INNOCENZO GASPARINI INSTITUTE FOR ECONOMIC RESEARCH

Institutional Members: CEPR, NBER and Università Bocconi

WORKING PAPER SERIES

Robust Political Equilibria Under Plurality and Runoff Rule

Matthias Messner and Mattias Polborn

Working Paper n. 288

April 2005

IGIER – Università Bocconi, Via Salasco 5, 20136 Milano –Italy
http://www.igier.uni-bocconi.it

Robust Political Equilibria under Plurality and Runoff Rule*

Matthias Messner
Bocconi University

Mattias K. Polborn†
University of Western Ontario

November 1, 2002

Abstract

A central problem for the game theoretic analysis of voting is that voting games have very many Nash equilibria. In this paper, we consider a new refinement concept for voting games that combines two ideas that appear reasonable for voting games: First, trembling hand perfection (voters sometimes make mistakes when casting their vote) and second, coordination of voters with similar interests. We apply this refinement to an analysis of multicandidate elections under plurality rule and runoff rule.

For plurality rule, we show that our refinement implies Duverger's law: In all equilibria, (at most) two candidates receive a positive number of votes. For the case of 3 candidates, we can completely characterize the set of equilibria. Often, there exists a unique equilibrium satisfying our refinement; surprisingly, this is even true, if there is no Condorcet winner. We also consider the equilibria under a runoff rule and analyze when plurality rule and runoff rule yield different outcomes.

Keywords: strategic voting, runoff rule, plurality rule, equilibrium refinement, trembling hand perfection, coalition-proofness.
JEL code: D720

*We would like to thank Martin Osborne for helpful comments.

†Corresponding author: Mattias Polborn, Department of Economics, University of Western Ontario, London, Ontario, N6A 5C2, Canada; email: mpolborn@uwo.ca

1 Introduction

It is well known that there are many Nash equilibria in voting games, even if one applies standard refinements. One part of the reason for this result is that in many voting profiles, each voter's vote is not pivotal (so any deviation does in general neither increase nor decrease utility, as the outcome remains unchanged). Related to this, the concept of Nash equilibrium only considers unilateral deviations from an equilibrium strategy profile, but not the possibility of a coordinated joint action of several voters.

In this paper, we consider a new refinement concept for voting games, which we call "robust political equilibrium", and apply it to an analysis of multicandidate elections under plurality rule and runoff rule. Basically, a robust political equilibrium must satisfy two conditions: First, we want that the equilibrium is stable against the possibility that some voters "tremble" and cast votes that are recorded differently than intended by the voter. In our model, each voter chooses which candidate he wants to vote for, but the vote is only counted with probability $1-\epsilon$; with probability ϵ, the vote remains uncounted, either because the voter makes a mistake, because of technological difficulties or because of human error in vote counting.[1] We require that in equilibrium, each voter behaves optimally, given that ϵ is very small, but positive. Clearly, this implies that voters should not use weakly dominated strategies, but we show that this is not the only consequence of robustness against trembles.

Second, we want to give voters an opportunity to coordinate their actions with other voters that have the same preferences over candidates, in order to exclude equilibria that are the consequence of implausible coordination failure. In our model, there are two degrees of coordination. First, we assume that voters who have identical preferences over all candidates (and hence in all circumstances the same interests) can coordinate their voting behavior. Second, two groups of voters with non-identical interests are able to cooperate in a joint action that makes the voters in both groups better off, provided that the agreement to play the joint action is self enforcing, i.e. none of the involved groups has an incentive to cheat, if the other group keeps to their part of the agreement.

Specifically, we model the coordination possibility as follows: Every preference group (i.e, all voters who have the same preference ordering over all candidates) has a leader with exactly the same preferences over candidates (think of a party leader or president of a political action committee), who can give a recommendation to the members of his group as to for which candidate to vote. While this recommendation is of course non-binding, it is also easy to see that the members of a preference group have nothing to lose (and sometimes can gain) from coordination with other voters who have the same preferences as they themselves, and so we assume that voters follow their leaders' recommendations (provided that there is no other action that would increase a voter's utility). Effectively, leaders play a simultaneous game which recommendations to give, and we look for the coalition proof equilibrium of this game, that is, there

[1]The probability that votes are not counted was surprisingly high in the last U.S. presidential election. According to the Caltech-MIT voting technology project, 4 to 6 percent of the votes cast were not counted.

should be either no subset of leaders who could jointly switch their recommendations and increase their utility, or, if such a group of potential deviators exists, then it should be instable.

Our main results are as follows. First, we show that the possibility that some votes are not counted implies Duverger's law: In all equilibria, (at most) two candidates receive a positive number of votes. This result is similar to Palfrey (1989) and does not depend on any coalition proofness considerations. On the other hand, we show that iterated elimination of weakly dominated strategies is not sufficient to get this result.

Under plurality rule, we identify a condition that guarantees that the candidate who has the "strategic majority" is the unique outcome of a robust political equilibrium. Loosely speaking, candidate A has a strategic majority if the number of people who have A as their most preferred candidate is larger than the number of people who prefer candidate x to A, for any other candidate $x \neq A$. The requirement of a strategic majority is stronger than being the Condorcet winner, but considerably weaker than the condition that is necessary and sufficient for a game to be dominance solvable (derived by Dhillon and Lockwood (2002)). Second, even if no candidate has a strategic majority, our refinement concept often gives rise to a unique equilibrium outcome. For example, when there are three candidates and voters' preferences lead to a Condorcet cycle, a unique robust political equilibrium exists. On the other hand, it is surprising that being the Condorcet winner is not sufficient for being a robust equilibrium outcome under plurality rule.

In the case of three candidates, being the Condorcet winner is a sufficient condition for being the unique robust political equilibrium outcome under runoff rule. Of course, since having a strategic majority implies being the Condorcet winner, having a strategic majority is also sufficient under runoff rule to be the unique robust political equilibrium. Moreover, in the case of three candidates, runoff rule also generates a unique equilibrium, if no Condorcet winner exists. With more than three candidates, we provide sufficient conditions for a Condorcet winner to be a robust political equilibrium. In particular, we can show that in the important case of pure horizontal differentiation of candidates (i.e., voters have singled peaked preferences on a line), the candidate who is preferred by the median voter is the equilibrium outcome; this is not necessarily true under plurality rule.

The first one to propose a dominance-based refinement concept in voting games has been Farquharson (1969). In his *Theory of Voting*, he developed the notion of *sophisticated* voting, which, as shown in McKelvey and Niemi (1978), amounts to requiring voters to apply only sequentially undominated voting strategies.[2] Farquharson's main result is that all sequential binary voting games have a unique sophisticated voting outcome.[3] However, this very strong result does not generalize to non-binary voting procedures like plurality rule or runoff rule, which are used for the elections to most political offices. In fact, Dhillon and Lockwood's (2002) characterization shows that

[2] See also Brams' (1975) concept of *admissibility* and Moulin's (1979) analysis of dominance solvable games.

[3] Sequential binary voting games have the property that, at each stage, only two alternatives are voted on, and the loser of the stage election is eliminated from further consideration.

the class of dominance solvable plurality voting games is rather small.

De Sinopoli (2000) analyzes plurality voting games and considers the classical equilibrium refinement concepts perfection, properness and Mertens stability. However, his main objective is not to give a characterization of the set of equilibria which survive these refinements, but just to relate them to iterated elimination. Only for the set of perfect equilibria a characterization is given, which highlights the weakness of that concept in plurality voting games: Perfection can exclude at most one of the candidates (the Condorcet loser) as equilibrium outcome.

Surprisingly, there have been very few attempts in the literature to propose refinements which go beyond the iterated elimination of sequentially dominated strategies in voting games. One exception is Feddersen (1993), who applies the concept of coalition-proof Nash equilibrium to a plurality voting game with a continuum of alternatives. In his paper, citizens can vote for any policy in $[0, 1]$, and the policy that receives the most votes wins. Feddersen's main question is which policy is chosen in this model from the continuum of alternatives, if the equilibrium is required to be stable against group deviations (that are themselves stable against further deviations). In our paper, the set of feasible alternatives, interpreted as (citizen) candidates, who cannot credibly commit to any other position than their preferred one, is exogenously fixed. This also reduces the number of possible deviations. Moreover, while Feddersen deals with the case of horizontally differentiated candidates (i.e., voters have single peaked preferences), our model deals with the case of general preferences over candidates.

The paper proceeds as follows: In the next section, we introduce the political environment. In section 3, we discuss our refinement concept. In section 4, we show that, if the election technology is imperfect (i.e., sometimes not all votes are counted), then this implies Duverger's law. In section 5 and 6, we analyze 3 candidate elections under plurality and runoff rule, respectively, and provide a complete characterization of these cases. We generalize our results to elections with more than 3 candidates, for plurality rule in section 7, and for runoff rule in section 8. The last section concludes.

2 The model

2.1 Political Environment

A society of N individuals has to choose between the elements of a finite set of $\mathcal{C} = \{A, B, C, \ldots\}$. In the following, we will talk about the choices being (citizen) candidates running for office.[4] However, it should be clear that they could also be interpreted as different, mutually exclusive policy platforms.

In this section, we will present the model for the case of three candidates, $\{A, B, C\}$. This case will be the subject of analysis in sections 5 and 6. The extension of the notation to the case of more than three candidates (analyzed in sections 7 and 8) should be clear.

[4] As the literature on the citizen candidate model (see Osborne and Slivinski (1996), Besley and Coate (1997), we assume that candidates cannot commit to any other position than their favorite one.

Each voter i associates with each candidate a von Neumann Morgenstern utility value, denoted by $u_i(l)$, $l \in \mathcal{C}$. For simplicity of notation and without much restriction of generality, we assume that all individuals have a generic preference over candidates, i.e. there are no i, l, m such that $u_i(l) = u_i(m)$ for $l \neq m$.

Most of our analysis does not depend on the cardinal properties of the individuals' preference rankings, and so we will usually only refer to the ordinal preference rankings of the voters, which are shown in Table 1 for the case of three candidates. We denote the number of individuals who have j as their most preferred candidate and k as their second most preferred candidate (preference $\succ_{jk}$) by N_{jk}, $j, k \in \{A, B, C\}$.

$\succ_{AB}$	$\succ_{AC}$	$\succ_{BA}$	$\succ_{BC}$	$\succ_{CA}$	$\succ_{CB}$
A	A	B	B	C	C
B	C	A	C	A	B
C	B	C	A	B	A

Table 1: The six different preference types for three candidates

The following two definitions are very helpful: ***Core supporters of candidate*** x are those voters who rank candidate x highest, for example groups AB and AC for candidate A. The ***sympathizers of*** x ***against*** y are those voters who rank candidate x higher than candidate y, and therefore would vote for candidate x in a binary election with candidates x and y. For example, the sympathizers of A against C are the groups AB, AC and BA.

Definition 1 (Core support and sympathizers). *Let* N_x *be the number of* ***core supporters of candidate*** x*:*

$$N_x = \#\{i : x \succ_i y, \text{ for all } y \in \mathcal{C}/\{x\} \}$$

Let $S(x, y)$ *be the number of* ***sympathizers of*** x ***against*** y*:*

$$S(x, y) = \#\{i : x \succ_i y\}$$

In all of the following, we will assume for simplicity that there is a unique candidate with the largest core support. Formally, there exists $j \in \mathcal{C}$ such that $N_j > \max_{k \in \mathcal{C}/\{j\}} \{N_k\}$. We also assume that no binary election results in a tie, $S(k, l) \neq S(l, k)$. A sufficient condition for the latter is that the number of voters is odd, and the first condition is very likely to be satisfied if there are many voters. Moreover, it should be relatively clear how our results need to be modified if a knife edge case arises, however, in the interest of brevity, we concentrate on the generic cases.

The decision which candidate is chosen is taken in an election that is subject to either *plurality rule* or *runoff rule*: Under plurality rule, the candidate who receives a relative majority (i.e., more than any of his competitors) is elected; if there are two or more candidates tied for the most votes, one of them is chosen at random (with each of the tied candidates getting the same probability of winning). Under runoff rule, there are potentially two rounds of elections; in the first round, a candidate is elected if he achieves an absolute majority. If there is no such candidate in the first round, the two candidates who received the most votes in the first round face off against each other in the second round, and the candidate who gets the most votes in the second round, wins.

3 Refinement concept

In this section, we present our refinement concept for voting games and will argue why we think that it is reasonable. It is well known that in plurality voting games with three or more voters, all alternatives can be implemented as a winner in a simple Nash equilibrium, because the outcome is determined by (relative or absolute) majorities of players. This implies that one can easily find strategy profiles which are immune to deviations of single players. Though all such profiles therefore satisfy the conditions of a Nash equilibrium, many of them are not very convincing as equlibrium profiles. As an example, consider a voting game with just two candidates A and B, and suppose that a majority of voters prefer A over B. While honest voting by every voter is a Nash equilibrium in which A wins, there is also a Nash equilibrium in which B wins: If more than $N/2+1$ voters vote for B, no single voter can change the election outcome. However, in this Nash equilibrium, some of the voters (those who prefer A, but vote for B) use a weakly dominated strategy: Against any strategy profile of the other players, voting for A would never yield a worse, but sometimes a better outcome (i.e. A).

Economists have long recognized that the elimination of weakly dominated strategies might help to get rid of undesired equilibria. But while this refinement is very powerful in elections with just two candidates (or, more generally, under binary voting procedures), it is not so under plurality rule, if there are more than two alternatives. Dhillon and Lockwood (2002) show that, even if we apply elimination of weakly dominated strategies iteratively, the number of possible equilibrium outcomes is usually not decreased significantly, provided that there is not an extreme concentration in the preferences of the voters. Specifically, they show that in a three candidate race, 2/3 or more of the population have to agree on who is the worst candidate, for the game to be dominance solvable.

The intuition for this result is as follows: Initially, for each player there is only one weakly dominated strategy, namely to vote for his least preferred candidate. Even in cases where there are only three candidates, this reduction of the voters' strategy sets will typically not decrease the number of possible equilibrium outcomes. To see this, consider the three candidate voting game described by table 2. The numbers in the second row of the table give the number of voters with the preference profile in the corresponding column. In this example, an absolute majority of voters ranks candidate C highest, and honest voting is a Nash equilibrium in which C wins.

However, there is also an equilibrium in which a majority of people (for example, all) vote for A. Moreover, no player uses a weakly dominated strategy in this equilibrium, because only voting for one's worst outcome is weakly dominated. There is even an equilibrium in which the *Condorcet loser* B (i.e. the candidate who would lose any binary election) is the winner of a plurality voting game: All voters with preference types AB and BA vote for B, and voters of preference type CA split 3:3 between voting for C and A; again, no player uses a weakly dominated strategy here.

$\succ_{AB}$	$\succ_{BA}$	$\succ_{CA}$
2	3	6
A	B	C
B	A	A
C	C	B

Table 2: Example 1

The common justification for the elimination of weakly dominated strategies is that they are strictly suboptimal to play, if the other players make mistakes with a positive probability. However, the elimination of weakly dominated strategies is not the only implication of robustness against a small probability that players make mistakes when voting. To see this, suppose that each voter in Example 1 can only *intend* to give a vote for candidate ℓ, and will actually succeed to do so only with probability $1-\epsilon$; with probability ϵ, his vote will not be counted, either because the voter makes a mistakes that renders is ballot invalid, or because there are technical difficulties that prevent the vote from being registered.[5] As ϵ goes to zero, there can be no equilibrium in which candidate B wins: Consider the strategy profile from above, where all voters of types AB and BA intend to vote for B, three voters of type CA intend to vote for A, and the remaining 3 intend to vote for candidate C. The probability that candidate B wins in this perturbed game is[6]

$$(1-\epsilon)^{11} + \binom{11}{1}(1-\epsilon)^{10}\epsilon + o(\epsilon), \tag{1}$$

where $o(\epsilon)$ are terms that go to zero quicker than ϵ. If one of those type CA voters who intend to vote for A switches his intention to vote to C, the probability that B wins is instead[7]

$$(1-\epsilon)^{11} + \binom{11}{1}(1-\epsilon)^{10}\epsilon\left\{\frac{6}{11} + \frac{1}{2}\frac{5}{11}\right\} + o(\epsilon). \tag{2}$$

For sufficiently small ϵ, we can ignore the $o(\epsilon)$ terms, and the expression in (2) is smaller than (1). Also, the probability that candidate A wins goes down, and thus the new probability distribution first order dominates (for a voter of type CA) the old one. Therefore, a type CA voter who votes for A could increase his utility by changing his vote to candidate C, and hence the voting profile in the example is not an equilibrium

[5] As mentioned in the introduction, the probability that a vote is not registered as intended seems rather high in the U.S.: In the report of the Caltech-MIT voting technology project, it is estimated that this probability lies between 4 and 6 percent. However, we will here consider the limit of $\epsilon \to 0$.

[6] B wins in any case if at most one vote is not counted.

[7] As explanation of the terms in curly brackets: Given that exactly one voter trembles, B wins, (a) if the trembler is one who intended to vote for A or C (probability 6/11), (b) if the trembler intended to vote B, but B wins the coin toss that becomes necessary after two candidates get the same number of votes.

in the perturbed game, if the probability of making an error is sufficiently close to zero (but positive).[8] The consequence of this example for the refinement concept in this paper is that we want to require robustness against the possibility that voters make mistakes with a small probability, and not just eliminate weakly dominated strategies.

Example 1 also highlights the important role which the failure of coordination between voters plays for the multiplicity of equilibria: Only because individuals of group CA could not coordinate on voting for their most preferred candidate, the other two could win. Such extreme cases of coordination failure seem somewhat unrealistic for voting games. When groups of voters have similar interests, they often find ways to coordinate their behavior. For example, voters often organize in political parties or political action committees (PAC's), or at least they listen when these groups give a recommendation for whom to vote, and it appears that one of the roles of these organizations is to coordinate the political actions of its members and supporters.

This said, the question becomes how to incorporate cooperative group behavior into the equilibrium concept. Game theorists have considered different possibilities. Aumann (1959) proposed to allow for all possible ways to form coalitions. According to the so called "strong Nash equilibrium" concept which he defined, a strategy profile is a strong equilibrium, if no coalition of players has an incentive to deviate jointly. Note that, in many cases, there is no voting profile that is a strong equilibrium.

Bernheim, Peleg and Whinston (1986) have argued that allowing for *all* possible coalitions to deviate is a too strong requirement. Instead, they propose to consider only deviations that are internally stable, i.e. deviations of coalitions which would not trigger further deviations by subcoalitions of the deviating coalition. Our model will capture the spirit of their coalition proofness concept, although there are differences in some details that we will discuss below.

The timeline of our model is as follows (see Figure 1): Each of the $(\#\mathcal{C})!$ different

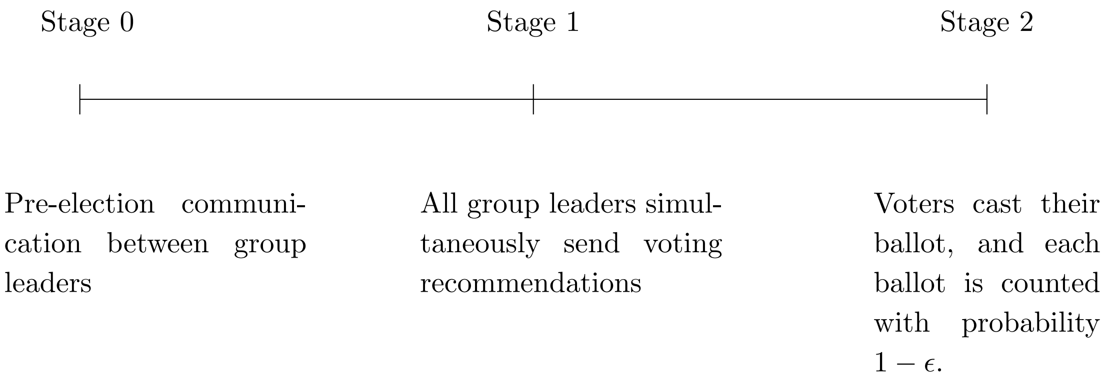

Figure 1: Timeline

preference groups has a leader; for example, in the case of 3 candidates, there are 6

[8] It is easy to check that there is also no other profile in which B wins; these would have to have either four type CA voters voting for A and two for C (in which case one of the type AB voters should deviate to vote for A), or vice versa, in which case one of the type CA voters who vote for A should switch to C.

different preference groups. Each leader can give a recommendation to the other voters of the same preference type (we refer to these voters of the same type also as the leader's "constituents"). In stage 0, before giving his recommendation, a leader can talk with other leaders about the election, but not sign binding contracts with them. In stage 1, the leader of group i sends a message m_i to his constituents, interpreted as a recommendation to vote. All leaders send their messages simultaneously, and each voter can only observe the message sent by his own leader. We assume that the leader can only send a message $m \in \{\mathcal{C}\}$, interpreted as his recommendation to vote for that candidate.[9]

In stage 2, voters of type i intend to vote according to the recommendations of their leader, and each voter succeeds in doing so with probability $1-\epsilon$; with probability ϵ the vote he casts is not counted. We assume that errors among different voters are independent random variables.

The assumption that voters try to follow their leaders' recommendations appears quite reasonable here. Leaders have exactly the same preferences as their constituents, so there is no agency problem. Moreover, in many cases it seems realistic that party or PAC leaders are better informed about the relative sizes of the other preference groups and hence are in a better position than ordinary voters to devise the most promising voting strategy for the group. The assumption excludes babbling equilibria in which people just ignore their leaders.

Before we state our equilibrium concept, we need some notation. Let D be the set of leaders, and D_C be a generic subset (C stands for coalition, and D_C is a group of leaders who could deviate from equilibrium play). For any set X, m_X is the strategy profile played by players belonging to X, and m_{-X} is the strategy profile played by all other players.

Part 1 of the following Definition 2 captures the notion that the equilibrium strategies should still be optimal, if there is a small chance that each *voter* makes a mistake when implementing his vote intention. Part 2 of this definition is our notion of coalition proofness. Consider an equilibrium, and suppose that there is a deviation coalition D_C of players (i.e., leaders) who could benefit from deviating together to some other strategy profile m'_{D_C}; then, the definition requires that this deviation is *internally instable.* By this, we mean that there is a subset of the deviating group that, given that the deviation occurs, could benefit further by changing their actions once more.

Definition 2 (Robust political equilibrium). *A strategy profile $M^* = (m_1^*, \ldots, m_{\#\mathcal{C}!}^*)$ is a robust political equilibrium, if both of the following conditions hold:*

1. *There exists an $\bar{\epsilon} > 0$ such that for all $\epsilon \leq \bar{\epsilon}$, there exists a profile M^ϵ such that m_j^ϵ is an optimal response to m_{-j}^ϵ, and M^ϵ converges to M^* for $\epsilon \to 0$.*

[9] Alternatively, one could assume that the leader can give personalized recommendations (i.e., an individual message to each group member), so that the leader could effectively split the votes of his constituents by recommending that some voters vote for, say A, and the other voters for B. Note, however, that this requires a much more sophisticated communication technology and is probably not a realistic possibility in large elections.

2. *Let $\mathcal{D}_C$ be the set of all coalitions of leaders that would benefit, for $\epsilon = 0$, from a joint deviation (D_C is a generic element of that set):*

$$\mathcal{D}_C = \{D_C | u_j(m'_{D_C}, m_{-D_C}) > u_j(m_{D_C}, m_{-D_C}), \text{ for all } j \in D_C\}$$

For every $D_C \in \mathcal{D}_C$, there exists a subset $I \subset D_C$ and a "deviation from the deviation", m''_I, such that

$$u_k(m''_I, m'_{D_C \setminus I}, m_{-D_C}) > u_k(m'_{D_C}, m_{-D_C}) \text{ for all } k \in I.$$

In our voting context, the effect of robustness is as follows. Consider an equilibrium in which, say, candidate A wins. First, if there is no coalition that is sufficiently big to change the outcome and would like to do so (i.e. everyone in the coalition would benefit from this change), then the equilibrium is robust.

Second, if there is a coalition that would prefer that, say, candidate B wins instead of A and that has the size that is necessary to make this happen (given that the other people continue to play their equilibrium strategy), then robustness of the equilibrium depends on whether this deviation is internally stable. Suppose that there is a subset of the deviation coalition consisting of one or more members who prefer candidate C even more than B, and if the first coalition deviates to B, then the subset has the possibility of making candidate C win. However, in this case, their partners in the original coalition will not believe that they will actually vote for B, and so the proposed deviation to B would be *internally instable*. If all coalitions that would be able and willing to change the outcome of the election are internally instable, then the equilibrium supporting A as outcome is stable. On the other hand, if there is (at least) one coalition that can and would like to make B win, and is internally stable, then the equilibrium supporting A as outcome is not robust.

Note that part 2 of definition 2 only considers coalition deviations that have the potential to change the outcome of the election, with a probability that is bounded away from 0, even as the probability that voters make a mistake, ϵ goes to 0. Put differently, we do not care about deviations that (almost certainly, for low ϵ) only change the relative positions of the losing candidates of the election. This is an intuitively plausible restriction: If forming a deviation coalition takes some effort, then this effort is only worthwhile, if there is a non-trivial chance of influencing the outcome of the election.

4 Effects of Trembling under Plurality Rule

In this section, we will start our analysis by considering just the effects of the trembling part of our definition of a robust political equilibrium: Which voting equilibria are robust to the introduction of a small chance ϵ that each voter's vote will not be counted? This question is interesting for two reasons: First, we will show that there is a difference between our way of modelling trembling and the iterated elimination of weakly dominated strategies that was analyzed by Dhillon and Lockwood (2002). Second, it shows which of our results depend on robustness against this type of election technology problems, and which depend on the second part of our definition of a robust political equilibrium, which assumes that some coordination among voters with the same or similar interests is possible.

Our first result is a simple and well known consequence of the fact that voting for the least preferred candidate is, under plurality rule, a weakly dominated strategy. Therefore, as long as there is some positive probability that one's vote is decisive, one should vote for one of the other candidates.

Lemma 1. *If $\epsilon > 0$, then voting for the least preferred candidate is a strictly dominated strategy for every voter.*

Proof. Omitted (obvious). □

Note that this is usually the only result that can be obtained from (even iterated) elimination of weakly dominated strategies. Dhillon and Lockwood (2002) show that a further round of elimination is possible in a three candidate race if and only if at least 2/3 of the voters agree on who is the worst candidate. Suppose that more than 2/3 of the voters think that candidate C is the worst of the three candidates. Then, since none of them will vote for C, C cannot win and even people who like C better would know that they would waste their vote, if they voted for C. Effectively, this reduces the three candidate race to a race with 2 viable candidates, and this game is dominance solvable.

Which additional consequences can be derived in our framework? Consider any equilibrium, and remember that each vote is counted with probability $1 - \epsilon$. Hence, there is a positive probability that the candidates who are on the second and third place in the voting *intentions* actually end up winning the election, because sufficently many votes cast for the other candidates are not counted. However, provided that ϵ is small, the probability that the third placed candidate wins is actually *much smaller* than the probability that the second placed candidate wins. Formally, "much smaller" means that the winning probability of the third placed candidate is a term of higher order (in ϵ) than the probability that the second placed candidate wins.

In a sense, a citizen who votes for a candidate who is not among the top 2 candidates (according to vote intentions) wastes his vote and should not do so, if he has a strict preference over the top 2 candidates: For example, if he prefers the candidate with the most votes over his pursuer, he should vote for him, because his vote is much more likely to be pivotal in a decision between the two candidates who receive the most vote intentions than in a decision involving any other candidate.

Lemma 2. *Let $v_i(s)$ be the number of vote intentions candidate i gets in the voting profile $s = (s_1, \dots, s_N)$, and let $V_1 = \max_i v_i(s)$ be the number of vote intentions for the top placed candidate, and let V_2 be the number of vote intentions for the second placed candidate, and so on.*

If individuals have a strict preference over all candidates, there exists $\bar{\epsilon} > 0$ such that a strategy profile with $V_1 \geq V_2 > V_3 > 0$ cannot be an equilibrium for all $\epsilon \in (0, \bar{\epsilon})$

Proof. See Appendix. □

The intution for this result is very similar to the one in Palfrey (1989). In his model, voter types are drawn from some known distribution. For the limit case of very many voters, he shows that there can only be equilibria in which just two candidates receive a positive measure of votes. While the details of Palfrey's and our model are different, there as here, the intuition for this result is that, conditional on a tie occurring, it is much more likely to be a tie between the most popular and the second most popular candidate than any other pair of candidates.

The next question is whether there can be any equilibria with two candidates being tied for the second place ($V_1 \geq V_2 = V_3$). Intuitively, this is also very unlikely. For such a constellation to be an equilibrium, it would have to be true that ***every*** voter for Candidate 2 prefers C_2 over C_1 (otherwise, switching the vote to candidate 1 is optimal), and C_1 over C_3 (otherwise, switching the vote to candidate 3 is optimal). Moreover, each citizen voting for C_2 would have to prefer a 50/50 lottery over Candidates 2 and 3 over the outcome which gives Candidate 1 for sure. The reason for this latter requirement is as follows: Given that Candidate 1 loses the election, Candidate 2 and 3 have equal chances to be the winner. On the other hand, in almost all cases in which C_1 loses the election, he loses because he ties with C_2 and C_3 and loses the following coin toss.[10] But in this case, a single voter who casts a vote for C_2 or C_3 could swing the election in favor of C_1 (who is his second-most preferred candidate). For this not to happen, it is necessary that ***each*** of the C_2 voters prefers the lottery between C_2 and C_3 to the certain outcome C_1.

The following assumption is sufficient to guarantee that there are only equilibria in which exactly two candidates receive a positive number of votes.

Assumption 1. *Assume that in the polity, for each of the* $(\#\mathcal{C})!$ *possible preference type, there is at least one individual who has that preference.*

Proposition 1. *Suppose that, in addition to the assumptions made in the model section, assumption 1 is satisfied.*
In every equilibrium, there are only two candidates who receive a positive number of votes. Every voter votes in the same way as he would, if only these two candidates were eligible (i.e., honestly according to his preference over the two candidates with a positive number of votes).

Proof. See Appendix. □

Proposition 1 has two interesting implications. First, it shows that the Condorcet loser can never be the equilibrium outcome, and, in fact, cannot receive any votes in equilibrium. Remember from Example 1 that this is not true if we only require iterated elimination of weakly dominated strategies. Second, it shows that plurality rule has a tendency to generate election outcomes in which all votes are concentrated on only two candidates. This is, of course, Duverger's law, a well known stylized fact of political systems operating under plurality rule. Note again that, only applying iterated elimination of weakly dominated strategies is *not sufficent* to generate this result.

[10] Any other scenario in which C_1 loses is less likely by at least one order of ϵ.

5 Equilibrium under Plurality Rule with 3 candidates

In this section, we analyze the equilibrium under plurality rule, for the special case of 3 candidates. The results so far were only a consequence of the imperfect election technology. We now turn to the additional results that follow in combination with the coalition proofness requirement.

We first state the concepts of a Condorcet winner (a candidate who would win all binary elections in which he participates) and Condorcet loser (a candidate who would lose all binary elections in which he participates), using our notation.

Definition 3. *Candidate x is a* ***Condorcet winner (CW)****, if and only if he would beat both other candidates, respectively, in a binary election; formally,*

$$S(x, y) > S(y, x) \text{ for all } y \neq x.$$

Candidate z is a Condorcet loser if he would lose against both of his opponents, respectively, in a binary election:

$$S(z, y) < S(y, z) \text{ for all } y \neq z.$$

The next concept is (to the best of our knowledge) new in the literature. We say that candidate x has a *strategic majority against candidate y*, if his core supporters outnumber those who prefer candidate y to x. This is a stronger requirement than "x beats y in a binary election", because there are people who would vote for x in an $x - y-$contest, but are not core supporters of x (they rank some other candidate z highest, but have $x \succ y$ as their second preference).

Definition 4 (Strategic majority). ***Candidate x has a strategic majority against candidate y****, if and only if*

$$N_x > S(y, x).$$

If $N_x > \max_{y \neq x} S(y, x)$ (i.e., candidate x has a strategic majority against all other candidates), we say that ***candidate x has a strategic majority.***

For example, candidate A has a strategic majority against B, if types AB and AC outnumber types BA, BC and CB. Clearly, if a candidate's core support is larger than half of the population, then he also has the strategic majority, but the reverse is not true; for example, in Table 3, candidate A has the strategic majority without being supported by more than half of the population.

AB	AC	BA	BC	CA	CB
23	25	10	17	12	13

Table 3: A has a strategic, but not an absolute, majority

On the other hand, if candidate A has the strategic majority, he is also the Condorcet winner (because $S(x, y) \geq N_x$), but not vice versa. For example, in Table 4, A

AB	AC	BA	BC	CA	CB
17	18	17	17	25	6

Table 4: A is the Condorcet winner, but has no strategic majority against B or C

is the Condorcet winner, as A would win against B and C, in binary elections, respectively, but in this example, A does not have a strategic majority against either B or C.

With the above concepts defined, we can now begin to state those results that rely on the coalition proofness part of our definition of a robust political equilibrium. The equilibrium depends on whether one candidate has a strategic majority (against both other candidates), or, if no candidate has a strategic majority, whether or not a Condorcet winner exists. Note that these three cases partition the parameter space: Every polity falls in exactly one of these cases. We start with the case that a candidate has the strategic majority.

Proposition 2. *Suppose candidate A has a strategic majority.*[11]

1. *If $N_{BA} > S(C, A)$, then the following is the unique robust political equilibrium: (A, A, B, B, A, B)*

2. *If $N_{CA} > S(B, A)$, then the following is the unique robust political equilibrium: (A, A, A, C, C, C)*

3. *If neither $N_{BA} > S(C, A)$ nor $N_{CA} > S(B, A)$, then both (A, A, B, B, A, B) and (A, A, A, C, C, C) are robust political equilibria.*

In any case, the unique equilibrium outcome is a victory of candidate A.

Proof. From Proposition 1, we know that we need to check the 3 voting profiles in which only two candidates get votes.

Consider first the profile (A, A, B, B, A, B). If $N_{CA} < S(B, A)$, no single leader can gain by switching his recommendation, provided that ϵ is sufficiently small. Moreover, there is no coalition that would like and is able to change the outcome of the election (the coalition of groups BC, CA and CB could change to C, but cannot change the outcome, because $N_A > S(C, A)$ by assumption). Hence, if $N_{CA} < S(B, A)$, (A, A, B, B, A, B) is a robust political equilibrium.

If $N_{CA} > S(B, A)$, then group CA is better off switching their vote to C: If CA switches to C, the probability that candidate C will win, is a term of order $N_A - N_{CA}$ in ϵ, while the probability that candidate B wins is a term of order $N_A - S(B, A) > N_A - N_{CA}$. Therefore, the increase in the chance to get C elected dominates the increase of the risk that B will be elected, and so CA is better off switching to C.

A similar consideration shows that the profile (A, A, A, C, C, C) is a robust political equilibrium, if and only if $S(C, A) > N_{BA}$. Since at most one of the conditions $N_{BA} > S(C, A)$ or $N_{CA} > S(B, A)$ can be satisfied, the first two claims follow.

[11]Certainly, if other candidates have a strategic majority, proposition 2 applies with the appropriate changes.

Last, consider the profile (B, C, B, B, C, C). This is not a robust equilibrium, since groups AB and AC could jointly switch to vote for A, which changes the outcome of the election to A and is certainly internally stable. □

Proposition 2 provides a condition that guarantees the victory of a particular candidate in a plurality election. Intuitively, the condition that candidate A has a strategic majority means that there is no candidate who can get more votes than A in an equilibrium, if A's core supporters vote for A; the most votes any other candidate x can get in this constellation is given by the number of people who prefer x to A, $S(x, A)$.

The condition of a strategic majority against both competitors is weaker than requiring that the candidate should have a core support that is larger than half of the electorate (which of course is also a sufficient condition). On the other hand, the condition of a strategic majority is stronger than being the Condorcet winner; indeed, we will show in the following propositions that there are cases in which a Condorcet winner, who does not have a strategic majority, is not guaranteed the victory.

Recall example 1 in Table 2, where there were equilibria in which the winner was not the candidate preferred by the majority of people. These equilibria survived iterated elimination of weakly dominated strategies, as the condition that 2/3 of the electorate agree on who is the worst candidate (which makes the game dominance solvable) was not satisfied. With our concept of a robust political equilibrium, only candidate C (who has a strategic majority against both of his opponents) can be the equilibrium outcome in this example.[12]

If no candidate has a strategic majority, two major cases have to be distinguished: Either, there is no Condorcet winner, or one of the candidates is the Condorcet winner (and has at most a strategic majority against only one of the other candidates).

We turn first to the case that there is no Condorcet winner among the 3 candidates (remember that this implies that no candidate has a strategic majority). For this case, we have a surprising result: There is a unique robust equilibrium. The winner is the candidate, who in pairwise comparison would win against the candidate with the largest core support. Hence, if A is the candidate with the largest core support, and A beats B, B beats C and C beats A, then a victory of candidate C is the unique robust equilibrium outcome.

Proposition 3. *Suppose that there does not exist a Condorcet winner, and let Candidate A have the largest core support, and let $S(A, B) > S(B, A)$, $S(B, C) > S(C, B)$ and $S(C, A) > S(A, C)$.*[13] *The voting profile (A, A, A, C, C, C) is the unique robust political equilibrium, and the equilibrium winner is candidate C.*

Proof. Let as start by checking that (A, A, A, C, C, C) is a robust political equilibrium: Groups AB, BA and BC would like to jointly deviate to B, but once groups BA and

[12] The reader may wonder whether, if 2/3 of the electorate agree on who is the worst candidate, one candidate must have a strategic majority. This is not the case, and we will discuss the precise relation below, at the end of this section.

[13] Clearly, given that there is no Condorcet winner, these assumptions as to which group is the largest and who beats whom are without restriction of generality. We just call the candidate with the largest core support "A", the candidate who would be beaten by A "B", and the candidate who beats A, "C".

BC vote for B, it is in group AB's interest to switch back to vote for A. Hence, the proposed deviation is not internally stable, and so (A, A, A, C, C, C) is a robust political equilibrium.

Now consider the two other candidate profiles for a robust political equilibrium. First, (A, A, B, B, A, B) can be upset by a deviation of groups BC, CA and CB to vote for C; moreover, this coalitional deviation is internally stable, because if group BC switch back to B, the only effect is that A wins. Hence, (A, A, B, B, A, B) is not robust. Second, (B, C, B, B, C, C) can be upset by a deviation of groups AB and AC, and this deviation is obviously internally stable. □

The intuition for the result is as follows: In order to break the (A, A, A, C, C, C) equilibrium in which C wins, sufficiently many voters have to switch to candidate B; the only coalition that can achieve such a majority (and has a motivation for replacing C with B) are groups AB, BA and BC. However, once this group changes their vote to candidate B, type AB voters would have an incentive to change again, to A; by assumption, A has the largest support and therefore wins, if the voting profile is (A, A, B, B, C, C). Hence, the leader of AB cannot credibly convince his potential coalition partners that he would in fact give a recommendation to his constituents to vote for B. In a nutshell, group A is too large for being trusted in deviations.

On the other hand, the other candidates for equilibrium can be upset by (internally stable) deviating coalitions. The profile (B, C, B, B, C, C) can be upset simply by A's core supporters (which is certainly stable), and the profile (A, A, B, B, A, B) can be upset by a joint deviation of types BC, CA and CB; this deviation is stable, because group BC (the only ones in the coalition who are not completely happy with the outcome after the proposed deviation) could only switch back to vote for B, but this would only hand the victory to A. Hence, the joint deviation is internally stable.

We now turn to the case that a Condorcet winner exists. In Proposition 2, we have dealt with the case that the Condorcet winner has a strategic majority against both of his opponents (the example in Table 3 provides a numerical example for this situation). The following Proposition 4 covers the remaining two cases that either the Condorcet winner has a strategic majority against only one competitor, or none at all (an example for the latter case is given in Table 4).

Proposition 4. *Suppose that a Condorcet winner exists, but that the Condorcet winner does not have a strategic majority. Let A be the candidate with the largest core support.*

1. *The CW is the unique equilibrium outcome in the following two cases:*
   (a) *A is the Condorcet winner and has a strategic majority against one of the other candidates.*
   (b) *B or C is the CW, and A is not the CL.*
2. *If A is the CL, then both B and C are robust equilibrium outcomes.*
3. *If A is the CW, but does not have a strategic majority against either opponent, then no robust equilibrium exists.*

Proof. 1. (a) Suppose A is the Condorcet winner, and has a strategic majority against B, but not against C. Note that this implies that $S(C, A) > S(B, A)$, and also $N_{CA} > N_{BA}$. Consider the profile (A, A, A, C, C, C) as a candidate for an equilibrium, in which A wins (by being the Condorcet winner). A deviation of BA, BC and CB fails to deliver a victory for candidate B, since A has a strategic majority against B.

For uniqueness of the outcome, note that from Proposition 1 and the assumption that A is the CW, the only possible equilibrium strategy profile in which A does not win, is (B, C, B, B, C, C). However, this is not a robust equilibrium, because groups AB and AC could jointly deviate to A, making A the winner (and this deviation coalition is also clearly internally stable).

(b) Suppose B is the Condorcet winner. Consider the strategy profile (A, A, B, B, A, B), in which candidate B wins. A deviation of groups AC, CA and CB could potentially change the outcome of the election to C, but is internally instable: AC could switch back to A and, since A's core support is the largest, A would win the election; hence the coalition that would deviate to C is not internally stable. The case that C is the Condorcet winner is symmetric.

To see that there cannot be an equilibrium in which another candidate wins, note first that C cannot be the winner (because it is the CL). For A to be the winner, the strategy profile would have to be (A, A, A, C, C, C). However, this is not a robust equilibrium, because groups BA, BC and CB could together switch to vote for B, and this deviation is internally stable.[14]

2. Suppose B is the CW and A is the CL. The proof that (A, A, B, B, A, B) is a robust equilibrium, with B as winner, is the same as above. Now consider the voting profile (A, A, A, C, C, C), in which C is the winner (since A is the Condorcet loser). A deviation from this profile of groups AB, BA and BC to vote for the Condorcet winner B would win, but is internally instable, as group AB would have an incentive to change back to A. Therefore C can be supported as the outcome of a robust equilibrium.

3. Suppose A is the Condorcet winner, but has neither a strategic majority against B, nor against C. Consider the profile (A, A, B, B, A, B). Groups BC, CA and CB can jointly deviate to C such that C wins. Moreover, this deviation is internally stable: If BC switches back to B, then A wins. Hence, (A, A, B, B, A, B) is not a robust equilibrium. An analogous consideration shows that (A, A, A, C, C, C) is not a robust equilibrium. Also, (B, C, B, B, C, C) is not robust, as AB and AC could jointly switch to A and make A win. □

Proposition 4 completes our results for plurality rule elections with three candidates. As perhaps expected, in most cases, the Condorcet winner is a robust political equilibrium, even if he does not possess a strategic majority against both competitors.

[14]If CB switches back to C, A wins, and BA and BC certainly have an incentive to stick to the deviation agreement.

If the Condorcet winner possesses the strategic majority against one competitor, then there is a robust equilibrium in which the Condorcet winner and the other candidate (against whom the Condorcet winner does not have a strategic majority), get a positive number of votes. This way, a deviation to the third candidate cannot be successful, because the Condorcet winner's core support can beat that deviation alone.

Things are more difficult, if the Condorcet winner does not have a strategic majority against either competitor. In this case, for any voting profile where the Condorcet winner and another candidate receive votes, there is always a coalition that strictly prefers the third candidate to the Condorcet winner and can beat the CW, if only the CW's core support votes for him. The question then turns out to be whether the deviation to the third candidate is stable.

If the Condorcet winner does not have the biggest core support, then the profile where the CW and the candidate with the biggest core support get votes, is stable, because a deviation to the third candidate is internally instable: It involves one of the groups that prefer the candidate with the largest core support. If the other groups involved in a possible coalition actually switch, then the supporters of the candidate with the largest core support would like to vote for their most preferred candidate; this makes the deviating coalition unsustainable.

However, if the CW is also the candidate with the largest support, then there exists a deviation that is internally stable and sufficient to beat the CW. Hence, the CW is not a robust equilibrium here, and no robust equilibrium exists.

There are instances in which a candidate, who is not the CW, can be a robust equilibrium outcome (apart from the CW who is also an equilibrium outcome in these cases). Equilibrium multiplicity arises, if the two smaller groups are in a certain sense closer to each other than to the big group, so that the candidate with the largest core support is at the same time the Condorcet loser. In this case, the two smaller groups have essentially a coordination problem with two solutions; moreover, because no one from the smaller groups can trust the big group in a possible deviation coalition, both candidates that do not have the largest core support are robust equilibrium outcomes.

As an example, consider the situation in Table 5. There are only 3 preference groups, the left (with a preference for the L candidate, then the M candidate and the R candidate comes at the bottom of their preference ranking), the right (which is symmetric to the left group), and a moderate group that likes its own candidate best, and prefers the left to the right candidate.

$\succ_1$	$\succ_2$	$\succ_3$
40	16	44
L^*	M	R^*
M	L^*	M
R	R	L

Table 5: Single peaked preferences with two robust equilibrium outcomes

While candidate M is a Condorcet winner in this voting game, both M and L are robust equilibrium outcomes. Basically, the first and the second group know that they need to coordinate on either candidate L or M, because they have no chance of winning against R, if they are separated. If they coordinate on L, the right group cannot lure the moderate group away from L into voting for M, by promising to vote for M themselves, because if that succeeded, the right group would have no interest to vote for M, but would rather vote R, so that R wins.

While equilibria in which another candidate than the CW wins may appear a bit strange at first glance, they are interesting for the explanation of some stylized facts. For example, in Britain, the Liberal Democrats are often perceived as being located between the Conservatives and the Labour Party. In constituencies where neither the Conservative candidate nor the Labour candidate have a core support that constitutes an absolute majority, the candidate of the Liberal Democrats should be the Condorcet winner. Nevertheless, Liberal Democrats win relatively rarely. This phenomenon can, of course, also be explained in a model where all people vote "honestly", if the Liberal Democrats core support is relatively small. However, the important result here is that such an equilibrium can survive even in an environment in which voters are rational and have some ability to coordinate their voting behavior with other voters.

The example in Table 5 is also interesting to illustrate the difference of our conception of coalition proofness from the standard game theoretic concept, if individual voters (rather than party leaders) are conceived as the players of this game. A coalition of 11 voters of group 2 and all voters of type 3 could upset the equilibrium in which candidate L wins, and this coalition is internally stable: Since there are still 49 citizens who vote for L, group 3 voters do not have any incentive to further deviate, as they would not be enough to change the outcome to R.

Note, however, that this deviation requires a very high degree of coordination. First, all individual voters (rather than just party leaders) need to know the composition of the electorate, and thus, need to know how many group 2 voters can switch to vote for M such that group 3 voters still have an incentive to stick to the agreement. Second, *all* group 2 voters need to know about the deviation proposal: Otherwise, group 3 voters would have an incentive to contact all group 2 voters and claim to each one individually that he is a part of the 11 group 2 voters that should switch to M; but if that worked, group 3 would of course vote for R. Third, given that all group 2 voters need to know about the deviation proposal, those who are supposed to still vote for L effectively make a mistake (they should rather vote for M such as to minimize the probability that L wins by chance). Summing up, we think that this is a degree of sophistication that is probably not realistic in most election settings, and thus we believe that our way of requiring coalition proofness only against deviations of entire groups is more reasonable.

It is also interesting to compare the condition of a strategic majority with the results of Dhillon and Lockwood (2002). They show that a necessary and sufficient condition for a voting game with three candidates to be dominance solvable is that at least 2/3 of the electorate agree on who the worst candidate is; if this condition is satisfied, a CW exists and the winner is the CW. Perhaps surprisingly, the set of polities that satisfy this Dhillon-Lockwood condition is not a subset of the economies where one candidate has a strategic majority. As an example, consider the following Table 6.

AB	AC	BA	BC	CA	CB
32	0	35	2	30	1

Table 6: Polity satisfies the DL condition, but no candidate has a strategic majority

Here, more than 2/3 of the electorate agree that candidate C is the worst choice. The CW is candidate A, but he has neither a strategic majority against B nor against C.

This raises the question whether there is always a unique robust political equilibrium in our model if the Dhillon-Lockwood condition is satisfied. The following Proposition 5 shows that this is in fact the case.

Proposition 5. *Suppose that more than* 2/3 *of the electorate agree on which candidate is the worst choice. Then, a CW exists and the unique robust equilibrium outcome is the CW.*

Proof. See Appendix □

6 Runoff rule with 3 candidates

Before we proceed to analyze the equilibrium under plurality rule with more than three candidates, we now turn to an analysis of the equilibrium under runoff rule, with three candidates. While this is slightly unsystematic, the advantage is that we can immediately compare the differences between plurality and runoff rule for the case of three candidates, that we can completely characterize under both rules. The reader who prefers a more systematic approach can, without any problem, skip to the next section (on the general candidate case under plurality rule) and come back to this point afterwards.

For runoff rule, we need to slightly adjust our equilibrium concept to a dynamic game setting. Fortunately, the outcome of the second round of elections is very easy to characterize, because honest voting is a weakly dominant strategy in the second round. Therefore, in case that the first round of elections does not produce a winner with an absolute majority, we just substitute the continuation utility from the result of the corresponding second round of elections. We then require that the first period election game (constructed in the way described) has a robust equilibrium.

In runoff election games, even fewer strategies are weakly dominated than in plurality elections. The reason is that it may make sense to vote for your least preferred candidate in the first round, if you believe that this makes it easier for your most preferred candidate to win in the runoff. Moreover, it is possible that vote-splitting by group leaders can be a reasonable strategy, for the same reason; if the preferred candidate cannot win in the first round, then rather than maximizing the preferred candidate's vote share in the first round, it might make sense to vote for another candidate in the hope that this candidate is easier to beat in the second round. In order to keep the results comparable to the plurality rule case, we will continue to assume that leaders can give only one simple recommendation for one of the candidates and effectively act as block voters who cast all votes of voters with the same preference for one of the candidates.

We will start our analysis with an analysis of the three candidate case, if there is a Condorcet winner. The following proposition will show that the Condorcet winner is the unique equilibrium outcome in this scenario.

Proposition 6. *Consider an election with three candidates. If candidate A is the Condorcet winner, then A is the unique robust political equilibrium outcome.*

Proof. We will first show that no other outcome can arise in equilibrium. Suppose first that A's core support is larger than 1/3 of the electorate. In this case, A's supporters can force A into a runoff (with winner A), unless all four other groups vote for the same candidate.[15] However, in a strategy profile where all other groups vote for B, group C1 (which ranks B lowest) does not behave optimally, and similarly for a profile in which all other 4 groups vote for C, so these profiles cannot be equilibria.

Suppose now that A's support is smaller than 1/3, and (without loss of generality) let C be the Condorcet loser. Assume that the outcome is B. Then, however, a deviation which results in groups A1, A2 and C1 voting for A in the first round is successful and internally stable.[16] Last, assume that the winner is C; note that this must be a victory in the first round of elections, and hence requires that groups A2, B2, C1 and C2 vote for C.[17] If $N_A > N_{B1}$, then a coalition of A1, A2 and B1 voting for A is successful and internally stable; if $N_B > N_{A1}$, then a coalition of A2, B1 and B2 voting for B is successful and internally stable. Since (by Assumption 1) at least one of the inequalities $N_A > N_{B1}$ or $N_B > N_{A1}$ must be satisfied, C cannot be a robust equilibrium.

We now show that A is the equilibrium outcome. Consider the strategy profile (A, A, B, B, A, B) in the first round of the election. The only feasible deviation is groups B2, C1 and C2 switching to C, but this will produce either a runoff between A and C (with winner A, so the deviation is not successful), or a runoff between B and C with winner B, which makes group C1 worse off than the original strategy profile. Hence, the strategy profile (A, A, B, B, A, B) in the first round of the election is a robust equilibrium. □

Under a runoff rule, every Condorcet winner is the unique equilibrium outcome, while with plurality rule, there are cases in which another candidate than the CW is the outcome of a robust equilibrium, and there are also cases in which a CW, but no robust equilibrium exists. What is the reason for these differences between plurality rule and runoff rule?

First, consider the equilibrium in the example of Table 5 in the previous section, in which candidate L (who is not the CW) wins under plurality rule. The reason that this is an equilibrium is that the potential deviation coalition, i.e. groups 2 and 3 voting for

[15] Since A is the CW, neither groups BA, BC and CB together, nor groups BC, CA and CB together are a majority of the population.

[16] The only group that could have any incentive not to keep to the coalition deviation is C1, but if they vote for C, the only possible change of the outcome would be that C loses in the runoff to B, so C1 does not have an incentive to deviate from the coalition.

[17] Since C is the CL, the support of C against A and the support of C against B are less than 50%.

M, is not internally stable, as group 3 would have an incentive to switch their votes to R. Under a runoff rule however, the coalition deviation of groups 2 and 3 in the first round is stable: If group 3 switch their vote to R under the runoff rule, they cannot secure a victory for candidate R, and all they achieve is that there will be a runoff election between L and R, which is won by L. Consequently, since the deviation coalition is successful and stable, L is not a robust equilibrium outcome under the runoff rule.

Next, consider the scenario in which no equilibrium exists under plurality rule. This happens, if the Condorcet winner A has the largest support, but does not have a strategic majority. The reason is that a deviation coalition to the third candidate is successful (as the CW does not have a strategic majority) and internally stable (since a (A, A, B, B, C, C) split would produce A as outcome). With a runoff rule, this problem does not arise in this parameter constellation. In fact, if the Condorcet winner has the largest support, then his supporters can easily secure alone that the CW reaches the runoff election, at which point he will beat any remaining competitor.

In cases where the CW has the smallest support, such that his supporters cannot unilaterally secure him a place in the runoff election, another argument applies. Note first that no candidate other than the CW can secure an absolute majority in the first round of the election. Second, one of the other candidates, say C, cannot win even in a runoff against the other non-CW. Therefore, C supporters who have the CW as their second choice know that if they don't vote for the CW (and the CW consequently drops out), then their least preferred candidate will win. Hence, their support for the CW is very stable in the first round election.

Similar to the plurality rule, the runoff rule produces a unique winner, if voters' preferences give rise to a Condorcet cycle. Under both rules, the candidate with the largest core support cannot be the winner if there is a Condorcet cycle. However, under plurality rule, the winner is that candidate, who beats the candidate with the largest support in a binary election, while under the runoff rule, the winner is the candidate who has the smallest core support. This implies that plurality rule and runoff rule sometimes produce different (and respectively unique) equilibrium winners.

Proposition 7. *Suppose that there are three candidates and that no Condorcet winner exists; let A be the candidate with the largest support, and assume that in a binary election, A beats B, B beats C and C beats A ($S(A, B) > S(B, A)$, $S(B, C) > S(C, B)$ and $S(C, A) > S(A, C)$).*

The candidate with the smallest core support is the unique equilibrium outcome: If $N_C < N_B$, then C, if $N_C > N_B$, then B is the unique equilibrium outcome.

Proof. Suppose first that $N_C < N_B$, and consider the voting profile (A, A, A, C, C, C) in the first election. This voting profile could be upset by a coalition of A1, B1 and B2 switching to B. However, this coalition is not stable, as group A1 would have an incentive to switch back to A which would result in a runoff between A and B and, hence, a victory for A. Therefore, (A, A, A, C, C, C) is a robust political equilibrium.

Is there another equilibrium outcome? Suppose that there is an equilibrium in which B wins. B could either win already in the first round, or in a runoff against C.

In both cases, a deviation by A1, A2 and C1 to vote for A is successful and stable.[18] (If C1 change back their votes, only a runoff between A and B can result.) A similar argument shows that there cannot be an equilibrium in which A wins: A deviation of B2, C1 and C2 to vote for C is successful and internally stable.

Suppose now that $N_C > N_B$. In this case, (B, C, B, B, C, C) in the first round is robust, because a deviation of A1, A2 and C1 to vote for A is internally instable (given that A1 and A2 vote for A, C1 would like to switch back to C, which yields a runoff between A and C with winner C).

Is there another equilibrium? (A, A, A, C, C, C) in the first election is not robust, as the coalition of A1, B1 and B2 changing to B is now stable: If A switches back to vote for A, a runoff between A and C with winner C follows. (A, A, B, B, A, B) is not robust either, because a deviation of C1 and C2 to vote for C will result in a runoff of A against C with winner C. □

What is the intutiton for the difference between the runoff rule and the plurality rule, if no Condorcet winner exists?

Under plurality rule, those who prefer the candidate with the largest core support have a strong incentive to vote for that candidate. If the supporters of the other two candidates are divided, the strongest supporter group can effect a victory for their preferred candidate by voting for him. Ironically, this strength means that no one can trust coalition deviations that include supporters of the candidate with the largest core support. Consequently, the obvious coordination point is the candidate who is preferred by the majority to the candidate with the largest core support.

Under runoff rule, the argument is slightly different. The largest supporter group cannot unilaterally win the election, if the two other groups are divided, by just standing together in the first round, as they do not have an absolute majority. Hence, it is possible that the candidate with the largest core support actually gets no votes in equilibrium.

Why will the candidate with the smallest core support win? The decisive feature here is the incentive for the "helper group", i.e. those voters that do not rank the winning candidate highest, but nevertheless vote for him (as they like the opponent who gets votes in equilibrium even less).[19] The helper group for the candidate with the smallest core support has a strong incentive to vote for this candidate; if they don't, then the smallest candidate drops out and the resulting runoff election is won by the candidate whom the helper group likes least.

7 Plurality rule with more than 3 candidates

In this section, we will analyze how our results for plurality rule can be extended to a setting with more than 3 candidates. While we cannot provide a complete characteri-

[18] If one or two of the groups A1, A2 and C1 already vote for A in the original voting profile, a deviation of the rest to vote for A is meant.

[19] For example, if $N_B < N_C$ in the theorem, then B and C get votes in equilibrium, and the helper group is group A1 with preference $A \succ B \succ C$.

zation of the equilibrium for all circumstances, some of our results from the 3 candidate case generalize quite nicely.

It turns out that the following concept of a *projection of the preference profile with respect to candidate set T* is very useful. By this, we mean the polity that would arise, if none of the candidates outside of set T were available for election. For example, if there are in total 4 candidates (A,B,C and D), and we consider the candidate subset $T = \{A, B, D\}$, then there are several preference groups that have the same preferences over this subset; for example, the groups with preference rankings (over the entire set) $C \succ A \succ B \succ D$, $A \succ C \succ B \succ D$, $A \succ B \succ C \succ D$ and $A \succ B \succ D \succ C$ all agree that A is the best candidate from the set T. We can then define these four groups to form one preference type with respect to the projection of the preferences to T, and similarly, we can define the core support of a candidate with respect to T.

Definition 5 (Projection of preferences). *Let $N_x(T)$ be the core support of candidate x, for $x \in T$, if preferences are projected to T:*

$$N_x(T) = \#\{ \text{ voters who rank candidate } x \text{ highest among all } y \in T.\}$$

Similarly, let $N_{xy}(T)$ be the number of people who rank x highest and y second highest among the candidates in T.

The following definition 6 generalizes the concept of a strategic majority.

Definition 6 (Strategic majority). ***Candidate x has a strategic majority against candidate y in the projection of preferences to T****, if and only if*

$$N_x(T) > S(y, x),\ x, y \in T.$$

If $N_x > \max_{y \in T, y \neq x} S(y, x)$ (i.e., candidate x has a strategic majority against all other candidates in T), we say that ***candidate x has a strategic majority in T.***

The relevant projections turn out to be those to candidate sets that have exactly 3 elements, so some notation is defined for this special case in the next definition.

Definition 7. *Let $\mathcal{C}$ be the set of all candidates, and let $\mathcal{T}$ be the set of those subsets of $\mathcal{C}$ that have exactly 3 elements. Moreover, let $\mathcal{T}_x$ be the set of those elements of $\mathcal{T}$ that have x as one of their 3 elements.*

We can now state the extension of Proposition 2 to the case of more than 3 candidates. It is rather easy to see that, if candidate A has a strategic majority with respect to the whole candidate set $\mathcal{C}$ (i.e. a core support that is larger than the number of sympathizers of any candidate against A, then A must be the unique robust political equilibrium outcome, by the same arguments as those presented to prove Proposition 2. However, the condition of having a strategic majority against *all* other candidates is certainly very much harder to fulfill, as the number of candidates increases. Therefore, it is important that the condition in the following Proposition 8 is considerably weaker.[20] It requires only that candidate A has a strategic majority in any 3 candidate projection of the preferences, in order to guarantee that A is the unique robust equilibrium outcome.

Proposition 8. *Suppose that, for all $T \in \mathcal{T}_A$, candidate A has a strategic majority in T_A. Then, candidate A is the Condorcet winner and the unique robust political equilibrium outcome.*

Proof. We prove first that A is the CW. Assume to the contrary that there is some candidate x that wins against A; but then, A cannot have a strategic majority in $\{A, x, y\}$ (for arbitrary y), the required contradiction.

To see that A is a robust political equilibrium, consider the strategy profile where only A and some other candidate y receive votes (i.e., everyone who prefers A over y votes for A and vice versa). Since A has a strategic majority in $\{A, x, y\}$ for all x, there is no other candidate to whom a successful coalition deviation could be made.

For uniqueness, suppose to the contrary that there is some other robust equilibrium outcome, say x, who wins against another candidate $y \neq A$ (the other candidate cannot be A, as A would win against x). Then, a deviation by those groups who prefer A to x and y is sufficient to change the result of the election; hence, in order for x to be a robust equilibrium outcome, it is necessary that this deviation is internally instable. This would be the case, if there exists z such that a subset of the original coalition would deviate to z and z would then win the election.

First, note that z cannot be equal to y, as A has by assumption a strategic majority in $\{A, x, y\}$. Second, is it possible that z is some other candidate? If the second deviation occurs, those types who have a preference ranking with $A \succ x$ and $A \succ z$ still vote for A, and those types who have a preference ranking with $z \succ A \succ x$ vote for z; the first group is A's core support in $\{A, x, z\}$, and the latter group not larger than $S(z, A; \{A, x, z\})$. Since A has a strategic majority in $\{A, x, z\}$, the second deviation to z cannot be successful, and hence the deviation to A is internally stable. Consequently, $x \neq A$ cannot be the outcome of a robust political equilibrium. □

The following Table 7 gives an example of a case in which A does not have a strategic majority with respect to the whole set of candidates (because his core support is smaller than the support of any other candidate against A), but where A nevertheless has a strategic majority in all subsets containing A and 2 further candidates.

A	B	B	B	B	B	B	C	C	C	C	C	C	D	D	D	D	D	D
(any	A	A	C	C	D	D	A	A	B	B	D	D	A	A	B	B	C	C
type)	C	D	A	D	A	C	B	D	A	D	A	B	B	C	A	C	A	B
	D	C	D	A	C	A	D	B	D	A	B	A	C	B	C	A	B	A
44	3	3	4	4	4	4	3	3	4	4	4	4	3	3	4	4	4	4

Table 7: A has a strategic majority in all projections to subsets of 3 candidates

[20] See Table 7 for an example in which A does not have a strategic majority but nevertheless satisfies the condition of Proposition 8.

Consider for example B's sympathizers against A. Its size is B's core support (22), plus 6 types of 4 voters each who have either C or D as their first choice, but rank B over A), hence 46 votes in total and thus greater than A's core support. However, in any projection of the preferences to a set containing A and 2 other candidates, A has a strategic majority. Consider for example the set $\{A, B, C\}$: Together with types $D \succ A \succ B \succ C$ and $D \succ A \succ C \succ B$, A's core support in $\{A, B, C\}$ is now 50 votes, while B's sympathizers against A are still 46.[21] Similarly, A's core support in the projection is now bigger than $S(C, A)$. The sets $\{A, B, D\}$ and $\{A, C, D\}$ are, of course, completely symmetric.

In order to guarantee that an equilibrium (rather than the unique equilibrium) in which A is the outcome exists, we can weaken the condition even more. (However, it is unclear whether the uniqueness part would still hold.)

Proposition 9. *Suppose that there exists a $y \in \mathcal{C}$ such that A has a strategic majority for all subsets of T_A of the form $\{A, y, x\}$, for all $x \in \mathcal{C}\backslash\{A, y\}$. Then, candidate A is the Condorcet winner and a robust political equilibrium outcome.*

Proof. We prove first that A is the CW. Assume to the contrary that there is some candidate x that wins against A; but then, A cannot have a strategic majority in $\{A, x, y\}$ (for arbitrary y), the required contradiction.

To see that A is a robust political equilibrium, consider the strategy profile where only A and y receive votes (i.e., everyone who prefers A over y votes for A and vice versa). Since A has a strategic majority in $\{A, x, y\}$ for all x, there is no other candidate to whom a successful coalition deviation could be made. □

Intuitively, this condition is quite weak (given that candidate A is the Condorcet winner), if there are many candidates. For example, if there is one candidate who is very unpopular among most people, then we can choose this candidate to be y, and A wins with near unanimity; this equilibrium cannot be upset by a coalition, if A is the Condorcet winner.

A similar conclusion holds if there is another candidate y whose supporters *hate* candidate A (i.e, all people who rank y higher than A actually have A at the bottom of their preference ranking). To see this, note that $N_A(\{A, x, y\}) = S(A, x, \{A, x, y\}) - \#\{\text{people with preference ranking } y \succ A \succ x\}$. If the second term on the right hand side is zero for all x, then it is sufficient that A is the Condorcet winner in order to also have a strategic majority in any 3-projection.

What happens, if no candidate satisfies the condition of Proposition 9? We did not succeed in obtainaing a result describing cases in which a robust equilibrium exists, in case preferences give rise to a Condorcet cycle. While a robust equilibrium always exists with 3 candidates and a Condorcet cycle, we have an example of a Condorcet cycle with 4 candidates that does not have a robust equilibrium. More work is needed, but (from preliminary results) it seems unlikely that any conditions guaranteeing existence can

[21] The size of $S(x, y)$ is never affected by a projection (as long as x and y are in the candidate subset considered, of course).

be expressed in terms of projections to 3 candidate sets, if no candidate has a strategic majority.

8 Runoff rule with more than 3 candidates

In this section, we will analyze how our results can be extended to a setting with more than 3 candidates. As in the case of plurality rule, we cannot provide a complete characterization of the equilibrium for all circumstances, but we can provide results for some important special cases.

As in the analysis of the plurality rule with more than 3 candidates, if there is no CW, it becomes very difficult to find simple conditions for equilibrium existence. We therefore concentrate on the case that a CW exists and look for additional conditions that guarantee that the CW is an equilibrium outcome, and also for (more restrictive) conditions that guarantee that the CW is the unique outcome.

Proposition 10. *Assume that candidate A is the Condorcet winner.*

1. *If, in all 3-candidate projections, A's support is at least* $1/3$ *of the electorate, then A is the outcome of a robust political equilibrium.*
2. *If, in all 3-candidate projections, A has a strategic majority, then A is the unique outcome of a robust political equilibrium.*
3. *If candidates have positions on the unit interval and voters have Euclidean preferences and A is the preferred candidate of the median voter, then A is the unique outcome of a robust political equilibrium.*

Proof.

1. Consider the profile in which those who support A versus B vote for A and vice versa; since A is the CW, A wins in this profile. Now consider a coalition deviation in the first round to candidate x. The minimum number of people who still vote for A is the number of people who prefer A to B and A to x, hence $N_A(\{A, B, x\})$; if this is larger than 1/3 of the electorate, then it is guaranteed that A will get into the runoff (which will be won by A), unless x achieves an absolute majority in the first round. However, the maximum number of votes for x is $N_x(\{A, B, x\}) \leq S(x, A) < S(A, x)$, so x cannot get an absolute majority in the first round and hence there is no successful deviation coalition.
2. Remember that, if A has a strategic majority, he has necessarily a core support that is larger than 1/3. Hence, the claim that A is the outcome of a robust equilibrium follows from 1. It remains to be shown that this outcome is unique. Assume to the contrary that there is a robust equilibrium in which B wins against $x \neq A$. However, if all those voters who prefer A over B change to vote for A in the first round, then A will win.

   Is it possible that this deviation is not internally stable? Suppose that there is another candidate, y, to whom a subset of the deviation coalition would like to deviate. In this case, A would still get the vote of his core support in $\{A, B, y\}$.

Candidate y gets at most $S(y, A; \{A, B, y\}) < N_A(\{A, B, y\})$ votes, so Candidate A would in any case reach the runoff round and there beat any opponent. Consequently, the deviation to y cannot be successful.

3. Consider a setting in which voters and candidates have positions on the unit interval and Euclidian preferences. Let A be the candidate ranked highest by the median voter, let B be the candidate who is ranked second highest by the median voter, and assume for concreteness that B's position is to the right of A.[22]

   Consider the voting profile in which all who prefer A over B vote for A, and the rest vote for B. Note that any possible coalition can contain either people who are to the left of A, or those who are to the right of A, but no coalition can contain both voters whose ideal position is to the left of A *and* voters whose ideal position is to the right of A (because these voters cannot agree on someone who is better for them than A). Furthermore, it is quite obvious that there cannot be a successful right deviation coalition: The right could possibly replace B by some more right candidate in the runoff, but that candidate also loses against A.

   How about a left deviation such that in the end, candidate C wins? This requires that neither A nor B reach the runoff stage, because C would lose against either of them (by our assumption that the median voter ranks A over B over any other candidate). Hence, there has to be some other candidate D against whom C can win in a runoff election, and those who prefer C over A must be able to secure that both C and D get more votes than A and B, respectively. Formally, let $G(\cdot)$ be the cumulative distribution function of voter types, and let x_i denote the position of candidate i.

   There are $G\left(\frac{x_A+x_C}{2}\right)$ voters who prefer C to A; if they deviate jointly to vote for C and D, the remaining voters for A are $G\left(\frac{x_A+x_B}{2}\right) - G\left(\frac{x_A+x_C}{2}\right)$, and there are $1 - G\left(\frac{x_A+x_B}{2}\right)$ voters who still vote for B. For the deviation coalition to be able to get both C and D in the second round, we must have that

$$G\left(\frac{x_A + x_C}{2}\right) > 2\max\left(G\left(\frac{x_A + x_B}{2}\right) - G\left(\frac{x_A + x_C}{2}\right), 1 - G\left(\frac{x_A + x_B}{2}\right)\right)$$

   or, equivalently,

$$G\left(\frac{x_A + x_C}{2}\right) > \max\left(\frac{2}{3}G\left(\frac{x_A + x_B}{2}\right), 2\left(1 - G\left(\frac{x_A + x_B}{2}\right)\right)\right).$$

   The right hand side of this expression is minimized for $G\left(\frac{x_A+x_B}{2}\right) = 1/2$, so $G\left(\frac{x_A+x_C}{2}\right) \geq 1/2$. This, however, contradicts the assumption that A is the CW, as A winning against C implies $G\left(\frac{x_A+x_C}{2}\right) < 1 - G\left(\frac{x_A+x_C}{2}\right)$.

   For uniqueness, suppose to the contrary that there is some robust equilibrium, in which candidate D≠A wins. Without loss of generality, assume that D is to

[22] Clearly, this is without loss of generality. If B is to the left of the median, all arguments given in the following apply with the appropriate changes.

the right of A. Since D is not the CW, there exists a nonempty set $W(D)$ that contains the candidates which would win in a binary election against D.

Let E be the candidate who is preferred by the "25th percentile voter"[23] among all candidates in $W(D)$: $||x_E - x_{25}|| < ||x_i - x_{25}||$ for all $i \in W(D)$.

Now consider a deviation by all those voters who are to the left of the median, plus an additional voter, to vote for E. This deviation is beneficial for all voters in the deviation coalition, successful and internally stable: Since x_E is the median candidate within the deviation coalition, there cannot be a majority of the deviation coalition which can agree on replacing E with some other candidate. □

While Proposition 10 gives sufficient conditions for the Condorcet winner to be the outcome of a robust equilibrium, it is unclear what the weakest sufficient conditions would be. However, we will now show that being the CW is alone not sufficient for being a robust equilibrium outcome.

Such a conclusion would seem at least somewhat intuitive, as the CW should be more stable under a runoff rule than under a plurality rule: Consider first the case of plurality rule, and suppose that the sympathizers of candidate x would like to replace candidate A, the CW, as winner of the election. All they have to do is to get more votes than A, but it is not necessary to get an absolute majority in favor of candidate x. In fact, since at most those people who prefer x to A vote for x in the coalition deviation, and this set of people is smaller than half of the electorate, it is impossible to achieve an absolute majority for candidate x in a deviation against an equilibrium in which A wins. But, as said, this is also not necessary, a relative majority for the deviation candidate x is sufficient under plurality rule to upset a voting profile as a robust equilibrium.

Under a runoff rule, upsetting a voting profile with the CW as winner appears much harder. As above, the deviators cannot hope for a victory by absolute majority in the first round, since less than half of the electorate prefers x over A. The only way how x could end up as the winner is that x gets into the runoff and A does not. Hence, for a successful coalition deviation the x supporter do not only have to make sure that, in the first round, x gets more votes than A, but also that there is a second candidate y who also gets more votes than A and who can be beaten by x in the second round. This appears much more difficult than the task under plurality rule, and so it is perhaps intuitive to expect that the CW is always a robust equilibrium outcome.

However, this is not true. Consider the preferences given in Table 8. In binary elections, D wins against any competitor with 50 to 49 votes (so D is the CW), and A beats B, B beats C and C beats A, with 66 to 33 votes, respectively;[24] hence, below the CW, we have a Condorcet cycle. There are 99 voters (so 50 is the absolute majority), and, if no candidate achieves a majority in the first round, a candidate needs at least 33 votes in order to proceed to the second round.

[23] The 25th percentile voter located at x_{25} is the voter who is more right than 25 percent of the population, and more left than 75 percent of the population; formally, $G(x_{25}) = 1/4$.

[24] The presentation is simplified by the fact that A, B and C have the same size of support, but nothing in the following really depends on this.

$\succ_1$	$\succ_2$	$\succ_3$	$\succ_4$	$\succ_5$	$\succ_6$
A	A	B	B	C	C
D	B	D	C	D	A
B	C	C	A	A	B
C	D	A	D	B	D
25	8	25	8	25	8

Table 8: D is CW, but not a robust equilibrium

Is there a robust equilibrium in which D wins in the first round against, say, A? Consider the voting profile (A, A, D, A, D, A); groups 4, 5 and 6 could jointly deviate to vote for C; then, there will be a runoff between A and C, and C is the overall winner. This deviation is also internally stable, as group 4 cannot achieve a victory of B.

Similarly, other voting profiles in which just two of the big groups (those with 25 members) vote for D, cannot be robust equilibria either. Note that the trembling hand part of our robustness definition implies that the 3 small groups do not vote for D (as this is a weakly dominated strategy for them).

Hence, the only remaining possibility how D could win is that D gets, in the first round, the votes of all 3 large groups. However, this profile again violates the trembling part of our robustness definition: Consider the profile (D, A, D, B, D, C); if group 1 changes to A, the probability that this will render A the final winner is a term of order ε^2: One vote for D is not counted in the first round, and there is another D vote not counted in the runoff round. On the other hand, the increased risk that either B or C become the final winners because group 1 switched is a term of order ε^{52}.[25] Hence, the benefits of the switch outweigh its risks.

Also for other combinations of the votes of the 3 small groups, there is at least one of the large groups that benefits from switching their vote to their most preferred candidate, and so a profile in which all large groups vote for D is not robust, either.

It is interesting to consider the following variant on this example. Suppose that the preference distribution is slightly modified such as to generate small differences between the sizes of the core supports (see Figure 9). Moreover, suppose that there are only 3

$\succ_1$	$\succ_2$	$\succ_3$	$\succ_4$	$\succ_5$	$\succ_6$
A	A	B	B	C	C
D	B	D	C	D	A
B	C	C	A	A	B
C	D	A	D	B	D
25	7	26	7	27	7

Table 9: D is CW, but A is the unique robust equilibrium outcome

[25] For B (or C) to reach the runoff after group 1's switch, at least 25 A votes have to be not counted, as well as 26 D votes (otherwise, D reaches an absolute majority of the votes already in the first round); in addition, there must be an additional mistake in the runoff round.

leaders, who make recommendations, one for the A supporters, one for the B supporters and the third one for the C supporters; one interpretation is that these leaders are the candidates themselves. Assume furthermore that the leader has the same preference as the respective big group, and that people only follow their respective leader, if that recommendation is not weakly dominated for them (there is no restriction imposed on the behavior of voters who don't follow leaders because their recommendation is a weakly dominated strategy for these voters). For example, if the A leader proposes to vote for D, then group 1 votes for D, and the voters in group 2 vote for any of the 3 other candidates (they are not required by this modified equilibrium concept to vote for the same candidate).

In this scenario, consider the following recommendations: A and C leader recommend to vote for candidate A, and the B leader recommends to vote for B. The B and C leader have two possible deviation that would benefit them both. First, they could recommend to their constituents to vote for D. However, if the B leader recommends to vote for D, then the C leader is better of recommending to vote for C: The first round result is then (at least) 32 votes for A, (at least) 34 votes for C, and 25 votes for D.[26] The runoff A versus C is won by C, and hence the C leader will not keep to the initial coalition agreement.

Second, the B and C leader could agree to recommend to vote for C. Then, however, the B leader would have an incentive to switch his recommendation to B, resulting in a B vs. C runoff and a B victory. Hence, neither of the successful deviations is internally robust, and the equilibrium with A as the winner is robust. Arguments similar to those presented above show that there is no other robust equilibrium outcome. Hence, even though a CW exists, the unique robust equilibrium outcome is another candidate in this example.

9 Conclusion

This paper analyzed the effects of two intuitive refinements for voting games. The first one is based on the idea that, for various reasons, not all votes cast are actually counted in elections. This is shown to imply Duverger's law in plurality elections. The second part of our refinement relies on the idea that voters with identical or similar interests should be able to coordinate their voting behavior, if this is beneficial for all involved parties and credible.

The refinement produces some strong and surprising results. For the case of three candidates, we can completely characterize the outcome under both runoff rule and plurality rule. Under runoff rule, there is always a unique equilibrium. Under plurality rule, a unique equilibrium obtains, if one of the candidates has a "strategic majority" or if there is a Condorcet cycle (the latter result is quite surprising).

As a general result, runoff rule appears to be better behaved than plurality rule. A robust equilibrium exists more often under runoff rule than under plurality rule, and,

[26] The exact behavior of the voters in group 4 does not matter, since they will not vote for D and therefore do not affect the identity of the candidates in the runoff.

at least in the case of three candidates, runoff rule delivers a unique equilibrium in some cases in which there are two equilibrium outcomes under plurality rule.

There are several interesting questions for future research. First, one can study the question of candidate entry in a citizen candidate model (Osborne and Slivinski (1996), Besley and Coate (1997)). For this question, it is a very desirable property of our model that it often produces a unique outcome (such that the result of the voting subgame is clear when candidates make their decision whether to run).

Second, one can study location games in which (other than in citizen candidate models) potential candidates can choose on which platform to run. It is quite clear that, under the runoff rule, the candidates will select the Condorcet winner (provided that one exists), but this is less clear under plurality rule and so convergence of candidates to that position is not guaranteed.

A third extension would be to study games in which only some voter groups can coordinate their behavior. For example, what happens if not every preference group has a leader to give a recommendation to group members, but only the candidates themselves can decide whether to withdraw and recommend to vote for some of their former opponents. Is the ability to coordinate with other voters with similar interests necessarily beneficial? The answer to this question may provide further insights into the theory of party formation.

References

[1] Aumann, R. (1959): "Acceptable Points in General Cooperative n-Person Games", in: *Contributions to the Theory of Games IV*, Princeton University Press, Princeton N.J. 1959.

[2] Bernheim, B. D. , Peleg, B. and Whinston, M. (1987): "Coalition-Proof Nash Equilibria I. Concepts", *Journal of Economic Theory*, 42, 1–12.

[3] Besley, T. and Coate, S. (1997): "An Economic Model of Representative Democracy", *Quarterly Journal of Economics*, 112, 85–114.

[4] Black, D. (1948): "The Decision of a Committee Using a Special Majority", *Econometrica*, 16, 245–261.

[5] Condorcet, M.d. (1976): "Essay on the Application of Mathematics to the Theory of Decision Making", *Condorcet, Selected Writings*, ed. by K. Baker, Indianapolis: Bobbs-Merrill.

[6] De Sinopoli, F. and A. Turrini (1999): "A Remark on the Besley-Coate Model of Representative Democracy", CORE DP 9927, Universite Catholique de Louvain, Belgium.

[7] De Sinopoli, F. (2000): "Sophisticated Voting and equilibrium refinements under plurality rule", *Social Choice and Welfare* 17, 655-672.

[8] Dhillon, A. and Lockwood, B. (2002): "When are Plurality Rule Voting Games Dominance Solvable?", *Games and Economic Behavior*, forthcoming.

[9] Dhillon, A. and Lockwood, B. (2001): "Multiple Equilibria in the Citizen-Candidate Model of Representative Democracy: A Note", mimeo.

[10] Farquharson, R. (1969): "Theory of Voting", New Haven. Yale University Press.

[11] Feddersen, T. (1993): "Coalition-proof Nash equilibria in a Model of Costly Voting Under Plurality Rule", mimeo.

[12] McKelvey R. D. and Niemi, R. G. (1978): "A Multistage Game Representation of Sophisticated Voting for Binary Procedures", *Journal of Economic Theory*, 18, 1–22.

[13] Moulin, H. (1979): "Dominance Solvable Voting Schemes", *Econometrica*, 47(6), 1337–1351.

[14] Niemi, R. G. (1983) "An Exegesis of Farquharson's Theory of Voting", *Public Choice*, 54, 187–190.

[15] Osborne, M. and Slivinski, A. (1996): "A Model of Political Competition with Citizen Candidates", *Quarterly Journal of Economics*, 111, 65-96.

[16] Palfrey, T. (1989): "A Mathematical Proof of Duverger's Law", in P. Ordeshook: "Models of strategic choice in politics", Ann Arbor. University of Michigan Press.

[17] van Damme E. (1983): "Refinements of the Nash Equilibrium Concept", Berlin: Springer Verlag.

Appendix

Proof of Lemma 2

Suppose that we have an equilibrium with $V_1 \geq V_2 > V_3 > 0$; in a slight abuse of notation, call the candidate who receives the ith-most voting intentions "Candidate i". In this equilibrium, the probability that Candidate 1 will win is $1 - o(V_1 - V_2)$ (where $o(x)$ is a term of order x in ϵ),[27] the probability that Candidate 2 will win is $o(V_1 - V_2)$, and the probability that Candidate 3 ends up as the winner is $o((V_1 - V_3) + (V_2 - V_3))$.

Consider citizen j who intends to vote for the third placed candidate, and fix all other citizens' vote intentions. The probability that j's vote is pivotal for a decision between Candidate 3 and Candidate 1, or between Candidate 3 and Candidate 2, is $o((V_1 - V_3) + (V_2 - V_3))$ in ϵ. On the other hand, the probability that j's vote is pivotal in a decision between Candidate 1 and Candidate 2 is $o(V_1 - V_2)$ in ϵ. In all other cases, voter j's vote does not matter. Since $V_1 - V_2 < (V_1 - V_3) + (V_2 - V_3)$, there exists an $\bar{\epsilon} > 0$ such that the action which maximizes expected utility for all $\epsilon \in (0, \bar{\epsilon}]$ is to vote for candidate 1, if $C_1 \succ_j C_2$, and to vote for candidate 2 if $C_2 \succ_j C_1$. Consequently, a vote for the third placed candidate cannot be optimal.

9.1 Proof of Proposition 1

First, it is easy to see that the profile in which just candidates i and j receive votes, and every voter who prefers i over j votes for i and vice versa, is an equilibrium. Suppose that $S(i,j) > S(j,i)$. The probability that an individual's vote is pivotal between i and j is a term of order $S(i,j) - S(j,i)$, while the probability that his vote, if cast for a third candidate, is pivotal between any other pair of candidates is a term of order $N - 1 > S(i,j) - S(j,i)$ (where N is the the total number of voters). Hence, there is no point in deviating to vote for a third candidate. Similarly, it is also obvious that it cannot be optimal for a voter who prefers i to j to vote for j, and vice versa.

We now turn to the proof that there are no other equilibria. First, note that there cannot be an equilibrium in which only one candidate receives votes. This follows immediately from Lemma 1. Second, we will show that there cannot be an equilibrium in which three (or more) candidates receive a positive number of votes. For simplicity, we give the argument for the case of three candidates, but it should be obvious that the argument applies for an arbitrary number of candidates.

Lemma 1 already shows that we cannot have $V_1 \geq V_2 > V_3 > 0$. The remaining possible configuration in which 3 candidates receive votes is $V_1 \geq V_2 = V_3$.

[27] The reason is that at least $V_1 - V_2$ votes need to be not counted in order that there is a possibility that Candidate 1 will not receive the most of the counted votes.

First, can we have $V_1 = V_2 = V_3$? In such an equilibrium, every core supporter of candidate i must vote for i (otherwise, a voter who preferred i, but voted for $j \neq i$ could switch to vote for i and secure i's victory). However, we assumed that there is one candidate who has the largest core support, so $V_1 = V_2 = V_3$ cannot be true.

Second, can we have $V_1 > V_2 = V_3$? Again, it must be true that the respective core supporters vote for their favorite candidates: This is obvious for the core supporters of the winner, but it is also true for the core supporters of one of the runners-up: Suppose it were not so, but rather a core supporter of (say) Candidate 2 voted for Candidate 3; in this case, switching the vote to Candidate 2 increases the probability that 2 wins by one order of magnitude (in ε), and in comparison to this effect, the change in 3's winning probability is negligible.

However, consider the voting behavior of an individual who prefers candidate 2 to 3 to 1 and who votes for 2 in equilibrium. Switching the vote to 3 increases the probability that 3 is elected from a term of order $V_1 - V_3$ to a term of order $V_1 - V_3 - 1$ in ε and decreases the probability that 2 will be elected to a term of order $V_1 - V_3 + 1$, but the overall effect of the vote switch on this individual's expected utility must be positive for ε sufficiently small. Hence, there cannot be an equilibrium with $V_1 > V_2 = V_3$.

Proof of Proposition 5

Existence of a CW is straightforward and was shown by Dhillon and Lockwood (2002): The candidate who is the worst candidate for more than 2/3 of the electorate must be the Condorcet loser, and therefore no Condorcet cycle can exist.

Without loss of generality, assume that candidate A is the CW and candidate C is the Condorcet loser. Note that C cannot have the largest support, because the largest support is greater than 1/3 of the electorate and then the DL condition could not be satisfied. From Proposition 4, we can conclude that there cannot be multiple equilibria, because for multiple equilibria to exist, the Condorcet loser has to have the largest support.

Could it be the case that no robust equilibrium exists? From Proposition 4, number 3, and the Dhillon Lockwood condition, the following inequalities would have to be satisfied:[28]

$$
\begin{array}{rcl}
w_{AB}+w_{AC} & > & w_{BA}+w_{BC} \\
w_{AB}+w_{AC} & > & w_{CA}+w_{CB} \\
w_{AB}+w_{AC} & < & w_{BA}+w_{BC}+w_{CB} \\
w_{AB}+w_{AC} & < & w_{BC}+w_{CA}+w_{CB} \\
w_{AB}+w_{AC}+w_{CA} & < & w_{BA}+w_{BC}+w_{CB} \\
w_{AB}+w_{AC}+w_{BA} & > & w_{BC}+w_{CA}+w_{CB} \\
w_{AB}+w_{BA}+w_{BC} & > & w_{AC}+w_{CA}+w_{CB} \\
w_{AB}+w_{BA} & > & 2(w_{AC}+w_{BC}+w_{CA}+w_{CB})
\end{array} \tag{3}
$$

We claim that there does not exist a solution to this system of inequalities. Assume to the contrary. The following Lemma shows that, if a solution to (3) exists at all, then there must also be a solution to (3) that has $w_{AC}=0$.

Lemma 3. *Suppose that* $(w^*_{AB}, w^*_{AC}, \ldots, w^*_{CB})$ *is a solution to the inequality system (3). Then,* $(w^*_{AB}+w^*_{AC}, 0, w^*_{BA}, \ldots, w^*_{CB})$ *is also a solution of (3)*

Proof. Clearly, shifting weight from w_{AC} to w_{AB} will not affect at all whether the first six inequalities are satisfied, and makes it easier to satisfy the last two inequalities. □

Hence, a necessary condition for a solution to (3) to exist is that there exists a solution to the following system (where we replaced w_{AC} by 0):

$$
\begin{array}{rcl}
w_{AB} & > & w_{BA}+w_{BC} \\
w_{AB} & > & w_{CA}+w_{CB} \\
w_{AB} & < & w_{BA}+w_{BC}+w_{CB} \\
w_{AB} & < & w_{BC}+w_{CA}+w_{CB} \\
w_{AB}+w_{CA} & < & w_{BA}+w_{BC}+w_{CB} \\
w_{AB}+w_{BA} & > & w_{BC}+w_{CA}+w_{CB} \\
w_{AB}+w_{BA}+w_{BC} & > & w_{CA}+w_{CB} \\
w_{AB}+w_{BA} & > & 2(w_{BC}+w_{CA}+w_{CB})
\end{array} \tag{4}
$$

The fourth and the last inequality together imply $w_{AB} < w_{BC}+w_{CA}+w_{CB} < w_{BA}$; but $w_{AB} < w_{BA}$ contradicts the first inequality.

[28] The first 2 inequalities follow since A must have the largest support. The next two inequalities describe that A has neither a strategic majority against B nor against C. The next 3 inequalities describe that A wins against B and C, and B wins against C in a binary election. The last inequality is the DL condition.