement of the extremists, no matter what their opponent does. If $\varepsilon_1 \in (\breve{\varepsilon} - \frac{\delta\alpha}{2}, \breve{\varepsilon}]$, then candidate 3 still finds it optimal to seek the endorsement of 4 no matter what 2 does; and given 3's behavior, 2 also finds it optimal to seek the endorsement of 1. The same conclusion holds, but with the roles of 2 and 3 reversed, if $\varepsilon_1 \in [-\breve{\varepsilon} + \frac{\delta\alpha}{2}, -\breve{\varepsilon})$.

Next suppose that $\breve{\varepsilon} < 0$. This then implies $\breve{\varepsilon} - \frac{\delta\alpha}{2} < \breve{\varepsilon} < 0$ and $-\breve{\varepsilon} + \frac{\delta\alpha}{2} > -\breve{\varepsilon} > 0$. This equilibrium is illustrated in Figure A2. If $\varepsilon_1 \in [\breve{\varepsilon}, -\breve{\varepsilon}]$, then both moderates find it optimal to seek no endorsement, no matter what their opponent does. If $\varepsilon_1 \in [\breve{\varepsilon} - \frac{\delta\alpha}{2}, \breve{\varepsilon})$, 3 does not seek for an endorsement as $\varepsilon_1 < \breve{\varepsilon} < -\breve{\varepsilon}$, and given 3's behavior then candidate 2 also seeks no endorsement. If $\varepsilon_1 < \breve{\varepsilon} - \frac{\delta\alpha}{2}$ 2 seeks an endorsement no matter what 3 does, and 3 does not seek an endorsement for the same reason spelled above. The same conclusion holds, but with the roles of 2 and 3 reversed, if $\varepsilon_1 \in (-\breve{\varepsilon}, -\breve{\varepsilon} + \frac{\delta\alpha}{2}]$. Finally, if $\varepsilon_1 > -\breve{\varepsilon} + \frac{\delta\alpha}{2}$ then candidate 2 still finds it optimal to seek no endorsement no matter what 3 does, while 3 finds it optimal to seek the endorsement of 4 no matter what 2 does.

Finally, suppose that $\breve{\varepsilon} > 0 > \breve{\varepsilon} - \frac{\delta\alpha}{2}$. Suppose also that $\frac{\delta\alpha}{2} \geq 2\breve{\varepsilon}$, so that $-\breve{\varepsilon} + \frac{\delta\alpha}{2} \geq \breve{\varepsilon}$ and $\breve{\varepsilon} - \frac{\delta\alpha}{2} \leq -\breve{\varepsilon}$. This equilibrium is illustrated in Figure A3. As shown in the Figure, if $\varepsilon_1 \in [-\breve{\varepsilon}, \breve{\varepsilon}]$ each moderate candidate would accept to be endorsed only if the other moderate is also endorsed. However, by the assumed sequentiality of the endorsement proposals, the first moderate receiving an offer of endorsement by the closer extremist would always be better off by refusing this offer, knowing that this will induce the other moderate to refuse the offer by the other extremist as well. Hence, for $\varepsilon_1 \in [-\breve{\varepsilon}, \breve{\varepsilon}]$ no endorsement occurs. No endorsement also occur if $\varepsilon_1 \in [\breve{\varepsilon} - \frac{\delta\alpha}{2}, -\breve{\varepsilon})$ (or symmetrically, if $\varepsilon_1 \in [\breve{\varepsilon}, -\breve{\varepsilon} + \frac{\delta\alpha}{2}))$ as at least one moderate always prefers to run alone and the other accepts to be endorsed only if the other moderate is endorsed. Hence, in this case at most one moderate is endorsed, 2 if $\varepsilon_1 < \breve{\varepsilon} - \frac{\delta\alpha}{2}$ and 3 if $\varepsilon_1 > -\breve{\varepsilon} + \frac{\delta\alpha}{2}$. Suppose next that $\frac{\delta\alpha}{2} < 2\breve{\varepsilon}$ so that $-\breve{\varepsilon} + \frac{\delta\alpha}{2} < \breve{\varepsilon}$ and $\breve{\varepsilon} - \frac{\delta\alpha}{2} > -\breve{\varepsilon}$. This equilibrium is illustrated in Figure A4. Here, if $\varepsilon_1 \in [\breve{\varepsilon} - \frac{\delta\alpha}{2}, -\breve{\varepsilon} + \frac{\delta\alpha}{2}]$ for the previous argument, no candidate accepts to be endorsed. For $\varepsilon_1 \in (-\breve{\varepsilon} + \frac{\delta\alpha}{2}, \breve{\varepsilon}]$ 3 always accepts to be endorsed and 2 too accepts to be endorsed if he expect 3 to be endorsed. Hence, both candidates are endorsed. Symmetrically, for $\varepsilon_1 \in [-\breve{\varepsilon}, \breve{\varepsilon} - \frac{\delta\alpha}{2})$, 2 always accepts to be endorsed and 3 too accepts to be endorsed if he expect 2 to be endorsed. Hence, both candidates are endorsed. Finally, for $\varepsilon_1 > \breve{\varepsilon}$ or $\varepsilon_1 < -\breve{\varepsilon}$ only one candidate is endorsed, 2 in the former case and 3 in the latter. QED

A centrist party

Single round elections. Consider first the party formation stage. As stated in the text, if $\underline{\alpha} \geq \alpha^c$ and $\lambda > 1/4$, it is immediate to show that the extremist party always merges with the moderate. Whether the centrist party is also included or not in one of the coalitions, depends on its size. If α^c is small, then including it is not worth the cost of rents and policy accommodation that this would require in the subsequent bargaining stage. In this case, the equilibrium under single round elections is identical to that described in Section 3. If α^c is sufficiently large, then the increase in the probability of victory compensates for the cost of including it. In this case, whether c merges with $\{1,2\}$ or $\{3,4\}$ depends on the order of moves at the party formation stage.

Suppose that α^c is large enough and party $\{1,2,c\}$ has formed. How are policy and rents determined by this party? The first step is to compute the disagreement points of all candidates. Since disagreement implies unilateral party breakup, we have:

$$\bar{u}^P = -C(q^{34} - t^P), \qquad P = 1, 2, c$$

where as before q^{34} denotes the equilibrium policy set by party $\{3,4\}$. We thus have:

$$\begin{aligned} U^P(q, r^P) &= p[V(r^P) - C(|q - t^P|) - (1-p)C(q^{34} - t^P)], \qquad P = 1, 2, c \\ U^P(q, r^P) - \bar{u}^P &= p[V(r^P) - C(|q - t^P|) + C(q^{34} - t^P)], \qquad P = 1, 2, c \end{aligned} \tag{25}$$

where $p = \Pr[\eta \geq -\frac{\alpha^c}{2}]$ denotes the probability that party $\{1,2,c\}$ wins. The Nash bargaining outcome is the solution to the problem of maximizing $\{[U^1(q,r^1) - \bar{u}^1][U^2(q,r^2) - \bar{u}^2]$ $[U^c(q,r^c) - \bar{u}^c]\}$ by choice of q, r^1, r^2, r^c. After some transformations, at an interior optimum the first order conditions of this problem imply:

$$\begin{aligned} \sum\nolimits_{1,2,c} \frac{U_q^P(q,r^P)}{U_r^P(q,r^P)} &= 0 \\ \frac{U_r^2(q,r^2)}{U_r^c(q,r^c)} &= \frac{U^2(q,r^2) - \bar{u}^2}{U^c(q,r^c) - \bar{u}^c} \\ \frac{U_r^2(q,r^2)}{U_r^1(q,r^1)} &= \frac{U^2(q,r^2) - \bar{u}^2}{U^1(q,r^1) - \bar{u}^1} \end{aligned} \tag{26}$$

Manipulation of these conditions can be shown to imply $q > t^2$, confirming the intuition stated in the text that the centrist party has more bargaining power than the extremist, despite its possibly smaller size.

Runoff elections. Repeating the same logic as in Section 3, it is easy to show that moderate candidates never want to merge with the extremists, since they can capture the extremists vote at the second round. What about a merger with the centrist? If say candidate 2 merges with the centrist, its expected utility cannot exceed $pV(R-\bar{r})-(1-p)C(|t^3-t^2|)$ (recall that in the equilibrium under runoff elections, candidate 3 does not merge with 4), where as above $p=\Pr[\eta \geq -\frac{\alpha^c}{2}]$. If instead no such merger takes place and both 2 and 3 run alone, then candidate 2's expected utility is $\frac{1}{2}[V(R)-C(|t^3-t^2|)]$. Combining these two expressions, we obtain that the moderates prefers to merge with the centrist candidate rather than to run alone, given that the other moderate is running alone, if

$$\Pr(\eta > -\frac{\alpha^c}{2}) > \frac{V(R)+C(t^3-t^2)}{2[V(R-\bar{r})+C(t^3-t^2)]} \tag{27}$$

If p exceeds the threshold on the RHS of (27), then there is a feasible combination of q and r^2 that leaves candidates 2 and c better off with a merger than without it - of course candidate c has nothing to loose from such merger.[21]

If condition (27) is satisfied and party $\{2,c\}$ is formed, then policy and rents inside this party are set according to the Nash bargaining outcome. Repeating the procedure in the previous proofs, the Nash bargaining solution implies:

$$\begin{aligned} \sum\nolimits_{2,c} \frac{U_q^P(q,r^P)}{U_r^P(q,r^P)} &= 0 \\ \frac{U_r^2(q,r^2)}{U_r^c(q,r^c)} &= \frac{U^2(q,r^2)-\bar{u}^2}{U^c(q,r^c)-\bar{u}^c} \end{aligned} \tag{28}$$

where now $U^P(q,r^P)-\bar{u}^P = p[V(r^P)-C(|q-t^P|)+C(t^3-t^P)]$, which is the same expression as in (25), except that q^{34} has been replaced by t^3. Manipulation of these first order conditions can be shown to imply that $q > \frac{\lambda}{2}$, which is the mid point between t^2 and $1/2$.

Comparing single round vs runoff elections. Clearly single round elections have a smaller equilibrium number of parties than runoff elections, since in the latter the extremists are always on their own. What about policy moderation?

Obviously party $\{3\}$ under runoff has a more moderate policy than party $\{3,4\}$ under single round elections. The comparison between party $\{2,c\}$ under runoff and party $\{1,2,c\}$ under single round is more subtle, however. On the one hand, the extremist candidate is

[21] Note that the LHS of (27) increases in α^c, and the RHS equals 1/2 if $\bar{r}=0$ while it rises above 1/2 as $\bar{r}$ increases above 0. Hence (27) is certainly consistent with $\alpha^c \leq \underline{\alpha}$ for sufficiently small $\bar{r}$ or large $\underline{\alpha}$.

only included under single round elections, and this pushes party $\{1, 2, c\}$ towards a more extreme policy than party $\{2, c\}$. This can be seen formally by noting that the summation of the marginal rates of substitutions between rents and policies includes the extremist candidate in (26) but not in (28). On the other hand, the bargaining power of the centrist candidate is stronger under single round elections than under runoff. The reason is that the opponent runs on a more extreme policy under single round (q^{34}) than under runoff (t^3), and the threat of electoral defeat is less fearsome for the centrist candidate than for the other party members, the more so the more extreme is the policy platform of the opponent. This can be seen by comparing the remaining expressions in (26) vs (28). Hence, a priori and without additional restrictions on functional form we cannot rule out the possibility that, despite the inclusion of the extremist candidate, party $\{1, 2, c\}$ under single round elections enacts a more moderate policy than party $\{2, c\}$ under runoff.

Relaxing the restrictions on party formation

Here we allow the formation of parties consisting of up to *three* adjacent candidates, and show that Propositions 1-3 still identically hold provided that candidates care sufficiently about policy relative to rents and that polarization (i.e., λ) is sufficiently high. Though we assume that there are no attached voters, although with suitable changes to the proofs and conditions the results with attached voters would also go through. Below we discuss the possible formation of a three candidate party consisting of $\{1, 2, 3\}$; given symmetry, the proposition below holds identically for a party consisting of $\{2, 3, 4\}$.

Before going through a formal proof, 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. Specifically, the policy set by $\{1, 2, 3\}$ would have to satisfy $q \geq 2\lambda$, where $q = 2\lambda$ is such that $t^4 - t^3 = t^3 - q$, so that group 3 voters are indifferent between q and t^4. If this constraint is satisfied, then party $\{1, 2, 3\}$ wins the election with certainty, otherwise it wins with probability 1/2. But it only makes sense to form party $\{1, 2, 3\}$ if it wins with certainty, because otherwise the extremist and at least one moderate candidate would be strictly better off with the symmetric two party system $\{1, 2\}$ and $\{3, 4\}$. Of course, the constraint $q \geq 2\lambda$ 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.

When discussing the possible formation of party $\{1, 2, 3\}$, we need to be explicit about what is the disagreement point under which Nash bargaining is conducted inside this party. It is natural to assume that disagreement implies that the party breaks up and no further renegotiation about party formation is possible.

Consider first single round elections. We start with the following (r^{*P} and q^* denote the equilibrium outcomes described in Proposition 1).

Lemma 3 *If the following condition is satisfied*

$$Max\left\{[C(2\lambda-t^2)+\frac{1}{2}V(R)-\frac{1}{2}C(t^3-t^2)],[C(2\lambda)+\frac{1}{2}V(r^{*1})-\frac{1}{2}C(q^*)-\frac{1}{2}C(1-q^*)]\right\}$$
$$> \; V(R-2\bar{r}) \tag{29}$$

then under single round elections there is no feasible outcome under party $\{1,2,3\}$ *that leaves both candidates 1 and 2 better off than in the equilibrium outcome of some other feasible party system.*

Proof

Consider candidate 1. His most favorable outcome under party $\{1,2,3\}$ is that he gets all the feasible rents, $r^1 = R-2\bar{r}$, and the policy is as low as possible subject to the constraint of winning with certainty, namely, $q = 2\lambda$. In this case the utility of candidate 1 is $V(R-2\bar{r}) - C(2\lambda)$. His best alternative to party $\{1,2,3\}$ is a symmetric two party system. By Proposition 1, in the equilibrium outcome of a two party system, candidate 1 gets rents r^{*1} (if $\{1,2\}$ win the election) and the policy is q^* if $\{1,2\}$ win and $1-q^*$ otherwise. Hence in the symmetric two party equilibrium the expected utility of candidate 1 is: $\frac{1}{2}V(r^{*1})-\frac{1}{2}C(q^*)-\frac{1}{2}C(1-q^*)$, and candidate 1 prefers this symmetric equilibrium outcome to any feasible outcome under party $\{1,2,3\}$ if:

$$C(2\lambda)+\frac{1}{2}V(r^{*1})-\frac{1}{2}C(q^*)-\frac{1}{2}C(1-q^*) > V(R-2\bar{r}) \tag{30}$$

Next consider candidate 2. His most favorable outcome under party $\{1,2,3\}$ is that he gets all the feasible rents, $r^2 = R-2\bar{r}$, and the policy is again as low as possible, namely $q = 2\lambda$. In this case the utility of candidate 2 is $V(R-2\bar{r})-C(2\lambda-t^2)$. From his perspective, the best alternative to party $\{1,2,3\}$ is a four party system in which all candidates run alone. Candidate 2 expected utility in this case is: $\frac{1}{2}V(R)-\frac{1}{2}C(t^3-t^2)$. Hence, candidate 2 prefers the four party system to any feasible outcome under party $\{1,2,3\}$ if:

$$C(2\lambda-t^2)+\frac{1}{2}V(R)-\frac{1}{2}C(t^3-t^2) > V(R-2\bar{r}) \tag{31}$$

Combining these two inequalities we get (A3). QED

We are now ready to state our first result.

Proposition 7 *Consider single round elections. If (A3) is satisfied, a party resulting from the merger of three candidates cannot formed in equilibrium, and Propositions 1 and 2 hold.*

Proof

Start with single round elections and suppose that $\lambda > 1/4$. Let e and m (or e', m') denote an extremist and moderate candidate respectively, and index by s the substages of the party formation stage. Since there are four candidates, and at most each one of them has a proposal right, $s = 1, 2, 3, 4$ (or less if a partition is reached before everyone has made a proposal). We first prove the following:

Lemma 4 *If party* $\{e, m, m'\}$ *is formed, this can only happen for* $s = 1, 2$.

Proof of Lemma 4

Consider the last substage, $s = 4$. If it is reached with a four party system, then the proposer will propose party $\{e, m\}$ (or $\{e', m'\}$) and this proposal will be accepted. This being the last substage of the game, such a party will win the election for sure, leaving both e and m better off than in the four party system.

Now consider $s = 3$, and suppose again that it is reached with a four party system. Anticipating the outcome $s = 4$, a party consisting of $\{e, m\}$ (or $\{e', m'\}$) will again be formed at $s = 3$. The reason is that leaving a four party system to whoever will make a proposal at $s = 4$ is suboptimal (strictly or weakly depending on the identity of the proposers). And proposing party $\{e, m, m'\}$ is also suboptimal, because either m or m' will reject this proposal, anticipating that in $s = 4$ they will be able to merge with the nearby extremist, and thus win the election for sure. This completes the proof of Lemma 3.

Now consider $s = 1$ or 2. We now prove the following:

Lemma 5 *If (A3) holds, the extremist candidates (say e) will never propose party* $\{e, m, m'\}$ *in* $s = 1, 2$, *and will always say no to any proposal to form such a party.*

Proof of Lemma 5

Condition A3 says that either: (i) candidate e prefers the symmetric two party equilibrium to the outcome under party $\{e, m, m'\}$; or (ii) candidate m or m' prefers the four party system to the outcome under party $\{e, m, m'\}$. Consider case (i). By Lemma 4 and the proof therein, if party $\{e, m, m'\}$ is not formed in $s = 1, 2$, it will also not be formed in later substages. By the reasoning in Proposition 1, we will then have a symmetric two party system, which by case (i) of (A3) is better than the outcome under $\{e, m, m'\}$ for extremist candidates. Hence the extremist will never allow party $\{e, m, m'\}$ to be formed. Next, consider case (ii). Here, if party $\{e, m, m'\}$ was formed, it will not survive once the bargaining stage is reached, since by case (ii) of (A3) candidate m is strictly better off by breaking the party and moving to the four party system (recall the restriction that a party must consist of adjacent candidates, and if a party breaks up then no renegotiation can take place amongst candidates). But this disagreement outcome is worse than the two party equilibrium from the perspective of extremist candidates, who will thus veto the formation of party $\{e, m, m'\}$ in case (ii) as well. This completes the proof of Lemma 5.

If the party $\{e, m, m'\}$ is not formed in equilibrium and $\lambda > 1/4$, then Proposition 1 holds. Finally suppose that $\lambda \leq 1/4$. Then the centrist party $\{m, m'\}$ is viable and dominates any other party from the perspective of both moderates. Hence, Proposition 2 always holds. QED

Next, turn to runoff elections. Here the two party system cannot be reached in equilibrium, so we need the following result instead:

Lemma 6 *If the following condition is satisfied*

$$Max\left\{[C(2\lambda - t^2) + \frac{1}{2}V(R) - \frac{1}{2}C(t^3 - t^2)], [C(2\lambda) - \frac{1}{2}C(t^2) - \frac{1}{2}C(t^3)]\right\} > V(R - 2\bar{r}) \tag{A4}$$

then there is no feasible outcome under party $\{1, 2, 3\}$ *that leaves both candidates 1 and 2 better off than under a four party system.*

Proof of Lemma 6

Consider candidate 1. Under the four party system, candidate 1 gets an expected utility of $-\frac{1}{2}C(t^2) - \frac{1}{2}C(t^3)$. Hence candidate 1 prefers the four party system to any feasible outcome under party $\{1, 2, 3\}$ if:

$$C(2\lambda) - \frac{1}{2}C(t^2) - \frac{1}{2}C(t^3) > V(R - 2\bar{r}) \tag{32}$$

Combining (32) and(31), we get (A4). QED

We can then prove:

Proposition 8 *Consider runoff elections. If (A4) is satisfied, a party resulting from the merger of three candidates cannot be formed in equilibrium, and Proposition 3 holds.*

Here the proof is simpler, since if party $\{e, m, m'\}$ is not formed, then by Proposition 3 we end up with a four party system. But by condition (A4), either the moderates or the extremists prefer the four party system to the most favorable outcome under party $\{e, m, m'\}$. Hence under (A4) there is always a candidate who will veto the formation of $\{e, m, m'\}$, and thus Proposition 3 holds. QED

Note that (30) is less restrictive than (32), since candidate 1 prefers the symmetric two party equilibrium to the outcome under four parties, so that condition (A3) is less restrictive than (A4). Intuitively, runoff elections reduce the bargaining power of extremist candidates, and so extremists are more likely to favor party $\{1, 2, 3\}$ than under single round elections. That is, to rule out the formation of party $\{1, 2, 3\}$ we need to impose a more restrictive condition.

Finally, to better assess the implications of conditions (A3-A4), suppose that the function $C(x)$ takes the form $C(x) = \sigma x^2$. Then after some algebra conditions (A3) and (A4) can be rewritten respectively as:

$$Max\left\{[\frac{1}{2\sigma}V(R) + 7\lambda^2 - 3\lambda + \frac{1}{4}], [\frac{1}{2\sigma}V(r^{*1}) + 4\lambda^2 + q^*(1 - q^*) - \frac{1}{2}]\right\} > V(R - 2\bar{r})/\sigma \quad (33)$$

$$Max\left\{[\frac{1}{2\sigma}V(R) + 7\lambda^2 - 3\lambda + \frac{1}{4}], [3\lambda^2 - \frac{1}{4}]\right\} > V(R - 2\bar{r})/\sigma \quad (34)$$

Both conditions are more likely to be satisfied for values of λ above $1/4$ (i.e., a more polarized political system), and for high values of σ (i.e., if the value of policy relative to rents is high).

Figure A1

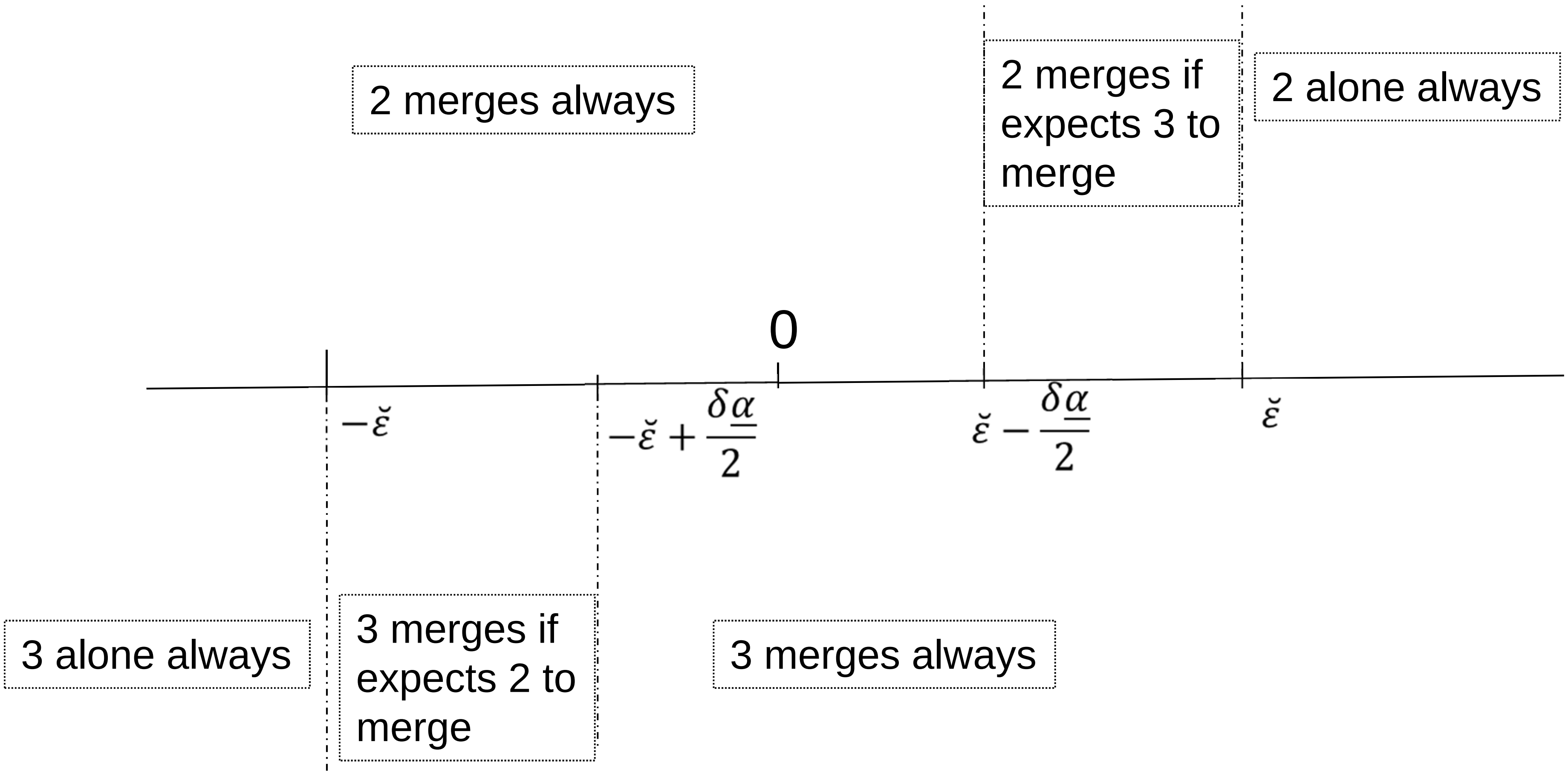
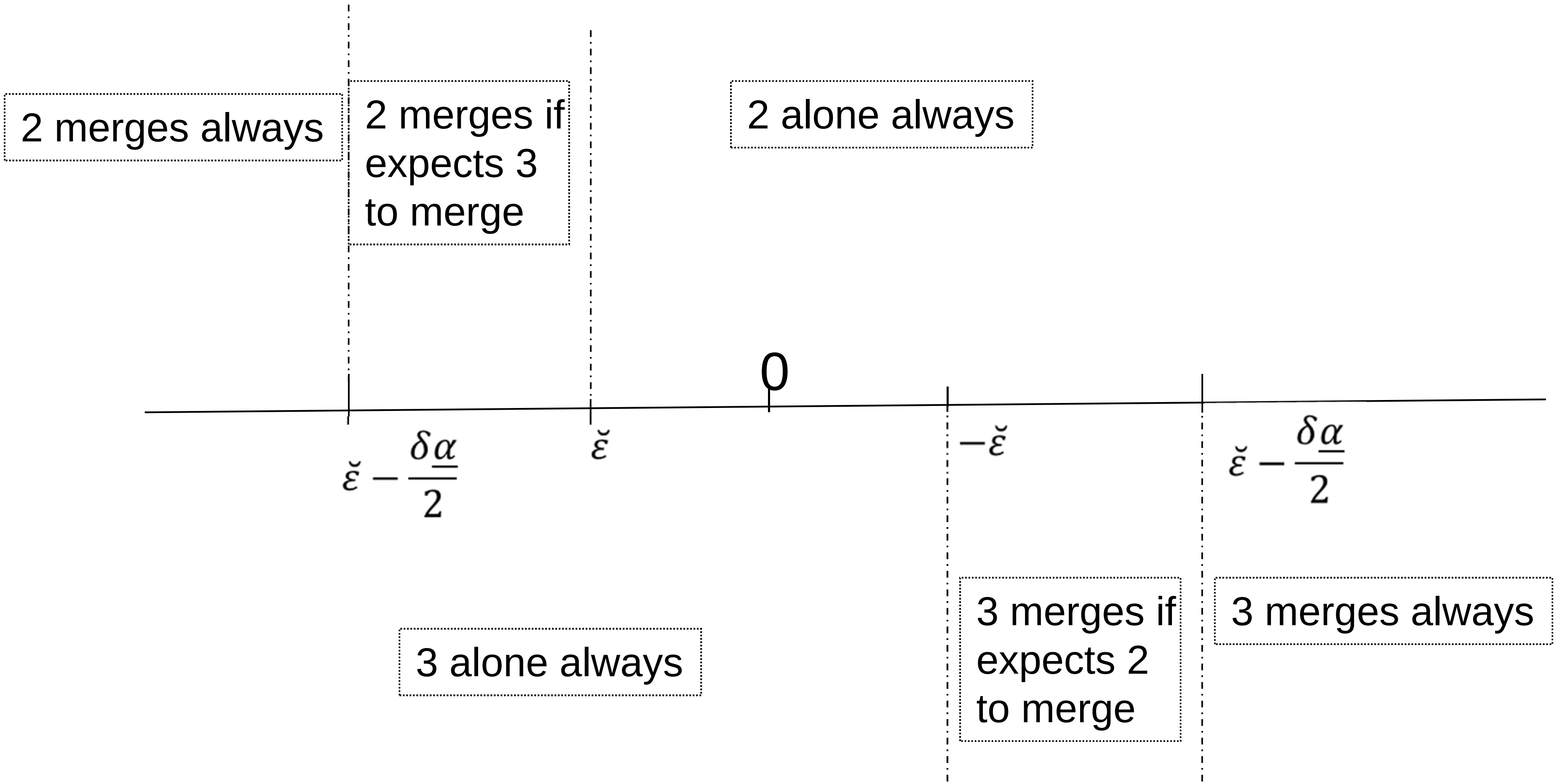

2 merges always
2 merges if expects 3 to merge
2 alone always
0
$-\breve{\varepsilon}$
$-\breve{\varepsilon}+\frac{\delta\underline{\alpha}}{2}$
$\breve{\varepsilon}-\frac{\delta\underline{\alpha}}{2}$
$\breve{\varepsilon}$
3 alone always
3 merges if expects 2 to merge
3 merges always

Figure A2



2 merges always
2 merges if expects 3 to merge
2 alone always
0
$\breve{\varepsilon} - \frac{\delta \underline{\alpha}}{2}$
$\breve{\varepsilon}$
$-\breve{\varepsilon}$
$\breve{\varepsilon} - \frac{\delta \underline{\alpha}}{2}$
3 alone always
3 merges if expects 2 to merge
3 merges always

Figure A3

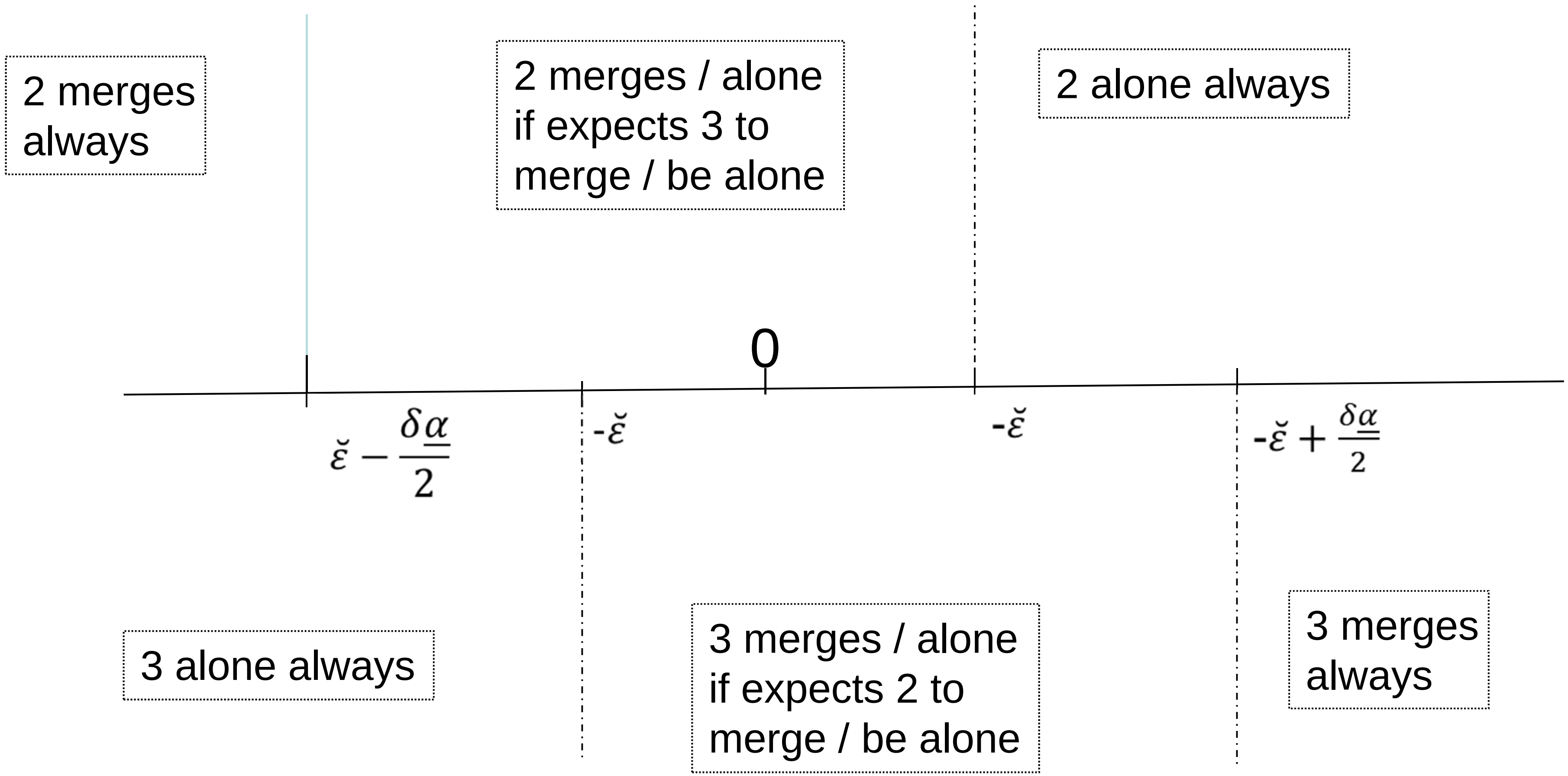
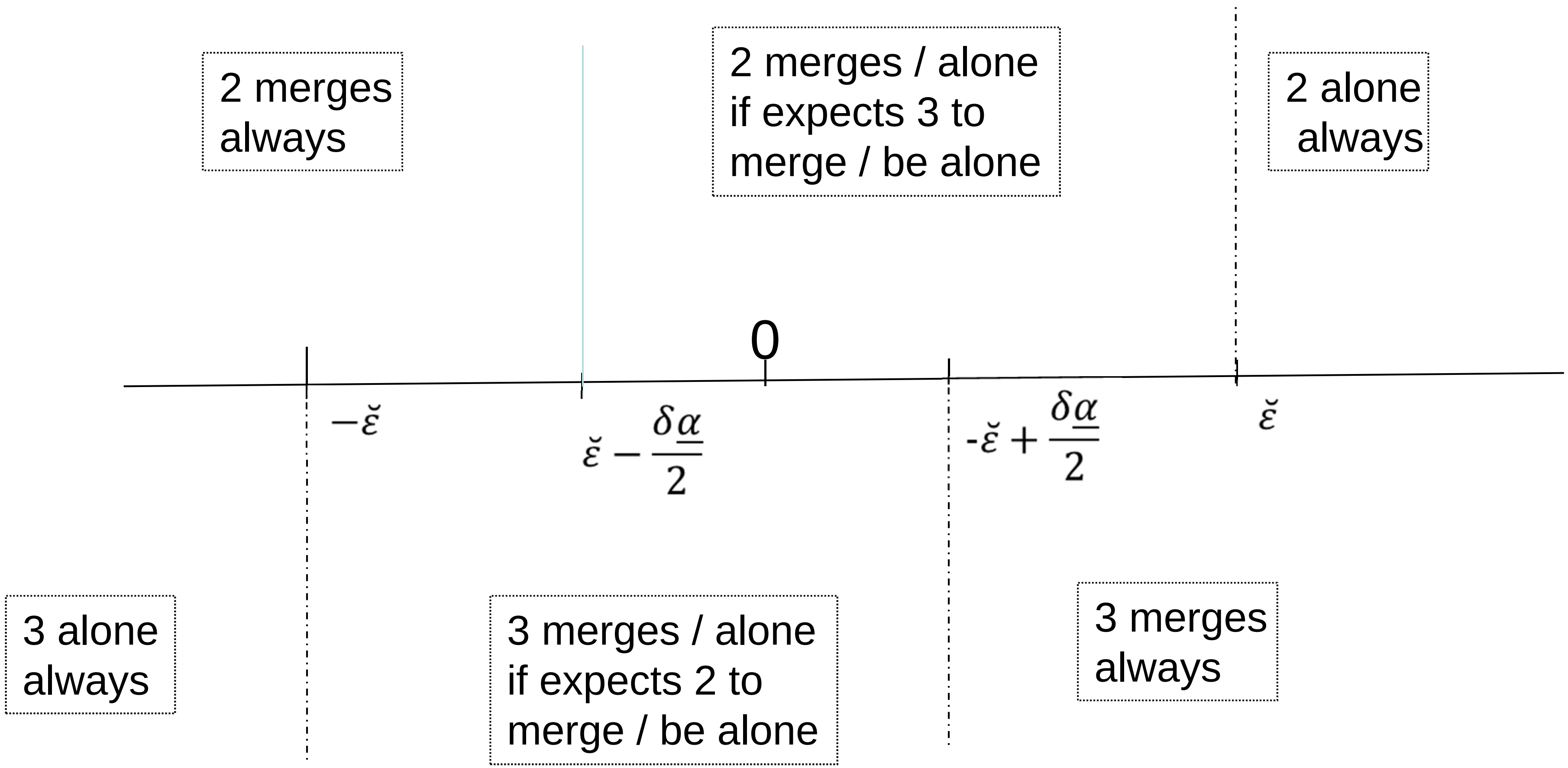

2 merges always
2 merges / alone if expects 3 to merge / be alone
2 alone always
0
$\breve{\varepsilon} - \frac{\delta \underline{\alpha}}{2}$
$-\breve{\varepsilon}$
$-\breve{\varepsilon}$
$-\breve{\varepsilon} + \frac{\delta \underline{\alpha}}{2}$
3 alone always
3 merges / alone if expects 2 to merge / be alone
3 merges always

Figure A4


2 merges always
2 merges / alone if expects 3 to merge / be alone
2 alone always
0
$-\breve{\varepsilon}$
$\breve{\varepsilon} - \frac{\delta \underline{\alpha}}{2}$
$-\breve{\varepsilon} + \frac{\delta \underline{\alpha}}{2}$
$\breve{\varepsilon}$
3 alone always
3 merges / alone if expects 2 to merge / be alone
3 merges always