Moderating Political Extremism: Single vs Dual Ballot Elections*

Massimo Bordignon† Guido Tabellini‡

This version: April 2008

Abstract

We compare single ballot vs dual ballot elections under plurality rule, assuming sincere voting and allowing for partly endogenous party formation. Under the dual ballot, the number of parties is larger but the influence of extremists voters on equilibrium policy is smaller, because their bargaining power is reduced compared to a single ballot election. The predictions on the number of parties are consistent with data on municipal elections in Italy, where cities with more (less) than 15,000 inhabitants have dual (single) ballots respectively.

*We thank Pierpaolo Battigalli, Massimo Morelli, Giovanna Iannantuoni, Francesco de Sinopoli, Ferdinando Colombo, Piero Tedeschi and participants at seminars held at the Universities of Brescia, Cattolica, Munich, Warwick, the Cesifo Workshop in Public Economics, the IIPF annual conference, the NYU conference in Florence, for helpful comments. We also thank Veruska Oppedisano and Paola Quadrio for excellent research assistance. Financial support is gratefully acknowledged, from the Italian Ministry for Research and the Catholic University for Massimo Bordignon, and from CIFAR and Bocconi University for Guido Tabellini.

†Defap, Catholic University, Milan and Ces-Ifo, Munich

‡IGIER, Bocconi University; CEPR; CES-Ifo; CIFAR

1 Introduction

In some electoral systems, citizens vote twice: in a first ballot they select a subset of candidates, over which they cast a final vote in a second ballot. The system for electing the French President, where the two candidates who get more votes in a first round run are admitted to the second round, is possibly the best known example. But variants of this *dual ballot* (or *run off)* system are used in many other countries, for example in Latin America, in the US gubernatorial primary elections, and in many local elections, including Italian municipal and regional elections (see Cox 1997 for examples, and below for the Italian case)[1]. How does the dual ballot system differ from the more common *single ballot* plurality rule election, where candidates are directly elected at the first round? In spite of its obvious relevance, this question remains largely unadressed, particularly when it comes to studying the policies enacted under these two systems.

This paper contrasts dual vs single ballot elections under plurality rule, focusing on the policy platforms that get implemented in equilibrium. We analyze a model with sincere voting where parties with ideological preferences commit to a one dimensional policy before the elections. The number of parties is partly endogenous. We start out with four parties. Before the elections, however, parties choose whether or not to merge, and bargain over the policy platform that would result from the merger. The central result is that the dual ballot moderates the influence of extremist parties and voters on the equilibrium policy. The reason is that the dual ballot reduces the bargaining power of the extremist parties that typically appeal to a smaller electorate. Intuitively, with a single ballot and under sincere voting, the extremist party can threaten to cause the electoral defeat of the nearby moderate candidate if this refuses to merge with him. Under a dual ballot this threat is empty, provided that when the second vote is cast some extremist voters are willing to vote for the closest moderate (rather than abstain). This result holds even if renegotiation among parties is allowed between the two ballots. Finally, the model also predicts that the number

[1] A variant of the run off system is the *alternative vote*, which is used for lower house state and federal elections in Australia. Here, voters only vote once, but list their preferences for all parties from first to last. The party with the fewest first preferences is eliminated and its votes are redistributed to the remaining parties according to the next preference listed. The process is repeated until there is only a party which is then declared the winner.

of parties running for elections is larger in the dual compared to the single ballot.

In light of these theoretical results, we then study data on Italian municipal elections. In Italy, Mayors in municipalities below (above) 15,000 inhabitants are elected according to a single (dual) ballot rule respectively. The data on dual ballot elections reveal that voters are indeed mobile between candidates: a relevant share of the voters supporting the excluded candidate seem to participate in the second ballot. Moreover, as predicted by the theory, the number of candidates for Major is larger under the dual ballot system, compared to the single ballot.

These results have important implications for the design of democratic institutions. Political extremism is often counterproductive, because it reduces ex-ante welfare if voters are risk averse, and because sharp disagreements could disrupt decision making in governments or legislatures. In this respect, dual ballot electoral systems have an advantage over single ballot elections, as they moderate the influence of extremist groups.

The existing literature on these issues is quite small. 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 Duvergers's Hypothesis on the equilibrium number of parties carry over to the run off system (Messner and Polborn, 2004, Cox, 1997 and, more recently Callander, 2005).[2]. Less attention has instead been devoted to the specific question of which policies are implemented in equilibrium. An exception is Osborne and Slivinsky (1996). In a model with sincere voting and ideologically motivate candidates, they study 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 in a dual ballot system. They also show that the number of active candidates is larger under the dual ballot (Callander (2005) reaches the

[2]The terminology is due to Riker (1982). Duverger's Law states that plurality rule leads to a stable two party configuration, while Duverger's Hypothesis suggests that several parties configuration should emerge from proportional representation. As the dual ballot system does not necessarily lead each party to maximize its vote share in the first round, the run off system has been traditionally grouped with proportional voting. 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 behaviour in single ballot elections under different electoral rules (Myerson and Weber, 1993; Fey, 1997). Very little is known about strategic voting in dual ballot elections (see however Cox 1997 and his cautios remarks).

opposite conclusion, however). But in keeping with the Duverger's tradition, their result is obtained in a long run equilibrium where all possibilities for profitable entry by endogeneous candidates are exausted. We are instead interested to discuss this issue in a shorter term perspective, where pre-existing policy oriented parties bargain over policy under the two different electoral systems. Finally, Wright and Riker (1989) present some empirical evidence suggesting that run off systems are indeed characterized by a larger number of running candidates.

The rest of the paper is organized as follows. Section 2 presents the basic model. Sections 3 and 4 study coalition and policy formation under single and dual ballot elections respectively, deriving the main results. Section 5 discusses an extension. Section 6 describes the Italian municipality electoral system and tests the predictions on the number of candidates. Section 7 concludes. Formal proofs are in the Appendix.

2 The Model

2.1 Voters

The electorate consists of four groups of voters indexed by $J = 1, 2, 3, 4$, with policy preferences:

$$U^J = -\left|t^J - q\right|$$

where $q \in [0, 1]$ denotes the policy and t^J is group J's bliss point. Thus, voters lose utility at a constant 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 $\frac{1}{2} \geq \lambda > \frac{1}{6}$. Groups 1 and 4 will be called "extremist", groups 2 and 3 "moderate". The assumption $\lambda > \frac{1}{6}$ implies that the electorate is polarized, in the sense that each moderate group is closer to one of the two extremists than to the other moderate group. We discuss the effects of relaxing this assumption in the next sections.

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 the shock η is to create some uncertainty about which of the two moderate groups is largest. Specifically, throughout we assume:

$$(\overline{\alpha} - \underline{\alpha}) > e \tag{A1}$$

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

Assumption (A1) implies that, for any realization of the shock η, any moderate group is always larger than any extreme group. Assumption (A2) 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. Again, we discuss the effects of relaxing these assumptions in Section 5. The realization of η becomes known at the election and can be interpreted as a shock to the participation rate.

Finally, throughout we assume that voters vote *sincerely* for the party that promises to deliver them higher utility.

2.2 Candidates

There are four political candidates, $P = 1, 2, 3, 4$, who care about being in government but also have ideological policy preferences corresponding to those of voters:

$$V^P(q, r) = -\sigma \left| t^P - q \right| + E(r) \tag{1}$$

where $\sigma > 0$ is the relative weight on policy preferences, and $E(r)$ are the expected rents from being in government. 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 split in proportion to the number of party members. Thus, $r = 0$ for a candidate out of government, $r = R$ if a candidate is in government alone, $r = R/2$ if two candidates have joined to form a two-member party and won the elections (as discussed below, we rule out parties formed by more than two party members). The value of being in government, $R > 0$, is a fixed parameter.

2.3 Policy choice and party formation

Before the election, candidates may merge into parties and present their policy platforms. We will speak of mergers between candidates as "parties", although they can be thought of as electoral cartels or coalitions 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, then the party can promise to deliver any policy lying in between the bliss points of its party members; thus, a party formed by candidates P and $P\prime$ can offer any $q^{PP'} \in [t^P, t^{P'}]$. But policies outside this interval cannot be promised by this coalition. This assumption can be justified as reflecting lack of commitment by the candidates. A coalition of two candidates can credibly commit to any $q^{PP'} \in [t^P, t^{P'}]$ by announcing the policy platform and the cabinet formation ahead of the election; to credibly move its policy platform towards t^P, the coalition can tilt the cabinet towards party member P. But announcements to implement policies outside of the interval $[t^P, t^{P'}]$ would not be ex-post optimal for any party member and would not be believed by voters (see Morelli (2002) for a similar assumption).

We assume that parties can contain at most two members, and these members have to be adjacent candidates[3]. Thus, say, candidate 2 can form a party with either candidate 3 or candidate 1, while candidate 1 can only form a party with candidate 2. This simplifying assumption captures a realistic feature. It implies that coalitions are more likely to form between ideologically closer 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. This gives moderate candidates an advantage - see below.

Candidates can bargain only over the policy q that will be implemented if they are in government. As we already said, rents from office are fixed and are split equally amongst party members[4]. Bargaining takes place before knowing the realization of the random variable η that determines the relative

[3]See Morelli (2002) for a similar modelling choice and Axelrod (1970) for a justification of this assumption.

[4]If rents were large and wholly contractible at no costs, then each coalition would form at the policy platform that maximizes the probability of winning for the coalition and rents would be used to compensate players and redistribute the expected surplus. But if rents were limited or contractible at some increasingly convex costs, then our results below would still hold qualitatively as coalitions would want to bargain over policies too.

size of groups 2 and 3, and agreements cannot be renegotiated once the election result is known.

Bargaining takes place according to a two stage process. At the first stage, candidates 2 and 3 bargain with each other to see if they can form a moderate party. Either 2 or 3 is selected with equal probability to be the agenda setter. Whoever is selected (say 2) makes a take-it-or-leave-it offer of a policy q^{23} to the other moderate candidate. If the offer is rejected, the game moves to the second stage. If the offer is accepted, then the moderate party is formed and the two moderate candidates run together at the election. Voters then vote over three alternatives: candidate 1, who would implement $q = t^1$; candidate 4, who would implement $q = t^4$; and the party consisting of candidates $\{2, 3\}$, who would implement $q = q^{23}$. Whoever wins the election then implements his policy and enjoys the rents from office.

At the second stage, the moderate and the extreme candidates, having observed the offers in the first stage, simultaneously bargain with each other (1 bargains with 2, while 3 bargains with 4) to see if they can form a moderate-extreme party. In each pair of bargaining candidates, an agenda setter is again randomly selected with equal probabilities. For simplicity, there is perfect correlation: either candidates 1 and 4 are selected as agenda setter, or candidates 2 and 3 are selected. This selection is common knowledge (i.e. all candidates know who is the agenda setter in the other bargaining pair). The two agenda setters simultaneously choose whether to make a take-it-or-leave-it policy proposal to their potential coalition partner, or to refrain from making any offer. This action is only observed by the candidate receiving (or not receiving) the offer, and not by his counterpart on the other side of 1/2. The candidates receiving the offer simultaneously accept it or reject it. If the proposal is accepted, the party is formed and the two candidates run together at the election on the same policy platform. If the proposal is rejected (or if no offer is made), then each candidate in the relevant pair stands alone at the ensuing election, and his policy platform coincides with his bliss point[5]. Again, whoever wins the election implements his policy and enjoys the rents from office.

Thus, this second stage can yield one of the following four outcomes. If both proposals are accepted, voters have to choose between two parties $(\{1, 2\}, \{3, 4\})$, each with a known policy platform. If both proposals

[5]Hence, we assume that a candidate (=party) always runs, either alone, or in a coalition with the other candidate (=party).

are rejected (or never formulated), then voters vote over four candidates $(\{1\},\{2\},\{3\},\{4\})$, each running on his bliss point as a policy platform. If one proposal is accepted and the other rejected, then voters cast their ballot over three alternatives: either $(\{1,2\},\{3\},\{4\})$, or $(\{1\},\{2\},\{3,4\})$, depending on who rejects and who accepts. Note that renegotiation is not allowed; that is, if say party $\{1,2\}$ is formed, but 3 and 4 run alone, candidates 1 and 2 are not allowed to renegotiate their common platform.

To rule out multiple equilibria in the second stage game sustained by implausible out of equilibrium beliefs, we impose the following restriction on beliefs. Call the player who receives the merger proposal the "receiving candidate". Each receiving candidate entertains beliefs about whether the other two players, on the opposite side of one half, have entered into a merger agreement or not. We assume such beliefs by each receiving candidate do not depend on the contents of the proposal that he received. Since each candidate only observes the proposal addressed to himself, and not the proposal that was made to the other receieving candidate, this is a very plausible assumption. This restriction corresponds to what Battigalli (1996) defines as independence property, and in a finite game it would be implied by the notion of consistent beliefs defined by Kreps and Wilson (1992) in their refinement of sequential equilibrium.

2.4 Electoral rules

The next sections contrast two electoral rules. Under a single ballot rule, the candidate or party that wins the relative majority in the single election forms the government. Under a closed dual ballot 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 ballot 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 dual ballot system. Section 5 discusses alternative assumptions about the relative size of extremist vs moderate voters.

3 Single ballot

We now derive equilibrium policies and party formation under the single ballot. The model is solved by working backwards.

Suppose that the second stage of bargaining is reached. Any candidate running alone (say candidate 1 or 2) does not have a chance of victory if he runs against a moderate-extremist party (say, of candidates $\{3, 4\}$ together). The reason is that, by assumption (A2), the size of voters in groups 3 and 4 together is always larger than the size of voters in group 1 or 2 alone, for any shock to the participation rate η. Moreover, given the assumption that $\lambda > 1/6$, voters in the moderate group 3 are ideologically closer to extremist candidate 4 than to the other moderate candidate, 2. Hence, all voters in groups 3 and 4 prefer any policy $q \in [t^3, t^4]$ to the policy t^2 (and symmetrically for group 1,2). In other words, the party $\{3, 4\}$ always gets the support of all voters in groups 3 and 4 for any policy the party might propose, and this is the largest group in a three party equilibrium. This in turn implies that a two-party system with extremists and moderates joined together is the only Nash equilibrium of the game. It also implies that the agenda setter, whoever he is, always proposes his bliss point, and his proposal is always accepted at the Nash equilibrium. Hence (a detailed proof is in the appendix):

Proposition 1 *If stage two of bargaining is reached, then the unique Nash 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 policy platform of each party is the bliss point of whoever happens to be the agenda setter inside each party. Hence, with equal probabilities, the policy actually implemented coincides with the bliss point of any of the four candidates.*

Note that, if all candidates run alone, the extremist candidates do not have a chance. By assumption (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 moderate candidates 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 in the former case they would not have to share the rents in case of a victory. But the two moderate candidates are caught in a prisoner's dilemma. In a four candidates situation, each moderate candidate 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. Here we model this by saying that with some probability they are agenda setters and impose their own bliss point on the moderate-extremist coalition. When this happens, the equilibrium policies reflect the policy preferences of extremist candidates, although their voters are a (possibly small) minority. But the result is more general, and would emerge from other bargaining assumptions, as long as the equilibrium policy platforms reflect the bargaining power of both prospective partners.[6]

Next, consider the first stage of the bargaining game. Here, one of the moderate candidates is randomly selected and makes a policy offer to the other moderate candidate. If the offer is accepted, the three parties configuration $(\{1\}, \{2,3\}, \{4\})$ results. If it is rejected, the two-party outcome in stage two described above is reached. Thus, the three party outcome with a centrist party can emerge only if it gives both moderate candidates at least as much expected utility as in the two party equilibrium of stage two. This in turn depends on the ideological distance that separates the two moderate candidates.

Specifically, suppose that $\lambda > 1/4$. In this case, the two moderate candidates are so distant from each other that they cannot propose any policy in the interval $[t^2, t^3]$ that would be supported by voters in both moderate groups. Hence, the centrist party $\{2,3\}$ would lose the election with certainty, and it is easy to show that both moderate candidates would then prefer to move to stage two and reach the two party system described above.

Suppose instead that $1/4 \geq \lambda > 1/6$. Here, for a range of policies that

[6] Without the restriction on beliefs introduced in the previous section, if $\lambda < 1/4$ there would be other equilibria sustained by implausible out of equilibtium beliefs. Specifically, the restriction is needed to rule out beliefs of the following kind; suppose that candidates 1 and 4 are the agenda setters; candidate 2 believes that 3 and 4 will not merge if candidate 1 proposes to 2 to merge on a platform $q^{12} \leq \hat{q}$, and he believes that 3 and 4 will merge if instead the offer received by 2 is $q^{12} > \hat{q}$. Such beliefs would induce a continuum of two party equilibria indexed by $\hat{q}$. But since the offers received by 2 reveal nothing about what players 3 and 4 are doing, such beliefs are implausible and violate the requirement of stochastic independence as discussed by Battigalli (1996).

depends on λ, the centrist coalition $\{2,3\}$ commands the support of moderate voters in both groups and, if it is formed, it wins for sure. From the point of view of both moderate candidates, this outcome clearly dominates the two party outcome that would be reached in stage two, since they get higher expected rents and more policy moderation. Hence, the centrist party is formed for sure, and its policy platform depends on who is the agenda setter in the centrist party.

We summarize this discussion in the following:

Proposition 2 *If* $1/2 \geq \lambda > 1/4$, *then the unique equilibrium outcome under the single ballot is as described in Proposition 1. If* $1/4 \geq \lambda > 1/6$, *then the unique equilibrium outcome under the single ballot is a three party system with a centrist party,* $(\{1\}, \{2,3\}, \{4\})$. *The centrist party wins the election with certainty, and implements a policy platform that depends on the identity of the agenda setter inside the centrist party.*

We can then summarize the results of this section as follows. If the electorate is sufficiently polarized ($\lambda > \frac{1}{4}$), the single ballot electoral system 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 $1/4 \geq \lambda > 1/6$, then a single ballot system would induce the emergence of a centrist party. Extremist candidates and voters lose the elections, and moderate policies are implemented.[7]

Finally, what happens if, contrary to our assumptions, $\lambda \leq 1/6$? This would mean that polarization is so low that the moderates' bliss points are closer to each other than to those of the respective extremists. In this case, the second stage game described above has no equilibrium (under the restriction on beliefs discussed in the previous section). Thus, to study this

[7]Because moderate voters are in larger number than extremists and $\lambda \leq \frac{1}{2}$, this also means that the sum of total expected losses by citizens from equilibrium policies are larger when $\lambda > \frac{1}{4}$ and the centrist party cannot be formed, than when $\lambda \leq \frac{1}{4}$ and the centrist party can be formed.

case we would need to relax the restriction on beliefs. Even in that case this second stage game would not be reached, however, since the two moderates would always find it optimal to merge into a centrist party at the first stage. The overall equilibrium would then be the same as with $1/4 \geq \lambda > 1/6$. The proof is available upon request.

4 Dual ballot

We now consider a closed dual ballot. The two candidates or parties that gain more votes in the first round are admitted to the second ballot, that in turns determines who is elected to office. To preserve comparability with the single ballot, we start with exactly the same bargaining rules used in the previous section. Thus, all bargaining between candidates is done before the first ballot, under the same rules and the same restrictions on beliefs spelled out in section 2. In particular, candidates can merge into parties only before the first ballot. Once a party structure is determined, it cannot be changed in any direction in between the two ballots. We relax this assumption in the next subsection.

The features of the equilibrium depend on other details of the model that were left unspecified in previous sections. In particular, here we add the following assumptions.

First, we decompose the shock η to the participation rate of moderate voters in two separate shocks, each corresponding to one of the two ballots. Specifically, we assume that in the first ballot the size of group 2 voters is $\bar{\alpha} + \varepsilon_1$, while the size of group 3 voters is $\bar{\alpha} - \varepsilon_1$. In the second ballot, the size of group 2 voters is $\bar{\alpha} + \varepsilon_1 + \varepsilon_2$, while the size of group 3 voters is $\bar{\alpha} - \varepsilon_1 - \varepsilon_2$. 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 section. In fact, it is convenient to define here $\eta = \varepsilon_1 + \varepsilon_2$. Exploiting the properties of uniform distributions, we obtain that the random variable η now is distributed over the interval $[-e, e]$, it has zero mean and a symmetric cumulative distribution given by

$$\begin{aligned} G(z) &= \frac{1}{2} + \frac{z}{e} - \frac{z^2}{2e^2} \quad \text{for } e \geq \ z \geq 0 \qquad (2) \\ G(z) &= \frac{1}{2} + \frac{z}{e} + \frac{z^2}{2e^2} \quad \text{for } -e \leq z \leq 0 \end{aligned}$$

Thus the first ballot reveals some relevant information about the chances of victory of one or the other moderate parties in the second ballot. This point is further discussed in the next subsection but plays no role here, since all bargaining is done before any voting has taken place.

Second, inside each extremist group, a constant fraction $\delta \leq 1$ of voters is ideologically "attached" to a candidate. These attached individuals vote only if "their" candidate participates as a candidate on its own or as a member of a party. If their candidate does not stand for election (on its own, or as a member of another party), then they abstain. This assumption plays no role under the single ballot, since all candidates always participate in the election, either on their own or inside a party. The fraction δ of attached voters should not be too large, however, otherwise there is no relevant difference between single and dual ballot elections. In particular, we assume:

$$2e/\underline{\alpha} > \delta \tag{A3}$$

We discuss the implications of this assumption below.

Finally, we retain assumptions $(A1)$ and (A2) in section 2. Clearly, these assumptions play an important role, because they determine who wins admission to the second round. In particular, assumption (A1) implies that 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.

This does not mean that moderate candidates always prefer to run alone, however. The reason is that, as spelled out above, a fraction δ of extremist voters is "attached" and will abstain in the second ballot if their candidate is not running. Merging with extremists thus presents a trade-off for the moderate candidates: a merger increases their chances of final victory, because it draws the support of these attached voters; but if they win, they get less rents and possibly worse policies. In the single ballot system, moderates faced a similar trade-off. But it was much steeper, because the probability of victory increased by 1/2 as a result of merging. Under the dual ballot, instead, the fall in the probability of victory is less drastic, and moderate candidates may choose to run alone. Whether or not this happens depends on parameter values, and on the expectations about what the other moderate candidate does.

Specifically, consider all possible party configurations before any voting has taken place, given that stage two of bargaining is reached. 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, the Appendix proves:

Lemma 1 *The probability that the moderate candidate (say 2) wins in the final round if it runs alone, given that his opponents (3 and 4) have merged, is* $1/2 - h$, *where* $h \equiv \frac{\delta\underline{\alpha}}{2e}(1 - \frac{\delta\underline{\alpha}}{4e})$.

Thus, the parameter h measures the handicap of running alone in a dual ballot system, given that the opponents have merged. Note that, by (A3), $1/2 > h > 0$. Thus, assumption (A3) 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 (A3) were violated, then the double ballot would not offer any advantage to the moderate candidates, and the equilibrium would be identical to the single ballot. 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, under (A1), (A2) and (A3), 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.

We are now ready to describe the equilibrium, if stage two of bargaining is reached. The appendix proves that it depends on whether the handicap of running alone, h, is above or below specific thresholds, $\bar{H} > \underline{H}$ and on the identity of the agenda setter inside the two prospective coalitions. More precisely:

Proposition 3 *Suppose that* $A(1)$, $A(2)$, $A(3)$ *hold and stage two of bargaining is reached. Then:*

(i) If $h < \underline{H} \equiv \frac{R}{4(2\sigma\lambda+R)}$ *the handicap of running alone is so small that both moderate candidates always prefer not to merge with the extremists. The unique equilibrium is a four-party system where all candidates run alone, and each moderate candidate wins with probability 1/2 with a policy platform that coincides with his bliss point.*

(ii) If $h > \bar{H} \equiv \frac{R}{4(2\sigma\lambda+R/2)}$, *the handicap of running alone is so large that both moderate candidates always prefer to merge with the extremists. The unique equilibrium is a two party system where moderates and extremists merge on both sides and each party wins with probability 1/2. If the moderate candidate is the agenda setter, then the policy platforms of each coalition*

coincide with the moderates' bliss points. If the extremist candidate is the agenda setter, then the policy platforms of each coalition lie in between the extremist and the moderate bliss points, and the distance between the equilibrium policy platforms and the moderates' bliss points is (weakly) decreasing in h*.*

(iii) If $\underline{H} \leq h \leq \bar{H}$*, then two equilibria are possible. Depending on the players' expectations about what the other candidates are doing, both a two party or a four party system can emerge in equilibrium. In a two party system, the policy platforms are as described under point (ii).*

These results are very intuitive. If the handicap of running alone is very large, the two electoral systems are similar, as moderates still always wish to merge with extremists, who then retain some bargaining power. But if this handicap is small, then the bargaining power of the extremists is entirely wiped out, and the dual ballot system induces that four party equilibrium which was unreachable under a single ballot because of the polarization of the electorate. In a sense, with the double ballot, voters are forced to converge to moderate platforms, by eliminating the extremist candidates from the electoral arena. In intermediate cases, anything can happen, given candidates expectations on other agents' behavior. But notice that even in a two party system, the coalitions between moderates and extremists generally form on a more moderate policy platform compared to the single ballot case. The bargaining power of moderates has increased, because a two-ballot system gives them the option of running alone without being sure losers, and this forces the extremist agenda setters to propose a more centrist policy platform.

Next, consider stage one of the bargaining game. As before, one of the moderate candidates is randomly selected and makes a take-it-or-leave-it policy offer to the other moderate. If the offer is rejected, the outcome described in Proposition 3 is reached.

As with a single ballot, the equilibrium depends on how polarized is the electorate. If voters are very polarized (if $1/2 \geq \lambda > 1/4$), 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 both moderates prefer to move to the second stage of the bargaining game. Hence, if $1/2 \geq \lambda > 1/4$ the final equilibrium is as described in Proposition 3.

Suppose instead that $1/4 \geq \lambda > 1/6$. Here the centrist party would win for sure for a range of policy platforms. But this needs not imply that the centrist party is formed, because such a party would still have to reach a policy compromise and dilute rents among coalition members. If the handicap from running alone is sufficiently small (if $h < \underline{H}$), then both moderate candidates know that the four party system emerges out of the second stage game (see Proposition 3). Hence, by linearity of payoffs, they are exactly indifferent between forming the centrist party with a policy platform of $q = 1/2$ or running alone in a four party system. A slight degree of risk aversion would push them towards the centrist party, but an extra dilution of rents in a coalition government compared to the expected rents if they run alone would push them in the opposite direction. If instead the handicap from running alone is sufficiently large ($h > \bar{H}$), then the moderates are strictly better off with the centrist party, since the continuation game would lead them to merge with the extremists. Finally, for intermediate values of the handicap (if $\underline{H} \leq h \leq \bar{H}$), both outcomes are possible, depending on players beliefs about the continuation equilibrium.

We summarize this discussion in the following:

Proposition 4 *Suppose that* $A(1)$, $A(2)$, $A(3)$ *hold.*

(i) If $1/2 \geq \lambda > 1/4$, *then the unique equilibrium outcome under dual ballot is as described in Proposition 3.*

(ii) If $1/4 \geq \lambda > 1/6$ *and* $h > \bar{H}$, *then the unique equilibrium outcome under dual ballot is a three party system with a centrist party,* $(\{1\}, \{2,3\}, \{4\})$. *The centrist party wins the election with certainty, and implements the policy platform* $q = 1/2$.

(iii) If $1/4 \geq \lambda > 1/6$ *and* $h \leq \bar{H}$, *then two equilibrium outcomes are possible under dual ballot: either the three party system with a centrist party described above, or the four party system described in part (i) of Proposition 3*

4.1 Dual ballot with endorsement

Here we allow some renegotiation to take place in between the two rounds of voting. Specifically, we retain the assumption that the policy cannot be renegotiated in between the two rounds. But we allow the excluded candidates to endorse one of the candidates admitted to the second round, if the latter approves. As a result of endorsing, the member of the winning coalitions share the rents from being in power; as in the previous sections, we assume that rents are divided in half. The restriction that policies cannot be renegotiated, although rents can be shared, is is in line with the interpretation that the policy is dictated by the identity (ideology) of the candidate, which cannot be changed after the first round. It is also coherent with the experience of many countries, including municipal elections in the Italian case (see below).

In our context, the consequence of an endorsement is to mobilize the support of the fraction δ of attached extremist voters. Under our assumption, these attached extremists vote for the neighboring moderate candidate in the second round only if there is an explicit endorsement by the extremist politician. Otherwise they abstain.

Clearly, an excluded extremist politician is always eager to endorse: by endorsing he has nothing to lose, but he can gain 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[8]. 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 cuts rents in half.

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.

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

$$\Pr(\varepsilon_2 > -\varepsilon_1) = \frac{1}{2} + \frac{\varepsilon_1}{e} \tag{3}$$

where the right hand side follows from the assumption that ε_2 has a uniform distribution over $[-e/2, e/2]$. Notice that (3) also describes the probability that candidate 2 wins if neither candidate is endorsed, as in this case, by symmetry, both moderate candidates lose the same number of attached extremist voters.

[8]In a more general dynamic setting with asymmetric information, an extremist candidate may prefer to signal his strength and refrain from endorsing, in order to strike a better deal in future elections (along the lines of Castanheira, 2004). This cannot happen in our model, as we assume that both $\underline{\alpha}$ and δ are known parameters and there is a single period.

Suppose instead that 3 has been endorsed by 4 while 2 did not seek the endorsement of 1. 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 (8), the probability that 2 wins is:

$$\Pr(\varepsilon_2 > \frac{\delta\underline{\alpha}}{2} - \varepsilon_1) = \frac{1}{2} + \frac{\varepsilon_1}{e} - \frac{\delta\underline{\alpha}}{2e} \tag{4}$$

if $\varepsilon_1 \geq \frac{\delta\underline{\alpha}}{2} - \frac{e}{2}$, and it is 0 if $\varepsilon_1 < \frac{\delta\underline{\alpha}}{2} - \frac{e}{2}$. Conversely, if 2 has been endorsed while 3 has not, then the probability that 2 wins is:

$$\Pr(\varepsilon_2 > -\frac{\delta\underline{\alpha}}{2} - \varepsilon_1) = \frac{1}{2} + \frac{\varepsilon_1}{e} + \frac{\delta\underline{\alpha}}{2e} \tag{5}$$

if $\varepsilon_1 \leq \frac{e}{2} - \frac{\delta\underline{\alpha}}{2}$ and it is 1 if $\varepsilon_1 > \frac{e}{2} - \frac{\delta\underline{\alpha}}{2}$.[9]

Hence, an endorsement increases the moderate's probability of victory by an amount proportional to the size of attached voters, $\delta\underline{\alpha}$. This gain in expected utility is offset by the dilution of rents associated with having to share power. It turns out that whether the gain in probability is worth the dilution of rents or not depends on the realization of ε_1 relative to the following threshold:

$$\breve{\varepsilon} \equiv \frac{\delta\underline{\alpha}}{2}(1 + \frac{4\sigma\lambda}{R}) - \frac{e}{2}$$

where $\breve{\varepsilon} \lesseqgtr 0$. If ε_1 is below the threshold, then the probability of victory for 2 is so low that he prefers to be endorsed even if this dilutes his rents. While if ε_1 is high enough, he is so confident of winning that he prefers no endorsement. Specifically, the appendix proves:

Lemma 2 *Irrespective of what candidate 3 does, candidate2 prefers to be endorsed by the extremist if $\varepsilon_1 < \breve{\varepsilon}$, and he prefers no endorsement if $\varepsilon_1 > \breve{\varepsilon} + \frac{\delta\underline{\alpha}}{2}$. In between, if $\breve{\varepsilon} \leq \varepsilon_1 \leq \breve{\varepsilon} + \frac{\delta\underline{\alpha}}{2}$, 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.*

Note that $\breve{\varepsilon}$ is increasing in $\sigma\lambda/R$ and $\delta\underline{\alpha}$, and decreasing in e. Thus, endorsements are more likely if the weight given to policy outcomes is large relative to rents ($\sigma\lambda/R$ is high), if there are more attached extremist voters

[9] By (A3), $\Pr(\varepsilon_2 > \frac{\delta\underline{\alpha}}{2} - \varepsilon_1) < 1$ and $\Pr(\varepsilon_2 > -\frac{\delta\underline{\alpha}}{2} - \varepsilon_1) > 0$ for any $\varepsilon_1 \in [-e/2, e/2]$.

($\delta\underline{\alpha}$ is high), or if there is less electoral uncertainty (e is small) - because this incrases the contribution of the extremist voters to final victory.

Based on Lemma 2, Proposition 7 in the appendix provides a complete description of equilibrium endorsements, in the event that both moderate candidates pass the first round.

Next, consider what happens before the first round. Again, start backwards, and suppose that the moderate candidates bargain with the extremists over party formation. Now, the moderates lose any incentive to merge with the extremists before the first round of elections. By (A2), 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, since the extremists are eager to share the rents from office. 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. Hence the option of waiting and running alone in the first round of elections is always preferred by the moderate candidates to the alternative of merging with the extremists.[10] We summarize this discussion in the following:

Proposition 5 *Suppose that stage two of bargaining is reached. 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 appendix.*

Finally, in light of this result, consider the first stage, where the two moderates bargain over the formation of a centrist party. If $\lambda > 1/4$, then as above the electorate is too polarized to sustain the emergence of a centrist party, and bargaining moves to stage 2 (and then to the four candidates running alone at the first electoral ballot). If instead $1/6 < \lambda \leq 1/4$, then the

[10] If (A2) did not hold and the moderates were unsure of passing the first round, then they might prefer to strike a deal with the extremists before any vote is taken. The equilibrium would then be similar to that of the previous subsection, without endorsements. Details are available upon request.

centrist party is feasible. By forming the centrist party the two moderate candidates win with certainty but have to share the rents in half and achieve some policy convergence. By giving up on this opportunity, the two moderate candidates know that they would end up in the equilibrium outcome described in Proposition 5. Here, each moderate candidate passes the post with probability $1/2$ on his preferred policy platform; but his expected share of rents is now strictly less than $R/2$, since with some positive probability the moderate party is forced to seek the endorsement of the extremist and this dilutes his expected rents (or alternatively, if the first ballot shock is so favorable that the moderate rejects the endorsement, his expected probability to win is less than $1/2$ since his opponent will accept the endorsement). Hence, forming the centrist party always strictly dominates the alternative of running separately at the first round of elections. The centrist party is formed with certainty on a policy platform that is tilted towards the bliss point of the agenda setter, whoever he is (since there are positive expected gains from forming the centrist party, these gains accrue to the agenda setter in the centrist party).

We summarize this discussion in the following:

Proposition 6 *(i) If* $1/2 \geq \lambda > 1/4$, *then the unique equilibrium outcome under dual ballot is as described in Proposition 5.*

(ii) If $1/4 \geq \lambda > 1/6$, *then the unique equilibrium outcome under dual ballot is a three party system with a centrist party* $(\{1\}, \{2,3\}, \{4\})$. *The centrist party wins the election with certainty, and implements a policy platform that depends on the identity of the agenda setter inside the centrist party.*

4.2 Single and dual ballot compared

Summing up, the equilibria unders single vs dual ballot entail relevant differences, both on the number of parties and on policy platforms.

Consider first the number of parties running for election, and suppose that the electorate is sufficiently polarized ($\lambda > 1/4$). The single ballot yields a two party system, each resulting from the merger of moderates and extremists. Under a dual ballot, instead, we either have four candidates for sure if endorsements are feasible, or we can have both a four candidate or a two party system in equilibrium (depending on parameter values), if endorsements are not feasible. Thus, if the electorate is sufficiently polarized, the number of candidates is generally smaller under the single ballot.

If instead the electorate is not very polarized ($1/4 \geq \lambda$), then an equilibrium with a centrist party always exists under both electoral systems and regardless of whether endorsements are feasible or not after the first ballot. Under the dual ballot and ruling out the possibility of endorsments, for some parameter values we also have a four candidate equilibrium. Thus, again the equilibrium number of parties is either the same or strictly smaller under the single ballot.

The differences are even more striking with respect to equilibrium policy outcomes. If the electorate is sufficiently polarized ($\lambda > 1/4$), then the dual ballot entails more moderate expected policies than the single ballot, irrespective of the number of parties running for election and whether or not endorsements are feasible after the first round. If instead the electorate is not very polarized ($1/4 \geq \lambda$), then equilibrium policy outcomes do not depend on the electoral rule, except in the four candidate equilibrium that also exists under the dual ballot but not in the single ballot (see above).

Below we take some of these predictions to the data. But before doing so, we discuss the robustness of these results to alternative assumptions about the relative size of moderate vs extremist groups.

5 Moderates as the smaller parties

Our assumption that there are more moderate than extremist voters is in line with the distribution of ideological preferences observed in most countries. Neverthless, the assumption plays a crucial role in the derivation of the result on policy moderation under the dual ballot. The reason is that this electoral system gives an advantage to the larger parties, who are more likely to pass the first round. This section briefly discusses whether the result on policy moderation survives under alternative assumptions about the relative size of extremists vs moderate voters.

Although anything can happen under very general assumption on the distribution of voters' preferences, there remains a reason why the dual ballot can induce policy moderation even if the moderate groups are smaller than the extremists. Moderates have an option that the extremists do not have: they can bargain with each other over the formation of a centrist party. The dual ballot can strengthen the incentives for the emergence of a centrist party, and in this way it can induce more policy moderation than the single ballot. The basic reason is that under the dual ballot what matters is not to win the first round, but to pass it and and to win the final elections. And a centrist party that manages to pass the first round has a larger probability to win the final elections, as it can then collect the voters of the excluded extremist party[11].

To illustrate this point, consider the following version of the model. Suppose 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. Suppose further that 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 centrist groups remains fixed at $\underline{\alpha}$. Everything else is kept unchanged, including the distribution of the shock, assumptions (A1-A3), and the sequence of bargaining. So, moderates first bargain among them and then (possibly) with the extremists, according to the rules described above. But we add a further assumption, namely:

$$\frac{e}{2} > (\overline{\alpha} - 2\underline{\alpha}) > 0 \tag{A4}$$

The second inequality implies that a single extremist group is larger (in expected value) than the sum of the two moderates. The first inequality implies that, at each ballot, electoral uncertainty is large enough to modify this ranking for some realization of the shock.[12] We also assume that $1/4 \geq \lambda > 1/6$, so that a viable centrist party is feasible (there exists a centrist policy platform which would be preferred by all moderate voters to the extremist bliss points). Consider then again the two electoral rules.

Under the single ballot, moderate candidates never form a centrist party at stage 1 and prefer to move to stage 2. The reason is that, under our assumptions on the distribution of the electoral shock and by (A.4), a centrist party, while viable, would always be defeated at the single ballot elections by one of the two extremists -whover happens to benefit from the electoral shock. On the other hand, if moderates decide to go on to stage 2, they now become essential players in the moderate-extremist coalitions, and it is easy to see that Proposition 1 goes through unchanged. Thus, a two party system with a coalition of extremists and moderates on each side will form, each winning with probability $\frac{1}{2}$, and each of the policies preferred by the four candidates will be implemented with equal probability.

[11] In a different modelling context, the same intuition explains the result of greater moderation of policy under the dual ballot system in Osborne and Slivinski (2001).

[12] Assumption (A4) is consistent with (A1-A2) if $\bar{\alpha}/2 > \underline{\alpha} > \bar{\alpha}/3$.

But consider now the dual ballot without endorsements. Suppose that a centrist party is formed. If one of the extremist parties is hit by a large enough negative shock (if $-\varepsilon_1 > \overline{\alpha} - 2\underline{\alpha}$), the centrist party passes the first round and goes to the second round. Given the assumptions on the distribution of ε_1, this occurs with probability $p_1 = 1 - \frac{2(\overline{\alpha}-2\underline{\alpha})}{e}$, a strictly positive number by (A.4). The centrist party will then win the final ballot if:

$$\overline{\alpha} + \varepsilon_1 + \varepsilon_2 < 2\underline{\alpha} + (1-\delta)(\overline{\alpha} - \varepsilon_1 - \varepsilon_2)$$

Let p_2 be the probability of this event, and notice that $p_2 = 0$, if $\delta \geq (2 - \frac{\bar{\alpha}}{\underline{\alpha}+\frac{e}{4}}) \equiv \overline{\delta}$ and $p_2 = 1$, if $\delta \leq \frac{2(\underline{\alpha}-e)}{(\overline{\alpha}-e)} \equiv \underline{\delta}$. Thus, for $\overline{\delta} > \delta > \underline{\delta}$ we have $1 > p_2 > 0$. In words, and quite intuitively, if the share of the attached voters is not too large, the centrist party could win the second round, although it had no chance of gaining plurality under the single ballot. The reason is that here the centrist party attracts the voters of the excluded extremist party. Next, consider the first stage of bargaining, where the moderates choose whether to form the centrist party or to negotiate with the extremists. This choice depends on their expected utility under the two scenarios. It can be shown that the moderates prefer the centrist party if $p_1 p_2 > \frac{1}{2}$. Inspection of p_2 shows this is certainly a possibility; for instance, for $\delta \leq \underline{\delta}$, this condition is satisfied if $\frac{e}{4} > (\overline{\alpha} - 2\underline{\alpha})$, that is, if the moderate voters, when joining forces, are sufficiently close in size to each extremist party.

This example is rather artificial, of course. Other examples could be constructed, with similar or different implications. But it illustrates a general insight. Moderate parties have an option that is precluded (or more difficult) to the extremists: namely, the two moderates can merge, while the two extremists cannot. The dual ballot increases the attractiveness of this option, because it allows the centrist party to gain the voters of one of the two extremists groups, if it can make it to the second round. Through this channel, the dual ballot can lead to less extreme policy outcomes even if the moderate voters are a minority[13].

[13]Moreover one would expect that the dual ballot system, in addition to make more likely the formation of a moderate party, would also strenghten the moderate bargaining power when contracting with the extremists, thus leading to more moderate policies even if the two coalitions moderate-extremist form. We cannot consider this issue here because of our assumed bargaining structure (if stage 2 of the game is reached, the moderates have already given up the possibility of forming a moderate party).

6 Evidence from Italian municipal elections

Here we test one of the theoretical predictions, namely that the number of parties standing for election is larger under the dual ballot compared to the single ballot. We exploit a reform in municipal elections in Italy, that introduced single and dual ballot elections for cities of different size. First we outline the institutions, then we analyze the data.

6.1 The electoral rules

Until 1993, municipal elections in Italy were ruled by a pure proportional parliamentary system. Voters voted for parties to elect the municipal legislature (Council); the Council then elected the Mayor and the municipal executive. Since 1993, the Major has been directly elected, with a single ballot for municipalities below 15,000 inhabitants, and with a dual ballot above this threshold. The system for electing the Council also changed as well.

Specifically, below the threshold, each party (or coalition) presents one candidate for Mayor and a list of candidates for the Council. Voters cast a *single vote* for the Mayor and his supporting Council list (they can also express preference votes over that list). The candidate who gets more votes, becomes Mayor and his list gains 2/3 of all seats in the Council.

Above the threshold, the electoral rule is more complex. Parties (or coalitions) present lists of candidates for the Council, and declare their support to a specific candidate for Mayor. But each candidate for Major can be supported by more than one Council list. There are two rounds of voting. At the first round, voters cast two votes, one for the Major and one for the Council list, and the two votes may be disjoint (i.e. they are allowed to vote for say Major A and a list supporting Major B); again, they can express a preference vote over the Council list. If a candidate for Major gets more than 50% of votes at the first round, s-he is elected. Otherwise, the two top candidates run again in a second round. In this second ballot, the vote is only over the Mayor, not the Counci lists. In between the two rounds of voting, Council lists supporting the excluded candidates for Major are allowed to endorse one of the remaining two candidates (if he agrees). Seats to the Council are allocated according to complicated rules, that generally entail a premium for the lists supporting the winning candidate for Major, but also take into account the votes received by the lists in the first round. Thus, this electoral rule is very similar to the dual ballot with endorsements described in the model, except for the complications concerning the Council lists.

6.2 The data

Our sample covers municipalities in Lombardia (the largest region in Northern Italy) from 1988 to 2004. Thus, the sample includes elections before and after the 1993 reform. To identify the effect of the electoral system, we focus on municipalities of similar size, neglecting the very small and very large municipalities. Thus, we only included municipalities with a population between 8,000 and 50,000 inhabitants (the threshold is 15,000). This gives us 118 municipalities below and 70 above the threshold. We also consider a smaller sample of cities with 12,000 to 18,000 inhabitants.[14]

How many voters are "attached"? An important assumption of the model is that at least some voters are not "attached", in that they vote for a second best candidate in the second round if their preferred candidate did not pass the first turn. If all voters were attached (if $\delta = 1$ in the model), then dual and single ballot would yield the same equilibria. To check that this assumption is consistent with the data, we compare the votes cast in the first and second round for each dual ballot election that had two rounds of voting. In Figure 1, we plot the total votes received in the first round by all the excluded candidates (on the horizontal axis), against the drop in participation between the first and second rounds (on the vertical axis). Voting for losers in the first round is substantial, ranging from 5% to almost 60%, with a median value around 30%. And the participation rate in the first round is always higher than in the second one, with an average drop in participation between the two rounds of about 15% of eligible voters. But most of the scatter plots lie well below the 45° line, meaning that in most elections the drop in participation between the two rounds is much smaller than the votes received by the excluded candidates. Thus, Figure 1 suggests that a large fraction of those who voted for losers in the first round participated again in the second round.

Under the assumptions that all those who voted in the second round also participated in the first one, and that all those who voted for the top two

[14] Unfortunately, we could not find a policy indicator with which to test the other central prediction of the model, concerning the moderating effect of the dual ballot. The reason is that the policy prerogatives of municipal governments in Italy cannot easily be captured by a policy indicator that reflects ideological differences between left and right.

candidates in the first round also participated in the second round, we can compute the fraction of attached voters (the parameter δ in the model) as the ratio between the drop in participation and the vote to the excluded candidates.[15] The median value of this ratio in our sample is 45%. Of course, a violation in one or the other assumption would result in an upward or downward bias in the estimate.

Altogether, these numbers suggest that a substantial share of those who vote for losers in the first round vote again for one of the surviving candidates in the second round, in line with the assumption of the model.

The number of parties One of the results of the theory is that the the number of parties standing for election is higher under the dual ballot than under the single ballot. Is this consistent with the evidence?

Table 1 presents summary statistics on the number of candidates for Major and number of lists for the Council, before and after the reform, and for municipalities above and below the threshold of 15,000 inhabitants. Before the reform, and after the reform in the single ballot elections (i.e., below the threshold of 15,000 inhabitants), each candidate for Major is supported by a single Council list, hence the number of candidates and the number of lists coincide. In the dual ballot elections, instead, a candidate for Major can be supported by several Council lists, hence the number of lists exceeds the number of candidates. As can be seen from Table 1, bigger municipalities have more candidates and more lists, and this is true both before and after the reform. The reform is associated with a reduction in the number of candidates (and lists) in the smaller municipalities, that adopted the single ballot. In the bigger municipalities, that adopted the dual ballot, the number of lists increased, but the number of candidates dropped. Overall, these statistics provide only limited support for the predictions of the theory. The number of lists clearly increases after the reform in the dual ballot elections, but the number of candidates does not. Dual ballot elections do have more candidates than single ballot, but the effect of the electoral rule is confounded by differences in size, and even before the reform bigger municipalities had a larger number of candidates.

To identify the effect of the electoral rule separately from the effect of

[15]In some cases, those participating in the second round may be attached voters who vote for a candidate endorsed by their party. But in our sample endorsements are a rare event.

city size, we then estimate a multivariate regression. Specifically, in Tables 2a and 2b we regress the number of lists (and candidates) on population and population squared, plus other dummy variables. Columns 1 and 3 refer to all municipalities in the sample, respectively after and before the reform. We are interested in the dummy variable *Above 15000*, that equals one if the municipality is above the threshold of 15,000 inhabitants, and 0 otherwise. Fixed effects are included for the election year, but not for the municipality because they are practically collinear with the dummy variable *Above 15000* (only one or two municipalities switch from below to above the threshold during this period). The estimated coefficient of *Above 15000* is significant and with a positive sign, as expected, but only after the reform. Columns 2 and 4 repeat the same excercise, restricting attention to municipalities of size between 12000 and 18000 inhabitants, to identify the effect of the electoral rule from the discontinuity in the threshold of 15000 inhabitants. Again, the variable *Above 15000* has a positive and significant estimated coefficient, but only after the reform. The estimates imply that switching from a single to a dual ballot increase the number of candidates of Major by one, and increase the number of Council lists by two or four, depending on the samples. The finding that the effect of the threshold is positive and significant only after the reform suggests that we are really identifying a causal effect of the electoral rule.

Finally, column 5 pools together all years and all municipalities. Here we are interested in the dummy variables *Dual ballot* and *Single ballot*, defined as 0 before the reform, and as 1 after the reform if above / below the 15000 inhabitants threshod respectively. Thus, the estimated coefficients on these variables capture the effetc of switching from the old proportional rule to the new plurality rule, dual or single ballot respectively. Here we include a municipality fixed effect, but not the year fixed effects (that are likely to be correlated with the adoption of the electoral reform). Thus, we identify the effect of the electoral rule from the within municipality variation only. The estimated cefficients reveal that switching from PR to pluralitry rule with a single ballot reduces both the number of candidates for Major as well as the number of Council lists. Switching from PR to a Dual ballot, instead, has opposite effects on the number of Council lists (that increases) and on on the number of candidates for Major (it goes down).

Overall, thus, these estimates are strongly consistent with the theoretical predictions concerning the number of parties in single ballot vs dual ballot elections.

7 Concluding remarks

We compared dual vs single ballot elections. With a highly polarized electorate, the dual ballot reduces the policy infuence of extremist groups. This happens because the dual ballot allows the moderate parties to run on their own policy platform without being forced to strike a compromise with the neighboring extremists. This also implies that the number of parties separately running for election is larger under the dual ballot than with a single ballot. The evidence from Italian municipal elections is consistent with this last prediction, on the number of candidates.

The model is built on two central assumptions. First, voters vote sincerely. Relaxing this assumption would lead to many equilibria, and would probably alter the main results. But we are not too apolegetic about our approach. While the assumption that everyone votes sincerely may be too simplistic, the opposite assumption that everyone votes strategically is even more implausible.[16] Allowing some but not all voters to be strategic would be interesting, but it would open the doors to even more ambiguities.

Second, although we study how pre-existing parties form alliances and merge, we don't allow entry of new candidates. Relaxing this assumption might yield additional insights, and our basic model could be adapted to allow for entry of new candidates, along the lines of Morelli (2002) and Osborne and Slivinski (2001). But it is not clear why this extension should radically change the main results on single vs dual ballot elections. We leave it as a promising direction for future research.

8 Appendix

Proof of Proposition 1 To formally prove Proposition 1, we need to compute the expected utilities of all parties in all possible party configurations. We need some extra notation. Let EV_i^P be the expected utility of party P under party configuration i, for $i = II, IIIa, IIIb, IV$, where: II refers to the two party configuration $(\{1,2\},\{3,4\})$, IV the four party configuration $(\{1\},\{2\},\{3\},\{4\})$, $IIIa$, the three party configuration $(\{1,2\},\{3\},\{4\})$;

[16] With reference to US elections in 1970-2000, Degan and Merlo (2006) estimate that only 3% of individual voting profiles are inconsistent with sincere voting, a figure well below measurement error. Sinclair (2005) estimates a bigger fraction of strategic voters in UK elections, but still of limited empirical relevance.

and $IIIb$,the three party configuration $(\{1\},\{2\},\{3,4\})$. These are the only possibile outcomes once the second stage of bargaining is reached. We now write down the players' expected utility in all party configurations, taking for granted that any merger will only form on platforms that satisfy (**??**).

4 parties $(\{1\},\{2\},\{3\},\{4\})$

Given assumption (A.1), the two extremist parties don't have a chance, and the election is won with probability 1/2 by one of the two moderate parties. Hence, by (1), the parties expected utilities are:

$$\begin{aligned} EV^1_{IV} &= EV^4_{IV} = -\frac{\sigma}{2} \\ EV^2_{IV} &= EV^3_{IV} = -\sigma\lambda + \frac{R}{2} \end{aligned}$$

3 parties $(\{1\},\{2\},\{3,4\})$.

By assumption (A2), groups 3 and 4 together are larger than either group 2 or group 1 alone, for all realizations of η. Moreover, given that $\lambda > 1/6$, voters in groups 3 and 4 always vote for the coalition $\{3,4\}$ rather than for candidate 2. This means that the coalition $\{3,4\}$ wins the election with certainty on the policy platform q^{34}. Expected utility for the four parties then is:

$$\begin{aligned} EV^1_{IIIb} &= -\sigma q^{34} \\ EV^2_{IIIb} &= -\sigma(q^{34} - \frac{1}{2} + \lambda) \\ EV^3_{IIIb} &= -\sigma(q^{34} - \frac{1}{2} - \lambda) + \frac{R}{2} \\ EV^4_{IIIb} &= -\sigma(1 - q^{34}) + \frac{R}{2} \end{aligned} \tag{6}$$

The other three party outcome $(\{1,2\},\{3\},\{4\})$ is symmetric to this one and can easily be computed

2 parties $(\{1,2\},\{3,4\})$.

If both coalitions form, each coalition wins with probability $\frac{1}{2}$. The equilibrium payoffs for the 4 parties depends on which policy is agreed upon in each coalition, and can be written as:

$$\begin{aligned} EV_{II}^1 &= -\sigma[\frac{q^{12}+q^{34}}{2}] + \frac{R}{4} \\ EV_{II}^2 &= EV_{II}^3 = -\sigma[\frac{q^{34}-q^{12}}{2}] + \frac{R}{4} \\ EV_{II}^4 &= -\sigma[1-\frac{q^{12}+q^{34}}{2}] + \frac{R}{4} \end{aligned} \tag{7}$$

Moderates as agenda setters.

It is easy to verify that the extremist is always better off to accept to merge with the nearby moderate than to say no, on any common policy platform and irrespective of what he expects the other two players to do. This is because under A.1 and A.3 the extremist can never win if he runs alone and the policy. Hence, if the moderates decide to merge with the extremists, they will always offer to do so at the moderates' bliss point. Comparing the previous expressions for the expected utilities under the possible party configurations, it can be shown that the moderate is also better off to merge on a platform that coincides with his own bliss point, rather than to run alone, irrespective of what the other two players on the opposite side of 1/2 are expected to do. Hence, the unique equilibrium is a two party configuration $(\{1,2\},\{3,4\})$, where each party runs on a platform that coincides with the moderate's bliss point.

Extremists as agenda setters

Comparing the previous expressions, we have:

i) $EV_{II}^2 > EV_{IIIb}^2$ for any $q^{34} \in [t^3, t^4]$ and any $q^{12} \in [t^1, t^2]$ In words, if 2 expects that 3 and 4 have merged, then he always prefer to merge with 1 on any feasible platform that does not entail losing the support of his moderate voters.

ii) $EV_{IIIa}^2 \gtreqless EV_{IV}^4$, depending on the value of $q^{12} \in [t^1, t^2]$. That is, if 2 expects 3 and 4 to run alone, then his preferred outcome depends on the common platform q^{12} that he is offered by 1. But there is a value of $q^{12} \in [t^1, t^2]$ that induces moderate party 2 to prefer to merge with 1. Clearly, EV_{IIIa}^2 is higher the closer is q^{12} to t^2.

To rule out multiple equilibria sustained by implausible beliefs by the moderates, here we have to invoke the restriction on beliefs discussed in the text (the independence property as defined by Battigalli 1996)). Namely, the moderate's (say 2) expectation about whether the other two players (3 and 4) will merge does not depend on the proposal he has received. Under this restriction, the only expectation by player 2 consistent with equilibrium is that the other two parties (3 and 4) will merge. The reason is that, as discussed above, the other agenda setter (say 4) always prefers to merge, on any policy platform acceptable by his moderate couterpart, and by ii), he can always find an acceptable proposal. Hence, the unconditional expectation that the other parties (3 and 4) will fail to merge is inconsistent with equilibrium behavior by 3 and 4. Given the unconditional expectation that 3 and 4 will merge, by i) the moderate party 2 is willing to merge with 1 on any proposed platform in the range $[t^1, t^2]$. Thus, here too, the unique equilibrium is a two party configuration, where the extremist agenda setters simultaneoulsy propose to their respective moderates to merge on a platfrom that coincides with the extremits' bliss points, and these proposals are always accepted by the moderates.[17]QED

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:

$$(1-\delta)\underline{\alpha} + \overline{\alpha} + \varepsilon_1 + \varepsilon_2 > \underline{\alpha} + \overline{\alpha} - \varepsilon_1 - \varepsilon_2 \tag{8}$$

or more succinctly if:

$$\eta \equiv \varepsilon_1 + \varepsilon_2 > \delta\underline{\alpha}/2$$

Since η is distributed over the interval $[-e, e]$ with distribution (2), this event has probability :

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

where, by (2), $h \equiv \frac{\delta\underline{\alpha}}{2e}(1 - \frac{\delta\underline{\alpha}}{4e})$.

Proof of Proposition 3 Suppose that the second stage of bargaining is reached. Extremist candidates are always better off in a two party system, since if they run alone they have no chances of winning. The issue is whether moderate candidates prefer to merge with the extremists or not, and on what policy platform.

[17]If $\lambda > 1/4$, the equilibrium would be unique even without the restriction on beliefs. The reason is that in this case the moderates would always be better off to merge on the extremist's platfrom, rather than to run alone, irrespective of their beliefs about what the other two players do.

Moderates as agenda setters Suppose first that the moderate candidates are the agenda setter inside each prospective coalition. Consider candidate 2, given that 3 and 4 have merged. If candidate 2 runs alone, as explained in the text, he wins with probability $1/2 - h$. If he wins, he implements his bliss point and enjoys the rents from office, R. If he loses, he gets no rents and the policy implemented is $t^3 = 1/2 + \lambda$. Hence, using the same notation as in the proof of Proposition 1, candidate 2's expected utility when running alone and given that 3 and 4 have merged is:

$$EV^2_{IIIb} = (\frac{1}{2} - h)R - 2\sigma\lambda(\frac{1}{2} + h)$$

If instead candidate 2 merges with 1 and implements its preferred policy, then their party wins with probability 1/2, but then candidate 2 has to share the rents from office with the other party member. Hence, candidate 2's expected utility when he merges with 1, given that 3 and 4 have merged is:

$$EV^2_{II} = (\frac{1}{4})R - \sigma\lambda$$

Comparing these two expressions, we see that 2 is indifferent between these two options if

$$h = \underline{H} \equiv \frac{R}{4(2\sigma\lambda + R)} \tag{9}$$

Hence, if $h < \underline{H}$, candidate 2 prefers to run alone, given that 3 and 4 have merged, while if $h > \underline{H}$, candidate 2 prefers to merge, given that 3 and 4 have merged.

Next, consider candidate 2's alternatives if candidates 3 and 4 do not merge. If 2 also runs alone, he wins with probability 1/2 and his expected utility is:

$$EV^2_{IV} = -\sigma\lambda + \frac{R}{2} \tag{10}$$

If instead candidate 2 merges with 1 and is the agenda setter inside his coalition, given that 3 and 4 have not merged, than party $\{1, 2\}$ wins with probability $(1 + h)$ and candidate 2's expected utility is:

$$EV^2_{IIIa} = (\frac{1}{2} + h)\frac{R}{2} - 2\sigma\lambda(\frac{1}{2} - h)$$

Comparing the last two expressions, we see that 2 is indifferent between these two options if

$$h = \bar{H} \equiv \frac{R}{4(2\sigma\lambda + R/2)} \tag{11}$$

For $h < \bar{H}$, candidate 2 prefers to run alone, given that 3 and 4 have not merged; while for $h > \bar{H}$, 2 prefers to merge with 1, given that 3 and 4 have not merged and that 2 is the agenda setter.

Comparing (9) and (11), we see that $\bar{H} > \underline{H}$. This makes sense: running alone is more attractive (i.e., the threshold of indifference is higher) if the opponents are also running alone. Hence, three cases are possible, depending on parameter values:

If $h < \underline{H}$, the handicap from running alone is so small that both moderate candidates always prefer *not* to merge with the extremists. In this case, if the second stage of bargaining is reached and the moderate candidates are drawn to be agenda setters, the equilibrium is unique and we have a four party system.

If $h > \bar{H}$, the handicap from running alone is so large that both moderate candidates always prefer to merge with the extremists. In this case, if the second stage of bargaining is reached and the moderate candidates are agenda setters, the equilibrium is again unique, and we have a two party system on the moderates' policy platforms.

Finally, if $\underline{H} \leq h \leq \bar{H}$, then multiple equilibria are possible, given that the second stage of bargaining is reached and the moderate candidates are agenda setters. Depending on the players' expectations about what the other candidates are doing, we could have both a two party or a four party system.

In all these cases, the policy platforms inside the coalitions coincide with those of the moderate candidates since the extremists are always willing to merge.

Extremists as agenda setters Next, suppose that extremist candidates are the agenda setters. Let $q^{34} \in [1/2+\lambda, 1]$ denote the policy proposal for party $\{3,4\}$ and $q^{12} \in [0, 1/2-\lambda]$ the policy proposal for party $\{1,2\}$. These policies need not coincide with the extremist candidates bliss points, since the extremists may have to deviate from their bliss points to get their proposals accepted. Our goal is to establish conditions under which such proposals might or might not be accepted by the moderate candidates. Again, we focus attention on candidate 2, under different expectations about what happens in the opposing party, since the extremists are alway better off when they merge.

Suppose that candidate 2 expects party $\{3,4\}$ to be formed on the policy platform q^{34}. Going through the same steps as above, candidate 2's expected utility if he rejects or accepts candidate 1's proposal of a platform q^{12} are respectively:

$$EV^2_{IIIb} = (\frac{1}{2} - h)R - \sigma(\frac{1}{2} + h)(q^{34} - \frac{1}{2} + \lambda)$$

$$EV^2_{II} = (\frac{1}{4})R + \frac{\sigma}{2}(q^{12} - q^{34})$$

Hence, candidate 2 is indifferent between these two alternatives for:

$$h = H(q^{12}, q^{34}) \equiv \frac{\sigma(\frac{1}{2} - \lambda - q^{12}) + R/2}{2\sigma(q^{34} - \frac{1}{2} + \lambda) + 2R} \tag{12}$$

Thus, if candidate 2 expect coalition 3,4 to be formed, he prefers to run alone (to merge) if $h < H(q^{12}, q^{34})$ (if $h > H(q^{12}, q^{34})$). Note that $H(.)$ is strictly decreasing in both arguments. Intuitively, as q^{12} increases it approaches candidate's 2 bliss point and the merger becomes more attractive; while as q^{34} increases it gets further away from candidate's 2 bliss point, and this too makes the merger more attractive for candidate 2 (since losing the election would cause more disutility).

By symmetry, if two parties are formed, in equilibrium the policy platforms agreed upon by each coalition must have the same distance from 1/2. Hence, $H(q^{12}, q^{34})$ can be rewritten (with a slight abuse of notation) as:

$$H^M(q) \equiv \frac{\sigma(\frac{1}{2} - \lambda - q) + R/2}{2\sigma(\frac{1}{2} + \lambda - q) + 2R} \tag{13}$$

for $q \in [0, 1/2 - \lambda]$ and where the M superscript serves as a reminder that 2 expects his opponents to merge. It is easy to see that $\underline{H} \leq H^M(q)$ for any $q \in [0, 1/2 - \lambda]$, where the first inequality is strict if $q < 1/2 - \lambda$ and it holds with the equal sign at the point $q = 1/2 - \lambda$. Moreover, $H^M_q(q) < 0$. Thus, the function $H^M(q)$ reaches a maximum at $q = 0$, where

$$H^M(0) = \frac{\sigma(\frac{1}{2} - \lambda) + R/2}{2\sigma(\frac{1}{2} + \lambda) + 2R}$$

The policy $q = 0$ is the point of most extreme symmetric extremism; at this choice, q^{12} and q^{34} coincide with the extremist candidates bliss points, 0 and 1 respectively. In words, as the policy q approved inside each coalition becomes symmetrically more extreme, a merger becomes less attractive for the moderate candidates, given that they expect a symmetric merger to be formed by their opponent. Hence, they will be more willing to run alone and refuse the merger, even if they expect a merger to occur in the opposing coalition.

Suppose now that candidate 2 does not expect a merger to occur in coalition 3,4. If he runs alone, either himself or the other moderate party wins with probability $\frac{1}{2}$. Hence his expected utility is the same as in (10) above.

If he instead accepts the offer from candidate 1 to form a coalition at policy q^{12}, his expected utility, given the expectation that the coalition 3,4 will not form, is:

$$EV_{IIIa}^{2} = (\frac{1}{2} + h)\frac{R}{2} - \sigma(\frac{1}{2} + h)(\frac{1}{2} - \lambda - q^{12}) - 2\sigma\lambda(\frac{1}{2} - h)$$

which is an increasing function of q^{12}. Candidate 2 will then be indifferent between accepting 1's offer or running alone, given his expectations on 3,4 , if:

$$h = H^{A}(q) \equiv \frac{\sigma(\frac{1}{2} - \lambda - q) + R/2}{2\sigma(q - \frac{1}{2} + 3\lambda) + R}$$

for $q \in [0, 1/2 - \lambda]$ and where the A superscript serves as a reminder that 2 expects his opponents to merge.Candidate 2 will then accept 1's offer if $h \geq H^{A}(q)$ and refuses it if $h < H^{A}(q)$. Clearly, $H_{q}^{A}(q) < 0$ and $\bar{H} \leq H^{A}(q)$, with equality at $q = \frac{1}{2} - \lambda$.

We are now ready to characterize the equilibrium if the extremists are agenda setters and stage two of bargaining is reached. Specifically:

If $h < \underline{H}$, then there is no feasible offer by an extremist that can induce a moderate candidate to merge with him, whatever the moderate's expectations about the other coalition. This can be seen by noting that, as discussed above, $\underline{H} \leq H^{M}(q)$, $H^{A}(q)$ for all $q \in [0, 1/2 - \lambda]$. Hence, the unique equilibrium is a 4 party system with all candidates running alone.

If $h > \bar{H}$, then the moderate candidate, say candidate 2, always prefers to merge with the extremist on at least some (though not necessarily all) feasible policy platforms, whatever his expectations on the other coalition's behaviour. This can be seen by noting that $H^{M}(q) \leq \bar{H}$ for at least some $q \in [0, 1/2 - \lambda]$, and $H^{A}(q) = \bar{H}$ at the point $q = 1/2 - \lambda$. By symmetry, candidate 2 will rationally expect that the other coalition will always be formed. He would then accept any offer q by candidate 1 such that $h \geq H^{M}(q)$. Hence, the unique equilibrium is a two party system with a merger between extremists and moderates taking place on both sides.

The extremists candidates who act as agenda setters will then impose the policy platforms closest to their bliss points, subject to getting their proposal accepted. Since $H^M(0) \lesseqqgtr \bar{H}$, the equilibrium platform in this case varies with the value of h. If $h \geq H^M(0)$, then both coalitions will form on the extremist candidates bliss points, 0 and 1 for coalitions $\{1,2\}$ and $\{3,4\}$ respectively. If $h < H^M(0)$, then coalition $\{1,2\}$ will form on the policy q^* $\in [0, 1/2 - \lambda]$ such that $h = H^M(q^*)$, while coalition $\{3,4\}$ will form on the symmetric policy $1 - q^*$. This can seen by noting that any policy $q' < q^*$ would not be accepted by candidate 2 (since by (12) $h < H(q', q^*)$), and any policy $q'' > q^*$ would be accepted by candidate 2 (since by (12) $h > H(q'', q^*)$) but suboptimal for candidate 1 who is the agenda setter. Since $H_q^M(q) < 0$, we have that $\frac{\partial q^*}{\partial h} = \frac{1}{H_q^M} \leq 0$, with strict inequality if $h < H^M(0)$. Thus, as h rises the equilibrium policy falls towards the extremists bliss point (or it remains constant if it is already at the extremist's bliss point).

Finally, if $\underline{H} \leq h \leq \bar{H}$, then two equilibrium outcomes are possible in pure strategies. (i) If the moderate candidate expects his moderate opponent to run alone, he also prefers to run alone (since $h \leq \bar{H} \leq H^A(q)$). Hence we have a four party equilibrium.(ii) If the moderate candidate expects his opponents to merge, then he also prefers to merge rather than running alone (since $\underline{H} = H^M(1/2 - \lambda) \leq H^M(q) \leq h$ for at least some q). Going through the argument in previous paragraph, the equilibrium policy platform in this case coincides with the extremist's bliss point if $h \geq H^M(0)$, and it is q^* such that $h = H^M(q^*)$ if $h < H^M(0)$. (Again, recall that $H^M(0) \lesseqqgtr \bar{H}$, depending on paramter values).QED

Proof of Lemma 2 Suppose that neither moderate candidate has been endorsed. Then the probability that 2 wins is given by (3) and 2's expected utility is:

$$(\frac{1}{2} + \frac{\varepsilon_1}{e})R - 2\sigma\lambda(\frac{1}{2} - \frac{\varepsilon_1}{e})$$

If instead candidate 2 has been endorsed while candidate 3 has not, then the probability that 2 wins is given by (5) and 2's expected utility is:

$$(\frac{1}{2} + \frac{\varepsilon_1}{e} + \frac{\delta\underline{\alpha}}{2e})\frac{R}{2} - 2\sigma\lambda(\frac{1}{2} - \frac{\varepsilon_1}{e} - \frac{\delta\underline{\alpha}}{2e})$$

provided that the first expression in brackets is strictly less than 1 and the second expression in brackets is stricly 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 probability that 2 wins is 1 and his expected utility reduces to $R/2$.[18]

Candidate 2 is indifferent between these two alternatives if:

$$\varepsilon_1 = \breve{\varepsilon} \equiv \frac{\delta\underline{\alpha}}{2}(1 + \frac{4\sigma\lambda}{R}) - \frac{e}{2} \tag{14}$$

If $\varepsilon_1 > \breve{\varepsilon}$ then candidate 2 strictly prefers no endorsement, given that 3 has not been endorsed. While if $\varepsilon_1 < \breve{\varepsilon}$ then candidate 2 strictly prefers to be endorsed, given that 3 has not been endorsed.

Next, suppose that both moderate candidates have been endorsed by the extremists. Then the probability that 2 wins is given by (3), and 2's expected utility is:

$$(\frac{1}{2} + \frac{\varepsilon_1}{e})\frac{R}{2} - 2\sigma\lambda(\frac{1}{2} - \frac{\varepsilon_1}{e}) \tag{15}$$

Suppose that 3 has been endorsed by 4, while 2 has not been endorsed. Then the probability that 2 wins is given by (4), and 2's expected utility is:

$$(\frac{1}{2} + \frac{\varepsilon_1}{e} - \frac{\delta\underline{\alpha}}{2e})R - 2\sigma\lambda(\frac{1}{2} - \frac{\varepsilon_1}{e} + \frac{\delta\underline{\alpha}}{2e})$$

provided that the first expression in brackets is strictly positive and the second expression in brackets is stricly less than 1, which 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 reduces to $-2\sigma\lambda$.[19]

Candidate 2 is then indifferent between these two options if

$$\varepsilon_1 = \breve{\varepsilon} + \frac{\delta\underline{\alpha}}{2} \tag{16}$$

If $\varepsilon_1 > \breve{\varepsilon} + \frac{\delta\underline{\alpha}}{2}$ then candidate 2 strictly prefers no endorsement, given that 3 has been endorsed. While if $\varepsilon_1 < \breve{\varepsilon} + \frac{\delta\underline{\alpha}}{2}$ then candidate 2 strictly prefers to be endorsed, given that 3 has been endorsed.

By symmetry, 3 has similar preferences, but in the opposite direction and with respect to the symmetric thresholds $-\breve{\varepsilon} - \frac{\delta\underline{\alpha}}{2}$ and $-\breve{\varepsilon}$ (eg. 3 prefers no endorsement, given that 2 has been endorsed, if $\varepsilon_1 < -\breve{\varepsilon} - \frac{\delta\underline{\alpha}}{2}$, and so on).QED

---

[18] Assumption (A3) implies that the first expression in brackets is always positive and the second one is always less than 1.

[19] By (A3), the first expression in brackets is always strictly less than 1 and the second expression in brackets is always positive.

Equilibrium with endorsements Here we 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. Specifically, under (A1-A3), we have:

Proposition 7 *(i) Suppose that $\breve{\varepsilon} > 0$. Then the equilibrium is unique and at least one of the two moderate candidates always seeks the endorsement of his extremist neighbor. If $\varepsilon_1 \in [-\breve{\varepsilon} - \frac{\delta\alpha}{2}, \breve{\varepsilon} + \frac{\delta\alpha}{2}]$ then both candidates seek the endorsement of their extremist neighbor. If $\varepsilon_1 > \breve{\varepsilon} + \frac{\delta\alpha}{2}$ then 3 seeks the endorsement while 2 does not. If $\varepsilon_1 < -\breve{\varepsilon} - \frac{\delta\alpha}{2}$ then 2 seeks the endorsement while 3 does not.*

(ii) Suppose that $\breve{\varepsilon} + \frac{\delta\alpha}{2} < 0$. Then the equilibrium is again unique and at most one of the two moderate candidates seeks an endorsement by his extremist neighbor. If $\varepsilon_1 \in [\breve{\varepsilon}, -\breve{\varepsilon}]$ then no moderate candidate seeks the endorsement of the extremist. If $\varepsilon_1 > -\breve{\varepsilon}$, then 3 seeks the endorsement of 4 while 2 seeks no endorsement. If $\varepsilon_1 < \breve{\varepsilon}$, then 2 seeks the endorsement of 1 while 3 seeks no endorsement.

(iii) Suppose that $\breve{\varepsilon} + \frac{\delta\alpha}{2} > 0 > \breve{\varepsilon}$. If $\varepsilon_1 \in [-\breve{\varepsilon}, \breve{\varepsilon}]$, then multiple equilibria are possible: either both moderate candidates seek an endorsement by their extremist neighbor or none of them does. For all other realizations of ε_1 the equilibrium is unique. If $\varepsilon_1 \in (-\breve{\varepsilon}, \breve{\varepsilon} + \frac{\delta\alpha}{2}]$ or if $\varepsilon_1 \in (\breve{\varepsilon}, -\breve{\varepsilon} - \frac{\delta\alpha}{2}]$ then both moderate candidates always seek the endorsement of the extremist. If $\varepsilon_1 > \breve{\varepsilon} + \frac{\delta\alpha}{2}$ then 3 seeks the endorsement of 4 while 2 does not seek any endorsement; and symmetrically, if $\varepsilon_1 < -\breve{\varepsilon} - \frac{\delta\alpha}{2}$ then 2 seeks the endorsement of 1 while 3 does not seek any endorsement.

Proof

Suppose first that $\breve{\varepsilon} > 0$. This then implies that $0 > -\breve{\varepsilon}$. This equilibrium is illustrated in Figure 2. If $\varepsilon_1 \in [-\breve{\varepsilon}, \breve{\varepsilon}]$, then both moderates find it optimal to seek the endorsement of the extremists, no matter what their opponent does. If $\varepsilon_1 \in (\breve{\varepsilon}, \breve{\varepsilon} + \frac{\delta\alpha}{2}]$, then candidate 3 still finds it optimal to seek the endorsment 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 $\varepsilon_1 \in [-\breve{\varepsilon} - \frac{\delta\alpha}{2}, -\breve{\varepsilon})$. Finally, if $\varepsilon_1 > \breve{\varepsilon} + \frac{\delta\alpha}{2}$ then candidate 2 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 (since a fortiori $\varepsilon_1 > -\breve{\varepsilon}$). By the same argument, the roles of 2 and 3 are reversed if $\varepsilon_1 < -\breve{\varepsilon} - \frac{\delta_\alpha}{2}$.

Next suppose that $\breve{\varepsilon} + \frac{\delta_\alpha}{2} < 0$. This then implies that $-\breve{\varepsilon} > -\breve{\varepsilon} - \frac{\delta_\alpha}{2} > 0$. This equilibrium is illustrated in Figure 3. If $\varepsilon_1 \in [\breve{\varepsilon} + \frac{\delta_\alpha}{2}, -\breve{\varepsilon} - \frac{\delta_\alpha}{2}]$, then both moderates find it optimal to seek no endorsement, no matter what their opponent does. If $\varepsilon_1 \in [-\breve{\varepsilon} - \frac{\delta_\alpha}{2}, -\breve{\varepsilon})$, then candidate 2 still finds it optimal to seek no endorsment no matter what 3 does; and given 2's behavior, 3 also finds it optimal to seek no endorsement. 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 $\varepsilon_1 > -\breve{\varepsilon}$ then candidate 2 still finds it optimal to seek no endorsement no matter what 3 does (since a fortiori $\varepsilon_1 > \breve{\varepsilon} + \frac{\delta_\alpha}{2}$), while 3 finds it optimal to seek the endorsement of 4 no matter what 2 does.

Finally, suppose that $\breve{\varepsilon} + \frac{\delta_\alpha}{2} > 0 > \breve{\varepsilon}$. This then implies $-\breve{\varepsilon} - \frac{\delta_\alpha}{2} < 0 < -\breve{\varepsilon}$. This equilibrium is illustrated in Figure 4. For $\varepsilon_1 > \breve{\varepsilon} + \frac{\delta_\alpha}{2}$ candidate 2 finds it optimal not to be endorsed, no matter what 3 does, while 3 finds it optimal to seek the endorsement of 4 no matter what 2 does (since in this case $\breve{\varepsilon} + \frac{\delta_\alpha}{2} > -\breve{\varepsilon}$). The same holds, but with the roles of 2 and 3 reversed, if $\varepsilon_1 < -\breve{\varepsilon} - \frac{\delta_\alpha}{2}$. If $\varepsilon_1 \in (-\breve{\varepsilon}, \breve{\varepsilon} + \frac{\delta_\alpha}{2}]$, then 3 still finds it optimal to be endorsed by 4 no matter what 2 does. And given 3's behavior, now 2 also finds it optimal to be endorsed. Again, the same holds, but with the roles of 2 and 3 reversed, if $\varepsilon_1 \in [-\breve{\varepsilon} - \frac{\delta_\alpha}{2}, \breve{\varepsilon})$. Finally, if $\varepsilon_1 \in [-\breve{\varepsilon}, \breve{\varepsilon}]$ then multiple equilibria are possible, since the optimal behavior of each moderate candidate depends on what his moderate opponent does. Hence, in equilibrium both seek the endorsement of their extremist neighbor or none of them does.QED

References

[1] Axelrod, R.M,, 1970 *Conflict of interest; A Theory of Divergent Goals with applications to policy,* Chicago, Markham

[2] Battigalli, P. , 1996 "Strategic Independence and Perfect Bayesian Equilibrium", *Journal of Economic Theory,* vol 70, n.1. July pp. 201-234.

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

[4] Castanheira M., 2003 "Why vote for losers?" *Journal of the European Economic Association,* 1(5): 1207-1238

[5] Cox, G., 1997, *Making votes count,* Cambridge, Cambridge University Press.

[6] Degan A. and Merlo A., 2006, "Do voters vote sincerely?" *mimeo,* Pennsylvania University.

[7] Feddersen T. 1992 "A voting model implying Duverger's Law and Positive Turnout" *American Journal of Political Science,* vol.36, n.4, pp.938-962;

[8] Fey M., 1997, "Stability and Coordination in Duverger's Law: A Formal Model of Preelection Polls and strategic voting", *The American Political Science Review,* vol.91, n.1, pp.135-147

[9] Fisichella, D., 1984, "The Double Ballot as a weapon against Anti system Parties" in Lijphart A. and Grfman B., eds., *Choosing an Electoral System: Issues and Alternatives.* New York: Praeger.

[10] Fudemberg D. and Tirole J., 1993, *Game Theory,* Cambridge, Mit Press.

[11] Kreps, D. and R. Wilson, 1992. "Sequential Equilibria", *Econometrica,* 50, pp. 863-94

[12] Messner M. and Polborn M. 2002 "Robust Political Equilibria under Plurality and Runoff Rule", *mimeo,* Bocconi university.

[13] Morelli M, 2002, "Party formation and Policy Outcomes under Different Electoral systems" Ohio State University.

[14] Myerson R. and Weber R., 1993 "A theory of Voting Equilibria" *The American Political Science Review,* vol.87, n.1, pp.102-114

[15] Osborne, M.J. and Slivinsk, A., 1996 "A model of Political Competition with Citizen-Candidates" *Quaterly Journal of Economics* 111: 65-96

[16] Palfrey T.R. 1989 "A Mathematical Proof of Duverger's Law" in *Models of Strategic Choices in Politics,* ed. P. Ordershook, Ann Arbor, University of Michigan Press, p.69-92.

[17] Persson T., Roland G., and Tabellini G. 2003 "How do electoral rules shape party structures, government coalitions, and economic policies?"NBER Working Paper No. 10176

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

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

[20] Sinclair, B.,2005, "The British Paradox: Strategic Voting and the Failure of the Duverger's Law"

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

Table 1 – Number of Candidates and Lists: Summary Statistics

	Pre reform			*Post reform*		
Population Size	< 15K	> 15K	Δ	< 15K	> 15K	Δ
N. Council lists	6.83	8.74	1.91	3.86	9.67	5.80
	(0.15)	(0.23)	(0.28)***	(0.06)	(0.19)	(0.20)***
N. candidates for Major	6.83	8.74	1.91	3.86	5.16	1.31
	(0.15)	(0.23)	(0.28)***	(0.06)	(0.11)	(0.12)***
Observations	120	73		318	221	

Standard errors in parenthesis, *** significant at 1%
Observations refer to elections
Δ = Number of (.) in municipalities > 15K minus Number of (.) in municipalities < 15K
Post reform electoral system: Dual ballot if population size is >15K, Single ballot if population size is < 15K

Table 2a: Electoral system and number of Council lists

	(1)	(2)	(3)	(4)	(5)
Dep. variable			*Number of Council lists*		
Above 15000	4.29	1.97	-0.39	-0.01	
	(0.38)***	(0.64)***	(0.46)	(0.88)	
Dual ballot					0.73
					(0.26)***
Single ballot					-3.16
					(0.17)***
Population	1.71	-57.45	3.47	-4.76	7.00
	(0.62)***	(19.01)***	(0.82)***	(26.90)	(2.40)***
Population Squared	-0.00	0.00	-0.00	0.00	-0.00
	(0.00)	(0.00)***	(0.00)***	(0.00)	(0.00)
Fixed effects	Year	Year	Year	Year	Municipality
Population size	All	12K -18K	All	12K -18K	All
Years	Post-reform		Pre-reform		All years
Obs.	539	163	193	49	732
Adj. R2	0.71	0.64	0.44	0.07	0.36
N. municipal.					188

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Population is measured in 10000 (10K)
Election year fixed effects included in columns (1-4)
Municipality fixed effects included in column (5)
In columns (1-4) the variable *Above 15000* equals 1 in the municipalities above 15000 inhabitants and 0 otherwise
In column (5) the variable *Dual ballot* equals 1 in the municipalities above 15000 inhabitants after 1992 and 0 otherwise, the variable *Single ballot* equals 1 in the municipalities below 15000 inhabitants after 1992 and 0 otherwise

Table 2b: Electoral system and number of candidates for Major

	(1)	(2)	(3)	(4)	(5)
Dep. variable			*Number of candidates for Major*		
Above 15000	0.90	1.11	-0.41	-0.01	
	(0.21)***	(0.39)***	(0.46)	(0.88)	
Dual ballot					-3.49
					(0.21)***
Single ballot					-3.06
					(0.16)***
Population	0.30	7.55	3.46	-4.76	1.53
	(0.43)	(10.92)	(0.82)***	(26.90)	(2.04)
Population Squared	-0.00	-0.00	-0.00	0.00	0.00
	(0.00)	(0.00)	(0.00)***	(0.00)	(0.00)
Fixed effects	Year	Year	Year	Year	Municipality
Population size	All	12K -18K	All	12K -18K	All
Years	Post-reform		Pre-reform		All years
Obs.	539	163	193	49	732
Adj. R2	0.33	0.25	0.44	0.07	0.58
N. municipal.					188

Robust standard errors in parentheses
* significant at 10%; ** significant at 5%; *** significant at 1%
Population is measured in 10000 (10K)
Election year fixed effects included in columns (1-4)
Municipality fixed effects included in column (5)
In columns (1-4) the variable *Above 15000* equals 1 in the municipalities above 15000 inhabitants and 0 otherwise
In column (5) the variable *Dual ballot* equals 1 in the municipalities above 15000 inhabitants after 1992 and 0 otherwise, the variable *Single ballot* equals 1 in the municipalities below 15000 inhabitants after 1992 and 0 otherwise

Figure 1

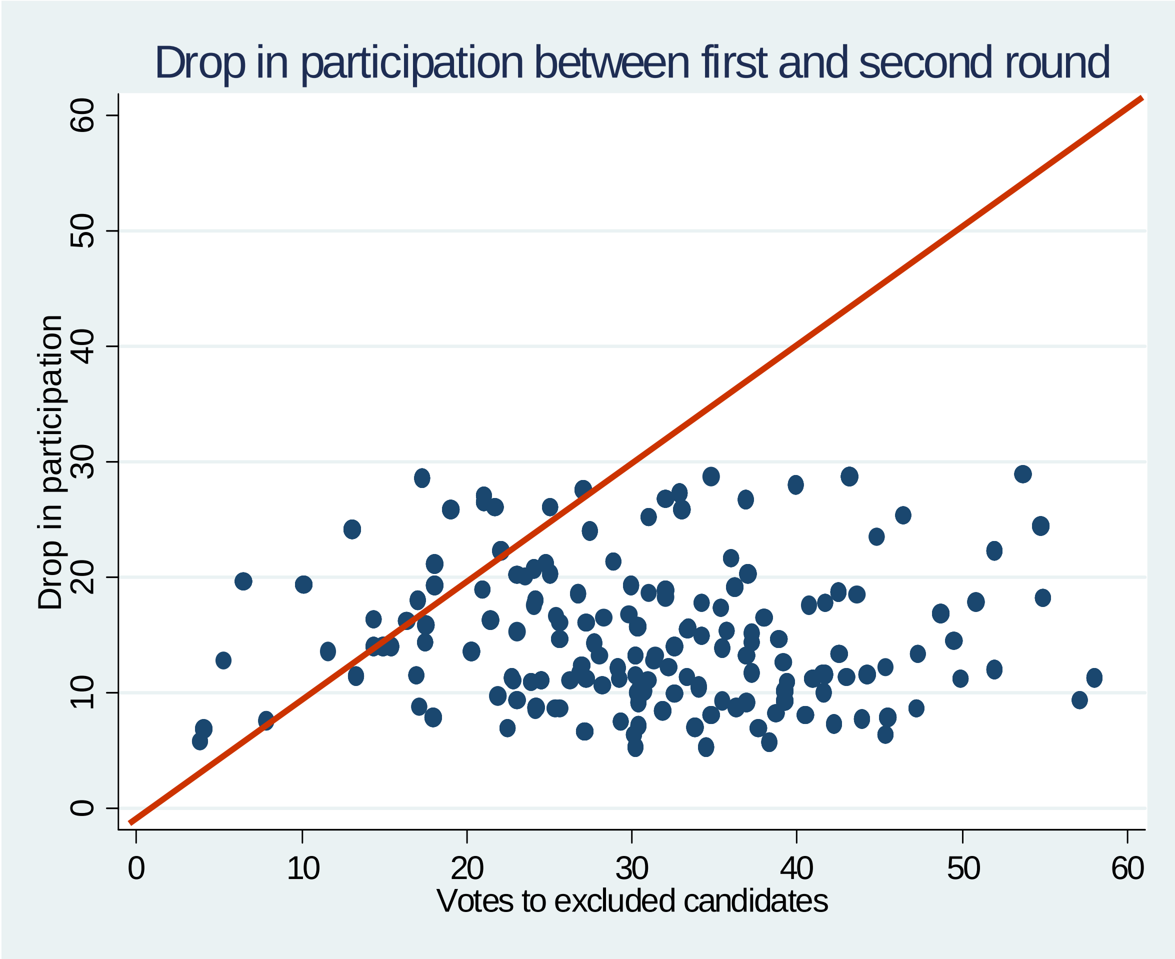

Drop in participation between first and second round
Drop in participation
0
10
20
30
40
50
60
0
10
20
30
40
50
60
Votes to excluded candidates

Figure 2

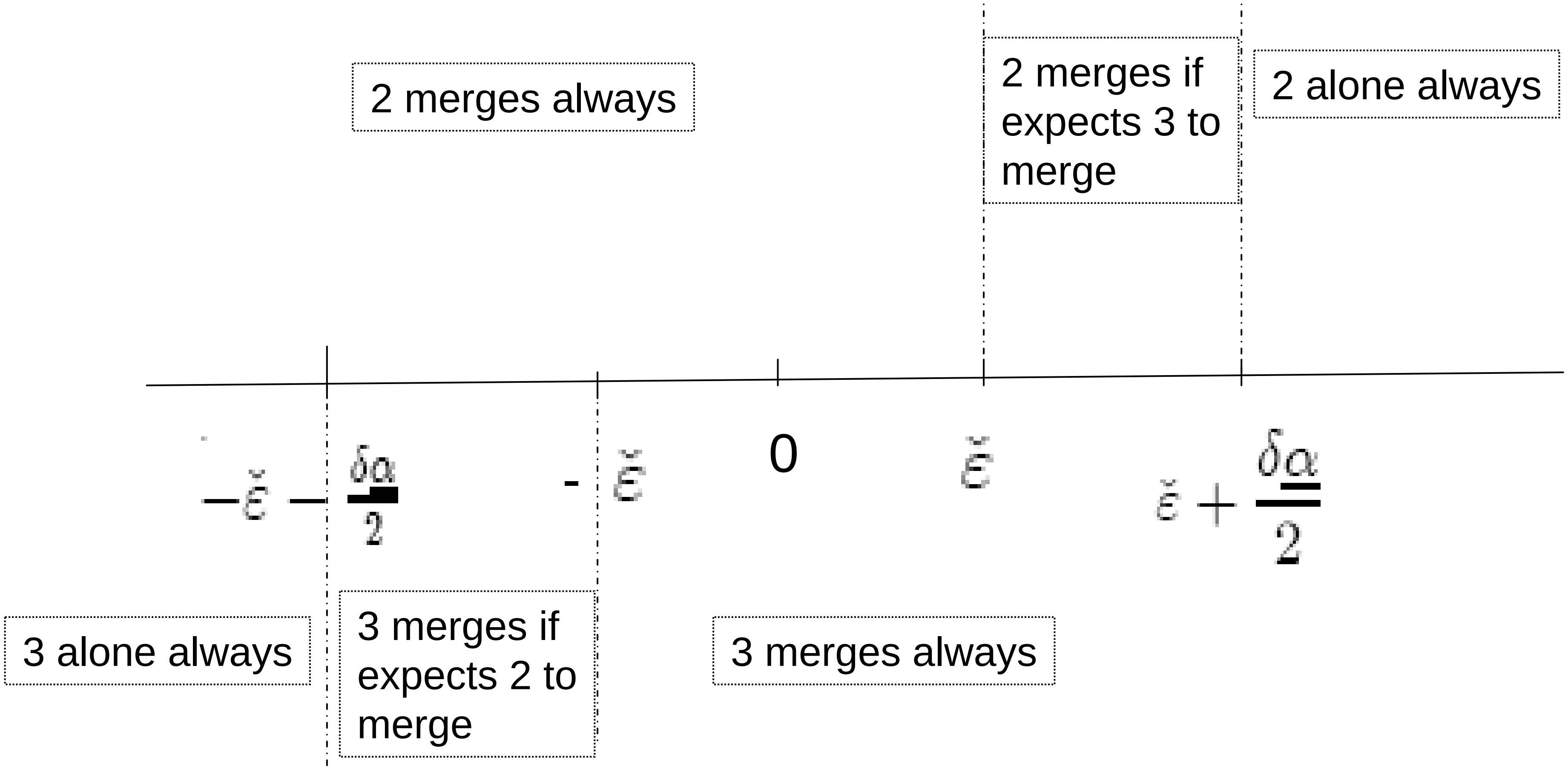

2 merges always
2 merges if expects 3 to merge
2 alone always
$-\breve{\varepsilon} - \frac{\delta\alpha}{2}$
$-\breve{\varepsilon}$
0
$\breve{\varepsilon}$
$\breve{\varepsilon} + \frac{\delta\alpha}{2}$
3 alone always
3 merges if expects 2 to merge
3 merges always

Figure 3

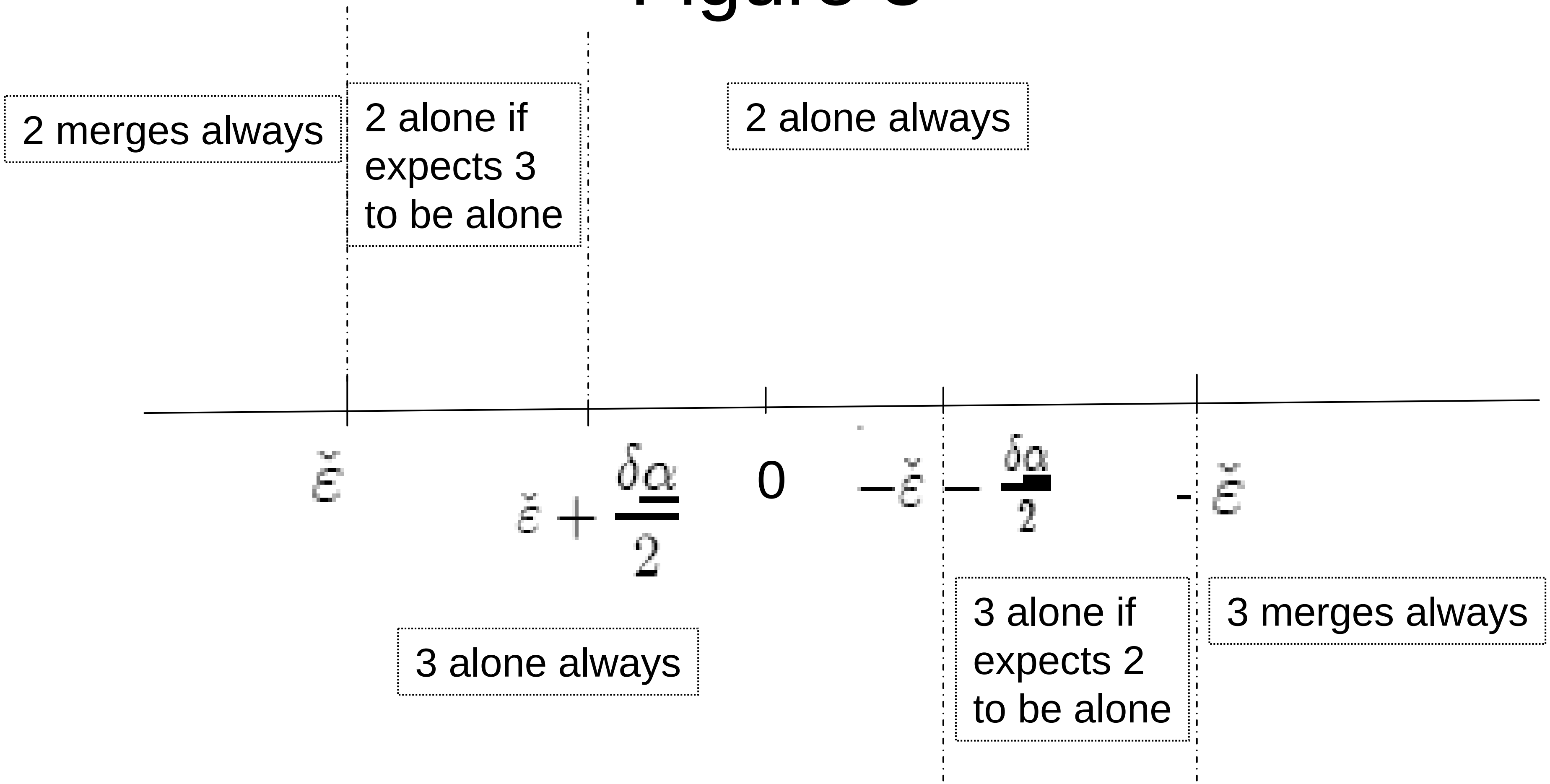
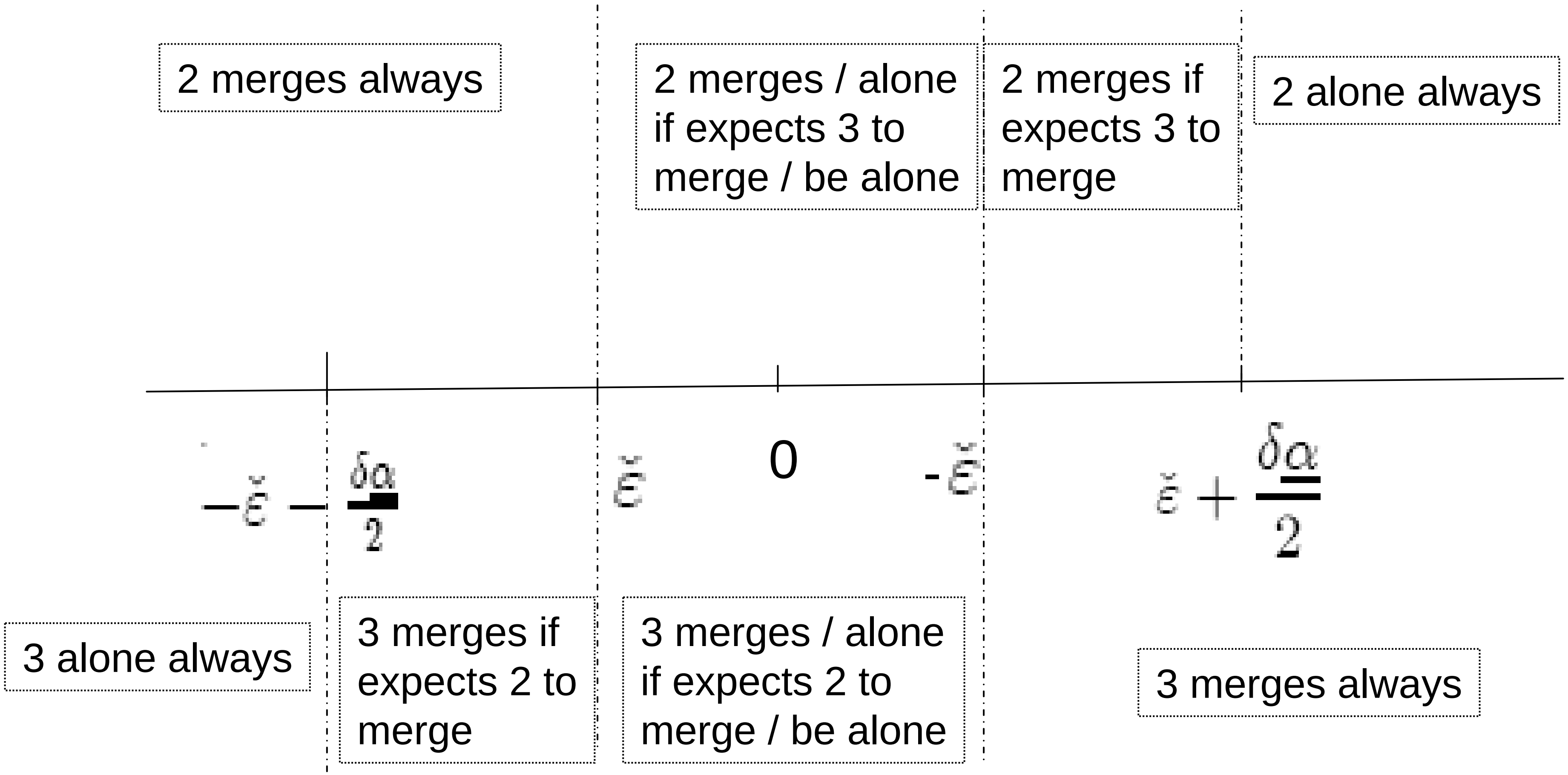

2 merges always
2 alone if expects 3 to be alone
2 alone always
$\check{\varepsilon}$
$\check{\varepsilon} + \frac{\delta\alpha}{2}$
0
$-\check{\varepsilon} - \frac{\delta\alpha}{2}$
$-\check{\varepsilon}$
3 alone always
3 alone if expects 2 to be alone
3 merges always

Figure 4


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