Runoff Elections in the Laboratory*

Laurent Bouton[†]

Jorge Gallego[‡]

Aniol Llorente-Saguer[§]

Rebecca Morton[¶]

June 16, 2021

Abstract

We study experimentally the properties of the majority runoff system and compare them to those of plurality rule. Our focus is on Duverger's famous prediction that the plurality rule leads to higher coordination of votes on a limited number of candidates than the majority runoff rule. We find strong coordination forces under both systems. However, as predicted by the theory, in some cases these forces are stronger under plurality. Despite these differences in voting behavior, we find small and mostly not significant differences in electoral outcomes and hence voters' welfare.

JEL Classification: C92, D70

Keywords: Multicandidate elections, Majority Runoff, Plurality, Laboratory Experiments

*Our dear friend and coauthor Rebecca Morton passed away in September 2020. We dedicate this article to her memory. We thank participants of the PEPS Chile 2017, Theem 2018, ESA 2018, APSA 2019, ELEW 2019, and seminar participants at Carlos III, CERGE-EI, Nanyang Technological University, National University of Singapore, Universidad del Rosario, University of Iowa, and Paris School of Economics. We extend particular thanks to Micael Castanheira, Boris Ginzburg, Gabriele Gratton, Jean-Francois Laslier, and Jan Zapal for insightful comments. We would also like to thank Mateo Vasquez and Taylor Mattia for their excellent assistance with running the experiments. This project has received funding from the European Research Council (Laurent Bouton) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 637662). The experiments reported in this paper have received the approval from the Institutional Review Board of New York University (IRB-FY2016-1401).

[†]Georgetown University, CEPR and NBER. Email: boutonllj@gmail.com.

[‡]Universidad del Rosario. Email: jorge.gallego@urosario.edu.co

[§]Queen Mary University of London, and CEPR. Email: a.llorente-saguer@qmul.ac.uk.

[¶]NYU-NYC and NYU-Abu Dhabi.

1 Introduction

With the spread of democracy worldwide since the end of World War II, there has been an increase in the use of majority requirements in which a candidate cannot secure election unless he or she has received a majority of the votes (see Bormann and Golder 2013, Figure 7). When no candidate receives a majority, then a second round is required between the two top candidates to determine which wins. This electoral rule is referred to as the *majority runoff rule* (hereafter runoff rule). In contrast, in a plurality rule or first-past-the-post election, whichever candidate receives the most votes is declared the winner. Majority runoff elections are not new: they appear to have existed as long as plurality rule elections in western countries.[1] They have traditionally been used for French presidential elections, and in many mayoral, legislative, and state-wide contests in the United States. More recently, they have been adopted during the democratic transition in Eastern Europe, Africa, and to a lesser extent Latin America. Runoff rules are also often used when committees or legislatures are selecting leaders and there are more than two candidates.[2]

Nevertheless, many elections continue to be conducted under plurality rule. The most obvious reason for not adopting majority requirements are the costs involved in a two-stage election process which increases (i) expenses paid by government agencies to conduct elections (doubling the variable expenses),[3] (ii) voting costs since voters must go to the polls twice,[4] and (iii) campaign costs due to the longer election process.[5,6] Clearly some have concluded that any advantages from runoff rules are not worth these costs.

But, what are the advantages? In this paper we focus on one in particular:[7] the

[1] Colomer (2007) records their use at various points in the 1800s by Belgium, France, Greece, Honduras, the Netherlands, Spain, and Switzerland.

[2] For example, in 2016 three members of the Louisiana House of Representatives vied to be House Speaker and a second round was held between the two top candidates when none received a majority in the first round. See Ballard (2016).

[3] In the 2013 NYC election for Public Advocate the 2nd round cost $13 million to conduct, although the office has an annual budget of only $2.3 million. (https://nyti.ms/2VDya1h).

[4] These costs are far from trivial: Stewart and Ansolabehere (2015, page 48) report that long lines at ballot places alone had "an estimated economic cost of $544.4 million [...] in 2012, which is about one-fifth the budget of local election offices in 2012 [...]"

[5] An extreme case of potential negative consequences is reported by Perez-Linan (2006, p. 136): "[...] additional financial needs created by the Colombian runoff campaign in 1994 led Ernesto Samper's operatives to accept contributions from the Cali drug cartel (Medina Serna, 1997)."

[6] Other drawbacks associated with the use of runoff rules include party fragmentation (Shugart and Taagepera 1994), the existence of potential harmful strategic behavior by voters in the first round called "push-over" (Smith 1973, Cox 1997, Saari 2003), and an increase in the risk that the median candidate is "squeezed" out of the second round (Van Der Straeten et al. 201, and Solow 2019).

[7] Some of the other advantages include: (i) Candidates elected through majority runoff elections are

perception that in runoff elections, voters in the first round can express their preferences more "sincerely", i.e., vote for their favorite candidate, than in a plurality rule contest. That is, in a plurality election, if a voter believes that her favorite candidate does not have enough support to defeat the other candidates, she will find it desirable to vote strategically for a less-preferred one. This reasoning leads to a concentration of votes on two candidates–the so-called *Duverger's Law* (Duverger 1959, Riker 1982, Palfrey 1989). In a runoff system, voters are arguably less concerned about whether their favorite candidate can defeat the rest of the field in the first stage, than whether the candidate can at least make it to the second round. Thus voters are believed to be more likely to vote sincerely for their favorite candidate in the first round–the so-called *Duverger's Hypothesis* (Duverger 1959, Riker 1982, Piketty 2000, Martinelli 2002).

However, recent theoretical literature has cast doubt on the extent that voters vote (more) sincerely in the first round of a runoff election. The forces toward coordination are in fact also present in those elections. They can be sufficiently strong as to prevent the existence of an equilibrium in which all voters vote sincerely, and even lead to a concentration of votes on only two candidates, as under plurality (see Bouton 2013, Bouton and Gratton 2015, and Bouton and Ogden 2021).

In this paper, we investigate the effects of the runoff rule on preference revelation using laboratory experiments. We find that, under the runoff rule, there are strong coordination forces that lead to insincere voting: a substantial fraction of the voters does not vote for their most preferred candidate to instead vote for their second choice. This is true even in our most extreme treatment in which voters' incentives to coordinate under the runoff rule are very weak. Yet, as predicted by theory, in some treatments, these forces appear stronger under plurality than runoff. This supports the theoretically-backed idea that runoff elections are more conducive to preference revelation than plurality elections only in some situations: when the minority candidate has sufficiently low support and when majority voters strongly disagree about which candidate is best.

We also explore the effects of the electoral system on electoral outcomes (i.e., which

assured of having received at least majority support and thus can contend that they have a mandate of democratic legitimacy (see the discussions in Perez-Linan 2006 and McClintock 2018); (ii) Candidates have incentives to cater to the interests of a broad coalition of voters (see, e.g., Lizzeri and Persico 2005); (iii) Voters can more easily transmit information about their policy preferences to candidates (Piketty 2000, Castanheira 2003, Blais 2004a,b).

alternative wins) and voters' welfare. There are two potential channels: outcomes can differ because (i) the system affects voters' behavior, and (ii) the two systems map votes into outcomes differently. However, we find small and insignificant differences between runoff and plurality systems. This lack of difference occurs because voters (partially) coordinate behind two alternatives under both systems. But, this result does not mean that they are equivalent. One difference our experiments highlight is that second rounds, which are costly in practice, take place frequently under the runoff system.

Our experimental approach is particularly well-suited for exploring the causal impact of the runoff system on the behavior of voters. As we discuss in detail below, we complement previous studies that cannot avoid confounding factors, such as changes in the behavior and the number of candidates/parties. This allows us to contribute to the literature along two dimensions. First, by showing that Duvergerian forces affect voters' behavior strongly under runoff, we complement the scarce empirical evidence on the topic. Second, by showing that Duvergerian forces affect voters more strongly under plurality than runoff only in some cases, we highlight the importance of voters to explain previously documented differences between the two systems (more on this in the next Section). Our experimental approach also allows us to assess the effect of the runoff system on electoral outcomes and voters' welfare. We find that differences in voting behavior do not necessarily lead to substantial differences along those two dimensions.

2 Background Literature

As mentioned in the introduction, the idea that runoff elections lead to more sincere voting and the viability of more electoral choices (either candidates or parties) is usually associated with Duverger (1959). Duverger argued that there are two sources for this difference. One source is a "mechanical effect" of parties not entering or forming coalitions of larger parties in anticipation of the difficulty of achieving enough votes to win. The other is a "psychological effect": in a plurality election, voters abandon their first choice if it is unlikely to contend for victory, and instead vote strategically for their favorite contender. According to Duverger, in the first round of majority runoff elections, both effects are less present, with more, smaller parties entering competition and voters not so attracted to abandon favorite choices when they are trailing behind. Our focus in this paper is on the psychological effect, and the extent to which it is less prevalent for the first round of voting in a majority runoff election than in a plurality election. Hence, we consider voter behavior holding candidate and party behavior fixed.[8]

The more recent, game-theoretic, literature refines Duverger's predictions and our understanding of voters' behavior under plurality and majority runoff.[9] In plurality elections, there are always multiple Duvergerian equilibria, in which only two candidates receive votes, and the sincere voting equilibrium does not exist generically.[10] By contrast, in the first round of majority runoff elections, the sincere voting equilibrium, or an equilibrium in which more than two candidates receive a positive share of the votes, exists for a broad range of situations. The conclusion of the early literature, i.e., that Duvergerian equilibria do not exist in majority runoff elections, critically depends on assumptions that are arguably unreasonable and counterfactual: there must be no uncertainty about the distribution of preferences in the electorate after the first round, and voters must perceive that there are no costs associated to holding a second round. Relaxing either of these assumptions leads to the existence of Duvergerian equilibria. This result holds not only in the pivotal voter model, but also in other models of voting.

Our paper contributes to this theoretical literature by extending some of the results to cases with a small number of voters. Although most of the literature on runoff rules has focused on large elections, runoff rules are often used in small group decision making as well when there are more than two choices such as when academic departments choose which of three or more candidates to hire for a single position, or when small governing councils and legislative bodies choose leaders.

There are many empirical studies exploring the presence of Duvergerian forces in plurality elections (Kawai and Watanabe 2013, Spenkuch 2015, 2018, and Pons and Tricaud

[8] See Osborne and Slivinski (1996), Callander (2005), Brusco et al. (2012), Buisseret (2017) and Shin (2019) for studies that examine the mechanical effect.

[9] See, e.g., Palfrey 1989, Myerson and Weber (1993), Cox (1997), Fey (1997), Piketty (2000), Myatt (2007), Martinelli (2002), Bouton (2013), Bouton and Gratton (2015), and Bouton and Ogden (2021).

[10] This statement requires two important clarifications. First, with aggregate uncertainty, (i) the sincere voting equilibrium exists for a broader set of distributions, and (ii) Duverger's law equilibria exist as long as the amount of aggregate uncertainty is not excessive (see Myatt 2007 and Bouton, Castanheira, and Llorente-Saguer 2017). Second, the literature also shows the existence of an equilibrium in mixed strategy in which more than two candidates receive a positive fraction of the votes (Myerson and Weber 1993, Fey 1997, Bouton, Castanheira and Llorente-Saguer 2017). The key condition of existence of that equilibrium is that both majority alternatives are equally strong contenders of the minority alternative. Without aggregate uncertainty, this equilibrium is unstable and often deemed unreasonable (see, e.g., Fey 1997).

2018 use observational data; Cain, 1978; Abramson et al., 1992; Alvarez and Nagler, 2000; Blais et al., 2001 use survey data; and Forsythe et al. 1993, 1996, Fisher and Myatt 2001, Bouton et al. 2016, 2017b use experimental data). Overwhelmingly, these studies find evidence in support of Duverger's prediction: some voters abandon their most-preferred candidate in order to support another candidate who is more likely to win. By contrast, there are much fewer studies exploring the strategic behavior of voters in the first round of runoff elections. Using survey data from Brazil, Plutowski et al. (2020) explore various types of strategic voting behavior in the first round of the 2018 presidential election. They find evidence of voters abandoning their declared preferred candidates in order to vote for another candidate who they deem more likely to win. We see our findings as complementary: while their approach has an advantage in terms of external validity, ours does not have to overcome potential issues with truthful reporting.

There are also some studies exploring the strategic behavior of voters under runoff in the laboratory (Morton and Rietz 2008, Van Der Straeten et al. 2010, 2016, and Tsakas and Xefteris 2020). Most of these experimental studies consider situations in which voters have complete information as to the distribution of voter preferences, which is empirically unrealistic and fails to capture the importance of uncertainty about the outcome in the second round for Duvergerian forces. The only exception is Tsakas and Xefteris (2020). They focus on a common-value setting in which voters have heterogeneous information. In line with our results that Duvergerian forces are sometimes stronger under plurality, they find that the effective number of parties is higher under runoff. In contrast with our result that there are no differences in voters' welfare (as long as we do not factor in the cost of a second round), they find that voters' welfare is higher under runoff. Combined, our results suggest that runoff and plurality differ more in terms of information aggregation than preference aggregation.

There are also empirical studies comparing Duvergerian forces under plurality and runoff using observational data. Fujiwara (2011) and Bordignon et al. (2016) use data from Brazil and Italy, respectively, and regression discontinuity designs to explore this question. They find contradictory results: Fujiwara (2011) finds that the concentration of votes on the two top candidates is stronger under plurality, whereas Bordignon et al.

(2016) find no significant difference.[11] Our results complement these findings in at least two ways.

First, these studies cannot cleanly identify if the effect of runoff is driven by changes in voters and/or candidates' behavior. The main issue impeding the ability of those studies to disentangle the source of the effect is that candidates' behavior, which is largely unobserved by the econometricians, can also be affected by the electoral system.[12] Moreover, even observed choices by candidates may be problematic. For instance, as explained by Bordignon et al. (2016), the decision by a candidate of whether to enter a particular race, which directly affects the concentration of votes on the two top candidates, may or not be driven by the "anticipation" of strategic voting by voters. Without knowing what motivates the decision of the candidate, it is not possible to determine which part of the effect on the concentration of votes is actually driven by the strategic behavior of voters. By contrast, in the laboratory, we can keep the behavior and the number of candidates fixed. This allows us to provide direct and conclusive evidence of the causal effect of the voting system on the behavior of voters. What we lose in terms of external validity, we gain in terms of cleanness of the causal identification.

Second, our results help us understand the contradictory findings of these two studies. Indeed, we find that the extent to which Duvergerian forces differ under plurality and runoff depends on the underlying distribution of preferences in the electorate. Differences appear in the laboratory only in some scenarios: when the minority candidate has sufficiently low support, and when majority voters strongly disagree about which candidate is best. It is possible that Brazilian and Italian elections differ systematically along those dimensions. Indeed, there are more candidates running in the sample of Brazilian races. This could imply systematically lower support for the candidates not in the top two, and more disagreement about which candidate is best.

[11] At first sight, their findings relative to the number of candidates running also appear contradictory. But, they are actually not that different. Fujiwara (2011) finds a positive but not statistically significant effect of runoff on the number of candidates. Bordignon et al. (2016) find a positive and significant effect. As acknowledged by Fujiwara (2011), the non-significance of his effect is most likely due to limitations in sample size.

[12] As stated by Fujiwara (2011): "It must be noted that providing direct evidence to rule out differential unobserved candidate quality and behavior under SB [plurality] and DB [runoff] is (by the definition of unobserved) not possible [...]." For instance, candidates can decide to exert different levels of effort or choose different platforms under runoff than plurality (and there is evidence that their observable behavior does vary, see, e.g., Chin 2019).

3 The Model

The main objective of our theoretical analysis is two-fold: (i) present some key results from the theoretical literature in a simple way, and (ii) extend those results to any electoral size, including small ones. To do so, we rely on a modified, simplified, version of the model in Bouton (2013), with some elements borrowed from Bouton and Gratton (2015). There are two main differences. First, we assume that the size of the electorate is fixed at n, instead of being randomly distributed according to a Poisson distribution. The main reason is that it is not practical to have a random and potentially enormous number of voters in the laboratory. Second, we consider any electorate size, which contrasts with most of the literature's focus on large electorates.

We consider the typical case of a divided majority, with three alternatives, $\{A, B, C\}$, and three types of voters: $t \in T = \{t_A, t_B, t_C\}$.[13] The electorate is split in two groups: majority and minority voters. Majority voters have a common view that C is the worst alternative. Yet, they disagree on which alternative is best: types-t_A prefer A whereas types-t_B prefer B. Minority voters are assumed to prefer C and, for the sake of simplicity, to be indifferent between A and B. Formally, the utilities of the different types are:

$$\begin{aligned} U(A|t_A) &> U(B|t_A) > U(C|t_A), \\ U(B|t_B) &> U(A|t_B) > U(C|t_B), \text{ and} \\ U(C|t_C) &> U(A|t_C) = U(B|t_C), \end{aligned} \tag{1}$$

where $U(W|t)$ denotes the utility of a type-t voter when W is the winning alternative.

Voters are assigned types by i.i.d. draws. The probability that a voter is assigned type t is $r(t)$, with $\sum_{t \in T} r(t) = 1$. These probabilities are common knowledge. In the case of the divided majority, we have:

$$r(t_A) + r(t_B) > \tfrac{1}{2} > r(t_C).$$

Alternative C is thus the (expected) Condorcet loser.[14] We focus on the case in which

[13] The divided majority setting is tractable but still captures some of the fundamental coordination problems that voters face in multicandidate elections. As a consequence, this setting is common in the literature on strategic voting in multicandidate elections (Myerson and Weber 1993, Cox 1997, Fey 1997, Piketty 2000, Myerson 2002, Dewan and Myatt 2007, Myatt 2007, Bouton and Castanheira 2012, and Bouton 2013).

[14] That is, in pairwise contests with the other two candidates, C is the loser. A Condorcet winner has

alternative C is a serious threat to the victory of a majority alternative when majority voters divide their votes: $r(t_C) > 1/3$.

By convention, we focus on the case in which the "more abundant" type among majority voters is t_A:

$$r(t_A) \geq r(t_B).$$

Alternative A is thus the (expected) Condorcet winner.

We consider two electoral systems: $S \in \{P, R\}$, where P refers to plurality and R to runoff. Under plurality, the election is as follows. There is one round of voting. Each voter casts a ballot in favor of one of the alternatives. The alternative that obtains the largest number of votes wins.[15] Under runoff, the election works as follows. In the first round, each voter casts a ballot in favor of one of the alternatives. If the alternative who ranks first obtains (strictly) more than 50% of the votes, it wins outright and there is no second round. Otherwise, a second round with the two alternatives that received the most votes in the first round is held. In this round, each voter casts a ballot in favor of one of the two and the alternative that obtains the most votes wins the election.

We know that the behavior of voters in the first round changes dramatically whether voters are indifferent between an outright victory of a given (majority) candidate in the first round and a second round with that candidate opposing the minority candidate. There are various reasons that make such an indifference unreasonable. First, there is the risk of an upset victory in the second round. This risk comes from the fact that the set of voters participating in the second round may differ from the set of voters participating in the second round (see, e.g., Wright 1989, Bullock III and Johnson 1992 for empirical evidence). This means that the distribution of preferences in the electorate remains uncertain after the first round. Second, even without such remaining uncertainty, a second round is costly for voters (there are some administrative costs, and, more importantly, they have to go to the polls a second time). They thus strictly prefer an outright victory of a given candidate in the first round than a victory of the same candidate in the second round.

For the sake of expositional clarity, we assume that, at the time of the first round, the probabilities of victory of all possible second rounds are given, positive (but potentially

the opposite property.

[15] For the sake of simplicity, ties are resolved by alphabetical order. Results hold if we assume that ties are resolved by the toss of a fair coin.

arbitrarily small), and constant.[16] We denote by $\Pr(i|ij)$, $i, j \in \{A, B, C\}$ the probability that alternative i wins a second round opposing alternatives i and j. For the sake of simplicity, we work under the assumption that $\Pr(A|AC) = \Pr(B|BC)$.[17] The focus under runoff is thus on the behavior of voters in the first round. We also assume that there are no costs associated to the second round for voters. Adding such costs would only reinforce the coordination effect identified in Proposition 1 below.

Both in a plurality election and in the first round of a runoff election, each voter may vote for one of the three alternatives who compete for election. The action set is thus the same under both rules: $\Psi = \{A, B, C\}$.[18]

A strategy under electoral system S is a mapping $\sigma^S : T \to \triangle(\Psi)$, the set of probability distributions over the action set. $\sigma_t^S(\psi)$ denotes the probability that a voter with type t plays action $\psi \in \Psi$ under electoral system S.

For the voting game under electoral system S, we analyze the set of symmetric Bayesian Nash equilibria[19] in which voters do not play weakly dominated strategies.[20] This directly implies that, in the equilibria we are considering, t_C-voters necessarily vote for C, whereas t_A- and t_B-voters do not vote for C.[21]

[16] The results are robust to various other specifications, e.g., some but not all voters participate in the two rounds. Our approach here includes (but is not limited to) any "realistic" restriction (e.g. the front-runner or the candidate with the largest (expected) number of supporters being more likely to win in the second round).

[17] It is a natural assumption in the sense that it is satisfied when we formally include the second round in the model and assume that there is a new draw of voters before the second round.

[18] Given that abstaining is weakly dominated, allowing for abstention would not affect our results.

[19] We focus on symmetric equilibria for three reasons. First, asymmetric equilibria require a level of coordination that seems unrealistic in voting settings; especially in large elections. Second, our experimental design with subjects being assigned types in every period, is not favorable to the emergence of asymmetric equilibria (exactly due to the coordination problem mentioned in the first point). Third, after having analyzed the experimental data, we conjecture that there is no asymmetric equilibrium that can help organize them. Indeed, for any strategy profile in which some t_B-voters always vote sincere, and other always vote Duvergerian (while t_A and t_C-voters vote sincerely), the incentives of sincere t_B voters to vote Duvergerian are stronger than those of Duvergerian t_B voters.

[20] The purpose of this assumption is to eliminate unreasonable equilibria in which all voters vote for the same alternative. These are equilibria because, for those strategy profiles, the probability that one ballot changes the outcome of the election is exactly zero. Voters are then indifferent about which alternative to support, even the one they like least.

[21] Voting for A or B is a weakly dominant strategy for t_C-voters under runoff only because of our assumption that $\Pr(A|AC) = \Pr(B|BC)$. Without that assumption, there would situations in which t_C-voters would prefer to exploit the non-monotonicity of the runoff system by using pushover tactics (in order to have a more favorable second round). We purposedly abstract from this possibility here.

4 Equilibrium Analysis

In this section, we analyze the behavior of voters under both plurality and runoff for any electorate size, including small ones. We focus on pure strategy equilibria under both systems. There are two types of pure strategy equilibria:

Definition 1 *The* sincere voting equilibrium *is such that all voters vote for their most preferred candidate:* $\sigma^S_{t_A}(A) = \sigma^S_{t_B}(B) = 1$.

Definition 2 *A* Duvergerian equilibrium *is such that only two candidates obtain a positive fraction of the votes: either* $\sigma^S_t(A) = 1$ *or* $\sigma^S_t(B) = 1$ $\forall t \in \{t_A, t_B\}$.

There are several reasons justifying our focus on pure strategy equilibria. First, it is sufficient to formally frame the existing debate in the literature about Duvergerian forces and sincere voting. Second, a full characterization of mixed-strategy equilibria for any electoral size is technically challenging. Third, for the cases under consideration in the experimental part, mixed-strategy equilibria both under plurality and runoff are not expectationally stable (see the discussions in Fey 1997 and Bouton 2013).[22] Last but not least, as we will discuss later, these mixed-strategy equilibria do not seem to help organize experimental data.

4.1 Preliminaries

Since voters are instrumental, their behavior depends on the probability that a ballot affects the final outcome of the election, i.e. its probability of being pivotal. The pivotal events and their probabilities given a strategy are different under the two systems. In this paper, we abstract from the formal definition of those pivotal events and the computation of their probabilities. Yet, we use those objects to explain the intuition of the different results. Therefore, it is important to understand the different types of pivotal events under plurality and runoff.

[22] Under plurality, when the sincere voting equilibrium does not exist, there is a mixed-strategy equilibrium in which t_A-voters mix between A and B, and t_B-voters vote for B. Under runoff, there are mixed-strategy equilibria only for some values of the parameters. In one of those equilibria, t_A-voters vote for A, and t_B-voters mix between A and B. In the other, t_A-voters mix between A and B, and t_B-voters vote for B. Figure D1 in Appendix D pins down the first type of equilibria for the values of the parameters considered in the various experimental treatments.

There is only one type of pivotal event under plurality: when a specific ballot changes the outcome of the election from a victory of alternative i to a victory of alternative j. In contrast, in the first round of a majority runoff election there are two types of pivotal events. First, a ballot can be threshold pivotal i/ij. This occurs when alternative i lacks one vote to obtain a strict majority of the votes in the first round. Thus, without an additional vote in favor of i, a second round opposing i to j is held. Second, a ballot can affect the final outcome if it changes the identity of the two alternatives participating in the second round. A ballot is second-rank pivotal ki/kj when alternative k ranks first (but does not obtain an absolute majority of the votes), and alternatives i and j (almost) tie for second place. An additional vote in favor of alternative i allows it, instead of j, to participate in the second round with k.

4.2 Duvergerian equilibria

From now on, we will often refer to the equilibrium in which all majority voters vote for A (B) as Duvergerian A (B) equilibrium. We can show that Duvergerian equilibria exist under both systems even when the electorate is small:

Proposition 1 *If there are strictly more than 4 voters, the two Duvergerian equilibria exist under both plurality and majority runoff.*

Proof. See Appendix A. ■

The intuition is exactly the same as for large electorates. Under plurality, supporters of the alternative expected to receive (almost) no votes, do not want to waste their ballot on their most-preferred alternative which is extremely unlikely to tie for first place, and instead prefer to vote for their second-preferred alternative in order to defeat their least preferred one. Under runoff, supporters of the alternative expected to receive no votes, do not want to waste their ballot on their most-preferred alternative (which has no path to victory). They prefer to vote for their second-preferred alternative in order to increase the probability it wins outright in the first round, and hence avoid the risk of an upset victory of their least-preferred alternative in the second round. This is true even if the risk of an upset victory of candidate C in the second round is arbitrarily small. Moreover, this intuition makes clear that, if the second round is costly for voters, the same result holds even if candidate C is certain to lose in the second round.

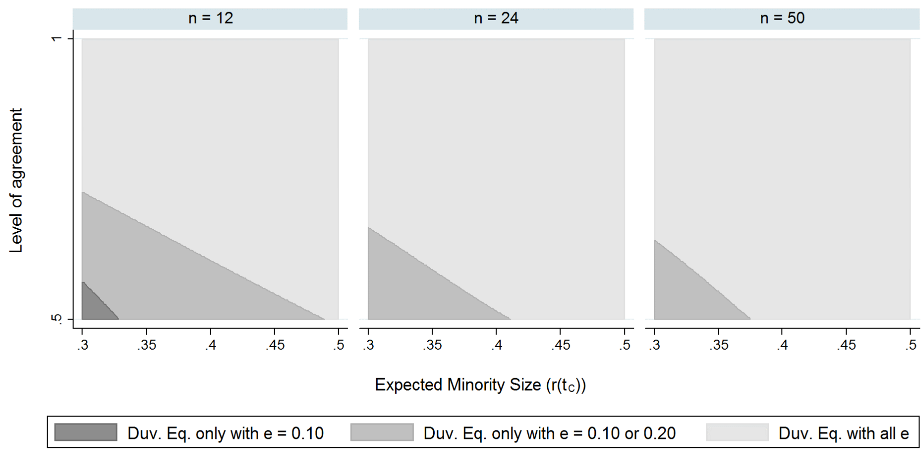


Figure 1: Existence of Duvergerian A equilibrium under Runoff elections for any combination of the size of the minority $(r(t_C))$ and the level of agreement $(r(t_A)/(r(t_A)+r(t_B)))$, for different sizes of the electorate and for sincere trembling frequencies by t_B-voters $\varepsilon \in \{0.1, 0.2, 0.5\}$.

The existence of the Duvergerian equilibria is not an artifact of one of the alternatives having a zero expected vote share. In most situations, majority voters indeed have strict incentives to abandon their most-preferred alternative even if the probability that other majority voters vote for it is not exactly zero. We illustrate these incentives in Figure 1. This figure shows whether the Duvergerian equilibria A exists when t_B-voters "tremble" by voting sincerely for B with probability $\varepsilon \in \{0.1, 0.2, 0.5\}$, for any combination of the size of the minority $(r(t_C))$ and the level of agreement $(r(t_A)/(r(t_A)+r(t_B)))$ and for different sizes of the electorate.[23] The figure shows that (i) Duvergerian equilibria exist with small electorates even with large trembles, (ii) the area for which Duvergerian equilibria do not exist shrinks with ε, (iii) the area for which Duvergerian equilibria do not exist shrinks with n.

Proposition 1 not only proves that coordination forces are strong both under plurality and runoff but also that these forces lead to the coexistence of multiple Duvergerian equilibria. Equilibrium multiplicity is a common and natural feature of voting games. It captures the risk of coordination failure that exists in multicandidate elections (see e.g. Myerson and Weber 1993, Bouton and Castanheira 2012). Yet, at first sight, the existence of a Duvergerian equilibrium in which a majority candidate preferred by the vast majority

[23] The utility values assumed for these simulations are 1 for the top-ranked candidate, 0 for the lowest-ranked candidate and 0.7 for the second-ranked candidate in the case of the majority voters. The same values are assumed for the other simulations presented.

of voters receives no votes may be unsettling. Why would a majority of voters abandon their most preferred candidate? As argued in Bouton (2013), this could happen when a challenger of high quality enters the field. Even if that challenger is of higher quality than the incumbent, the incumbent status could play a focal role: majority voters expect other majority voters to vote for the incumbent. Then, the incumbent is more likely to be in a close race with the minority candidate. Voting for the incumbent is thus individually rational. Other coordination devices such as polls, media, and parties, may help majority voters switch from that bad Duvergerian equilibrium to the other. But, as Bouton (2013, p. 1263) explains: "the existence of [...] Duverger's Law equilibria in runoff elections shows that the runoff system alone does not guarantee that such a coordination failure will never arise. Plurality elections feature the exact same weakness."

4.3 Sincere Voting Equilibrium

For some values of the parameters, the Duvergerian equilibria described in the previous section coexist with the sincere voting equilibrium, in which voters vote for their favorite candidate. In this section, we discuss the existence of such an equilibrium under plurality and runoff by considering numerical examples. The central message is that sincere voting is more prevalent in the first round of a runoff election than in a plurality election.

Figure 2 shows whether the sincere voting equilibrium exists under plurality and runoff for any combination of the size of the minority ($r(t_C)$), the level of agreement ($r(t_A)/(r(t_A)+r(t_B))$), and different sizes of the electorate. It is quite enlightening concerning the conditions under which an equilibrium exists and conveys several messages. First, if a sincere voting equilibrium exists under plurality, it also necessarily exists under runoff, but the converse is not true. This result holds for any size of the electorate. Second, the set of parameter values for which the sincere voting equilibrium exists under plurality shrinks quite rapidly when the electorate grows large. The same is not true under runoff. Finally, if the level of agreement in the majority is high enough, the sincere voting equilibrium does not exist under plurality nor runoff. By contrast, it can exist when the minority is expected to be quite large and the level of agreement among majority voters is sufficiently low.

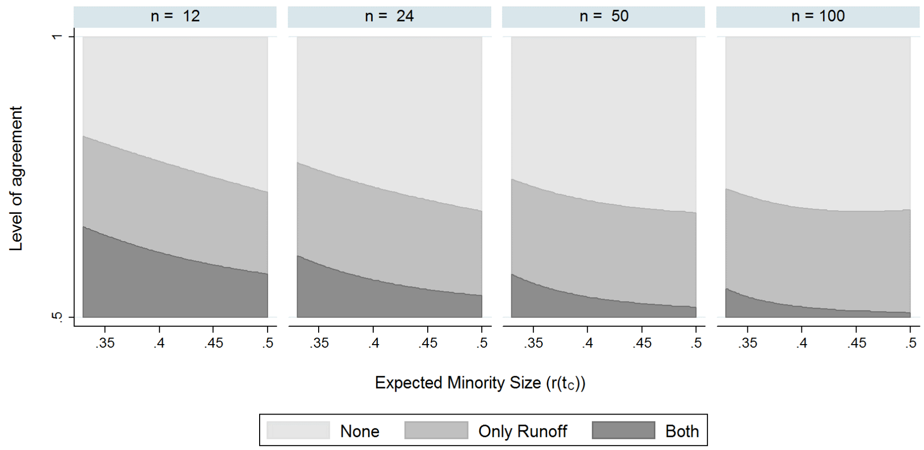


Figure 2: Existence of sincere voting equilibria under both Runoff and Plurality election for any combination of the size of the minority (r_C) and the level of agreement ($r_A/(r_A + r_B)$) and for different sizes of the electorate.

4.4 Electoral Outcomes

Electoral outcomes can be substantially different under plurality and runoff systems. An obvious case is when voters coordinate on different equilibria under the two systems. Differences in outcomes can also emerge when voters coordinate on sincere voting under the two systems, but not when they coordinate on the same Duvergerian equilibrium. To illustrate more concretely the differences in outcomes between plurality and runoff, we consider some numerical examples using the parameter values of the Baseline and Low Disagreement treatments from our laboratory experiments (see Table 1). The overall message is that, in comparison to plurality, runoff improves the selection between the two majority alternatives, A and B. But, this comes at the cost of increasing the risk of a victory of the minority alternative, C.

Under plurality, only the Duvergerian equilibria exist (in both cases). In those equilibria, one majority alternative, say A, is supported by all majority voters. It is thus very likely to win: 76.3% in both plurality treatments. Alternative C wins with the remaining probability (and B never wins). The issue with those equilibria is that there is no possibility of selection between the majority alternatives: only A can win, even if all majority voters prefer B over A. If we assume that the realized distribution of preferences of voters is the true one, then although the minority alternative is defeated when it should be, the plurality system deprives majority voters of the possibility of selecting the alternative with more support.[24]

Under runoff, the Duvergerian equilibria may coexist with the sincere voting equilibrium. This is the case in the Baseline configuration but not the Low Disagreement one. When voters coordinate on a Duvergerian equilibrium, the alternative receiving the support of all majority voters is again very likely to win. The probabilities are similar to those under plurality.[25] Thus, the two systems are similar. By contrast, when voters coordinate on the sincere voting equilibrium, the probability that A wins is sharply lower (49.91%), while the probabilities that B and C wins are higher (13.03% and 37.05% respectively). If we assume that the realized distribution in the first round is the true distribution of preferences (or closer to the true distribution than in the second round), then the higher probability of B winning is a desirable feature of the runoff system. By incentivizing voters to reveal their preferences, runoff systems lead to a better selection between the majority alternatives. This desirable feature comes at a cost: an increase in the probability that C wins either the first round or in the second. If there is a second round, which only happens when there is not a majority of voters supporting C in the first round, then there is a risk of an upset victory of C. Such victories of C, which we view as socially undesirable, happens 13.01% of the time.

5 Experimental Design and Procedures

To test our theoretical predictions, and in particular, to determine if preference revelation differs as a function of the type of electoral system, we conducted controlled laboratory experiments. Subjects played a game with the same structure as the one of the game presented in Section 3. We manipulated the relative support of each group of voters and the voting rule in a between-subjects design. The specific levels of agreement in

[24] Our approach relies on the implicit assumption that discrepancies between the expected distribution of preferences in the electorate (i.e., the $r(t)$ in our theoretical model) and the actual distribution of preferences among voters on the day of the election is welfare relevant. This is reasonable if we view the expected distribution of preferences as a noisy measure (e.g., coming from polls) of the actual distribution of preferences. An alternative view is that the expected distribution of preferences in the electorate is the true distribution, and discrepancies with the actual distribution of preferences among voters on the day of the election arise due to turnout decisions. In our model, that alternative approach would imply that C never has the support of a majority of the electorate ($r(t_C) < 0.5$), and that A is the majority alternative with more support in the electorate ($r(t_A) > r(t_B)$).

[25] The difference is entirely driven by the case in which A and C receive 6 votes each. Due to the alphabetical tie-breaking rule, A wins in that case under plurality. By contrast, under runoff, no alternative wins outright. In the second round, both A and C win with positive probability.

the majority and sizes of the minority that we chose are informed by the theory. We manipulated them to create situations in which sincere voting (i) is an equilibrium under neither systems, (ii) is an equilibrium only under runoff, and (iii) is an equilibrium under both systems (see additional treatments in Section 7). This allows us to test the theoretical predictions that Duverger's coordination forces exist under both plurality and runoff, but that, in some situations, these forces are stronger under plurality.

Design. In all treatments, participants interacted in the same group of 12 voters and played the game for 60 rounds. Alternatives A, B, and C were labeled Green, Purple and Yellow, respectively. Thus, Green and Purple were the two majority candidates and Yellow was the (expected) Condorcet loser. In the beginning of each round, participants were assigned a type by drawing (with replacement) a ball from a hypothetical urn which contained a commonly known composition of green, purple and yellow balls (and the color of the ball indicated the most preferred choice of the voter). The particular *composition* of balls was one of the treatment variables in our experiments. In the *Baseline* (B) set of parameters, the urn consisted of 34 Green (t_A), 22 Purple (t_B), and 44 Yellow (t_C) balls. In the *Low Disagreement* (LD) set we kept the (expected) size of the minority fixed, and increased the agreement between majority voters, by having 43 Green (t_A) balls and 13 Purple (t_B) ones. For the sake of expositional clarity, in the remainder of the paper, we use the same wording as in the theory section when referring to alternatives (A, B, and C) and types of voters (t_A, t_B, and t_C).

The second variable that we varied across treatments was the *electoral system*: either plurality or runoff. Under either voting rule, after learning their type, participants voted for one of the three alternatives. Under *Plurality* (P), the alternative with the most votes won the election, with ties broken alphabetically. Under *Runoff* (R) elections, one of the alternatives won if it received strictly more than 50% of the votes. In case no alternative won in the first round, the two alternatives with the most votes advanced to the second round which was computerized. In line with the theoretical setup, participants knew beforehand the probabilities of victory of each alternative in the second round under all possible scenarios. These probabilities were computed as if there was a new draw of 12 individuals from the same distribution as that of the initial urn, with all participants playing their dominant strategy (and type t_C abstaining when A and B proceeded to the

Treat.	Parameters	Rule	r_A	r_B	r_C	pAB	pAC	pBC	Groups	Sinc	Duv
R_B	Baseline (B)	Runoff	34	22	44	78	76	76	6	✓	✓
P_B	Baseline (B)	Plurality	34	22	44	-	-	-	5		✓
R_LD	Low Disagr. (LD)	Runoff	43	13	44	96	76	76	5		✓
P_LD	Low Disagr. (LD)	Plurality	43	13	44	-	-	-	5		✓

Table 1: Main treatments overview. r_i refers to the probability of becoming type i, i.e., $r_i = r(t_i)$. $p_i j$ is the probability that alternative i defeats alternative j if alternatives i and j advance to the second round.

second round).[26]

The payoffs were the same in all treatments: t_A and t_B-voters received \$11 if their top choice was elected, \$8 if their second-preferred alternative was elected, and \$1 if their least-preferred was elected. t_C-voters received \$11 if their top choice was elected, and \$1 otherwise.

Table 1 summarizes the main treatments and indicates the number of groups of 12 used for each one.[27] It also indicates which types of (pure strategy) equilibria exist: while Duvergerian equilibria exist in all treatments, the sincere voting equilibrium exists only in the R_B treatment—i.e. under the runoff rule with the baseline parameter configuration.

Procedures. Experimental sessions for all treatments were conducted at New York University between October 2017 and April 2018. The subject pool at the university is extremely diverse, drawing from a large population of international students (more than one-third of the student body) as well as students from across the United States. We conducted a total of 38 sessions, with either 12 or 24 subjects. No subject participated in more than one session. Subjects were recruited through the online recruitment software h-root (Bock, Nicklisch, and Baetge, 2014), and the experiment was programmed and conducted using the software z-Tree (Fischbacher, 2007). All sessions were organized under the same procedure: subjects received detailed written instructions, which an experimenter read aloud (see Appendix H). Before starting the experiment, subjects were asked to answer a questionnaire to confirm their full understanding of the experimental design.[28]

[26] The subjects in our experiments were only given the probabilities with which each alternative would win in a hypothetical second round. We did not explain how we computed those probabilities.

[27] Including the additional sessions that we discuss in Section 7 on Alternative Treatments, we conducted a total of 34 sessions. 32 sessions had 12 subjects (1 group) and 2 sessions had 24 subjects (2 groups). In the sessions with 24 subjects, subjects stayed in the same group throughout the session.

[28] Different types of questions were included in this questionnaire. We check whether participants understand: how many periods determine their earnings; how many participants compose each group

After the questionnaire, subjects began to play. At the end of each round, each subject received the following information: (i) the group decision, (ii) the number of votes for each alternative, and (iii) her payoff for that period.

To determine the payment at the end of the experiment, the computer randomly selected three periods and subjects were paid the total sum earned in these three periods. In total, subjects earned an average of \$35.03, a minimum of \$10, and a maximum of \$40. All these figures include a show-up fee of \$7.

5.1 Discussion of Design Choices

Fixed Matching. We used fixed matching to allow voters to better learn the strategies of others and allow them to coordinate. For example, Forsythe et al. (1993, 1996) observe that Duverger's Law equilibria might emerge among voters with a common history—see also Rietz (2008) and Bouton et al. (2016). Fixed matching has the additional advantage that, for a given cost, it delivers more independent units of observation than random matching. A typical drawback of fixed matching is that it favors repeated game effects. One might thus fear that outcomes based on fixed matching could display more cooperative behavior than what theory predicts. The design of the experiment is such that the difference in the ex-ante expected payoff of the Duvergerian A and the sincere voting equilibria is minimal (6.70 vs 6.64 in R_B and 6.92 vs 6.81 in R_LD).[29] Therefore, in our setup, there is no strong case for payoff dominance as a selection criteria between the Duvergerian A equilibrium and the sincere voting equilibrium. Importantly, this desirable feature of our design does not imply that subjects' incentives are weak. The expected utility differences at the interim stage, i.e. when types are realized, are substantial within the context of voting games as shown in Figure D1 in the Appendix D.

Computerized Second Stage. We chose to computerize the second stage for three reasons. First, in pilot sessions using a similar design with additional groups of 12 subjects who only voted in the second round, we found that the dominant strategy of voting for

and whether this composition changes throughout the session; whether their type changes throughout the session; the probability of being assigned to a certain type; the path of the election after two different results in the first round; the probabilities of victory in second-round elections; and payoff configurations as a function of types and election outcomes.

[29] These differences have a similar magnitude in treatments R_SM and R_NU introduced in Section 7. For the Duvergerian B equilibrium, the ex-ante payoff is 6.42 in R_B and 6.20 in R_LD.

the preferred choice was played more than 95% of the time. Therefore, our design reduced the cost with respect to the alternative design while preserving the properties of the second round. Second, a non-computerized second round might have induced subjects to converge to the Duvergerian equilibria in order to minimize the length of the experiment.[30] As mentioned in the theory section, a costly second round indeed gives incentives to voters to coordinate behind one of the majority candidates in the first round.[31] Finally, flexibility on the realized probabilities in the second round allows us to vary the second round probabilities without changing the instructions for the first round. We exploit this possibility in some of the additional treatments (see Section 7).

Multiple Periods. We decided to have subjects vote multiple times for three reasons. First, repetitions allow subjects to get familiar with the trade-offs of the environment. Second, despite fixed matching, we anticipated that subjects would need time to solve the coordination problems. Repetition allows us to (i) identify possible learning effects on those dimensions, and (ii) study how the systems perform once they had time to adjust their strategy. We can still capture possible coordination failures and their effects on systems' performances, by focusing on the first periods of the experiment. Accordingly, in the next section, we indicate whether the results are robust to considering only early periods.

6 Experimental Results

This section summarizes the voting behavior observed in the laboratory. The bottom line is that coordination forces appear to be strong both under plurality and runoff. Voters (partially) coordinate on the Duvergerian A equilibrium under both systems. Yet, as theory predicts, when majority voters strongly disagree about which candidate is best, these forces appear stronger under plurality than runoff. Despite these differences, we find very small

[30] Ex-post, we compared the results of our pilot to those of the experiments reported in this paper. In order to allow for comparibility between the two sets of data despite the differences in the two designs, we propose a measure of the incentives to vote strategically. Even when correcting for those incentives using our measure, we find no evidence of the presence of such additional Duvergerian forces in the case of a non-computerized second round (or any other significant difference in voting behavior). See Appendix G for a detailed discussion.

[31] With our design, the time difference between periods in which a second round took place and those in which it did not (because an alternative won outright) is minimal (only two additional "informational" screens where subjects had no decision to make).

and mostly insignificant differences in electoral outcomes (i.e., which alternative wins), and hence voters' welfare.

Note that we report parametric tests of differences in choices and outcomes based on regression analysis, summarized in the Appendix C (Tables C1-C9). These random effects regressions cluster standard errors at the matching group level, and in some specifications, we control for individual voter characteristics as discussed below. If they lead to different conclusions, we also report non-parametric tests, using averages at the group level as their unit of analysis. All these tests are quite stringent, since the total number of groups for each parameter configuration is small. In some cases, this leads to the conclusion that differences that are large in magnitude are only marginally significant or not significant at all.[32]

Aggregate Voting Behavior. Table 2 summarizes aggregate voting behavior in the four treatments. Two clear patterns emerge in all treatments. First, and in line with the theoretical predictions, t_C-voters overwhelmingly play their weakly dominant strategy to vote for alternative C. The frequency of playing this strategy is above 95% in all the treatments and there are no significant differences across electoral systems (p-value = 0.829 for Baseline treatments and p-value = 0.121 for LD treatments). Second, t_A-voters massively choose alternative A: the frequency is above 97% in all treatments and there are no significant differences across electoral systems (p-value = 0.137 for Baseline treatments and p-value = 0.615 for LD treatments). As a consequence, it is clear that there is no convergence to the Duvergerian B equilibrium. If there is convergence to a Duvergerian equilibrium it will be to the one in which majority voters coordinate on alternative A. This depends on the behavior of t_B-voters.[33]

Under plurality, if t_B-voters correctly anticipate the behavior of t_A- and t_C-voters, then they face no trade-off: they should abandon their favorite alternative and vote for alternative A. This is true even if all other t_B-voters vote sincerely for B. And indeed, we

[32] We also estimated models (available upon request) in which we used the Wild Cluster Bootstrap adjustment (Roodman et al. 2019), without the results changing substantially.

[33] Given the co-existence of two Duvergerian equilibria, it is somehow surprising that majority voters always coordinate behind alternative A. One potential explanation is that the status of A as the majority alternative with the higher (expected) support in the electorate, i.e., $r(t_A) > r(t_B)$, makes the Duvergerian A equilibrium the focal equilibrium. Another, related, possible explanation is that Duvergerian A equilibrium payoff dominates Duvergerian B equilibrium (6.70 vs 6.42 in R_B and 6.92 vs 6.20 in R_LD). Then, repeated play effects could favor coordination of majority voters on that Duvergerian equilibrium.

		Baseline			Low disagreement		
		% A	% B	% C	% A	% B	% C
Plurality	t_A	99.1	0.41	0.49	98.17	1.57	0.26
	t_B	75.25	24.14	0.62	71.49	27.92	0.59
	t_C	1.47	0.89	97.64	2.82	1.22	95.97
Runoff	t_A	97.21	2.38	0.41	97.34	2.14	0.52
	t_B	58.72	40.26	1.01	70.55	28.6	0.85
	t_C	1.34	0.97	97.69	1.58	0.69	97.73

Table 2: Aggregate behavior in main treatments.

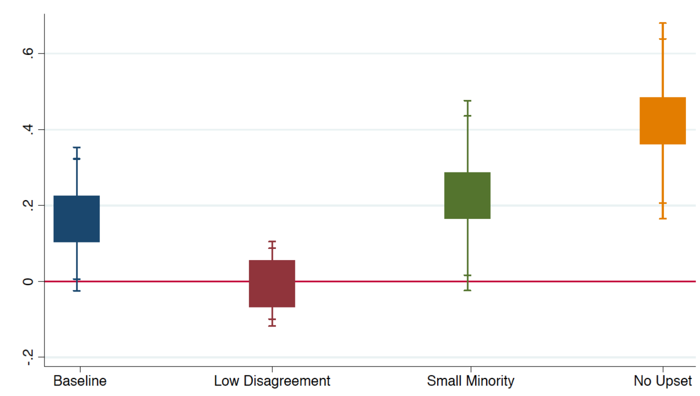


Figure 3: Treatment effects of the voting rule on the probability of strategic voting by t_B-voters. 90% and 95% confidence intervals are drawn from the regression in the Appendix C (pages 10-15).

find that t_B-voters play this best response a bit more than 70% of the time. Nevertheless, the frequency with which they vote for their favorite alternative is non-negligible: 24.14% in P_B treatments and 27.92% in P_LD treatments.

In treatment R_LD, if they correctly anticipate the behavior of t_A- and t_C-voters, the incentives for t_B-voters to vote sincerely under runoff are similar to those under plurality: the sincere voting equilibrium does not exist. In this treatment, t_B-voters vote sincerely as infrequently as in the P_B treatment: only 28.6% of the time. Figure 3 displays the treatment effects of the voting rule on the behavior of t_B-voters (in order to save space, it also includes the coefficient for two additional treatments, SM and NU, which we describe in Section 7). As the figure shows, in the LD treatments, the frequency of sincere voting by t_B-voters under runoff is essentially identical as under plurality (no significant difference, p-value = 0.727). This is in line with the theoretical predictions: only the Duvergerian equilibria exist under the two systems. This result holds not only for the whole sample, but also if we restrict the analysis to the first ten periods. Furthermore, as Figure 4 shows, the amount of sincere voting decreases over time in all groups.[34]

Unlike in the other treatments, in treatment R_B, the sincere voting equilibrium and the Duvergerian equilibria coexist. Taking as given the behavior of t_A-voters (who almost always vote for A), this means that there are two equilibria in pure strategies: one in which t_B-voters all vote for A, and another one in which they all vote for B. In this treatment, t_B-voters voted for their favorite candidate 40.26% of the time, which contrasts with the 24.14% observed under plurality. As Figure 3 shows, we find a significant effect at the 10% confidence level in the regression analysis. Table C1 in Appendix C shows that the runoff treatment always has a positive and significant effect, of about 18 percentage points, on the sincerity of t_B-voters.[35]

Given the multiplicity of equilibria, it could be that groups coordinate on different equilibria. Figure 4 shows that 4 out 6 groups feature a clear (but only partial) convergence to the Duvergerian equilibrium A, with more than 75% of t_B-voters voting for A in the last 10 periods. While this convergence is slow, it seems consistent with the theoretical

[34] In the R_LD treatment, sincerity of t_B-voters in group 1, for instance, decreases from 71% on average during the first five periods, to 33% over the last five; it goes from 57% to 40% for group 2; from 62% to 29% for group 3; from 33% to 14% for group 4; and from 29% to 10% for group 5.

[35] Despite the size of the difference, it is not significant with non-parametric tests (Wilcoxon, p-value = 0.144).

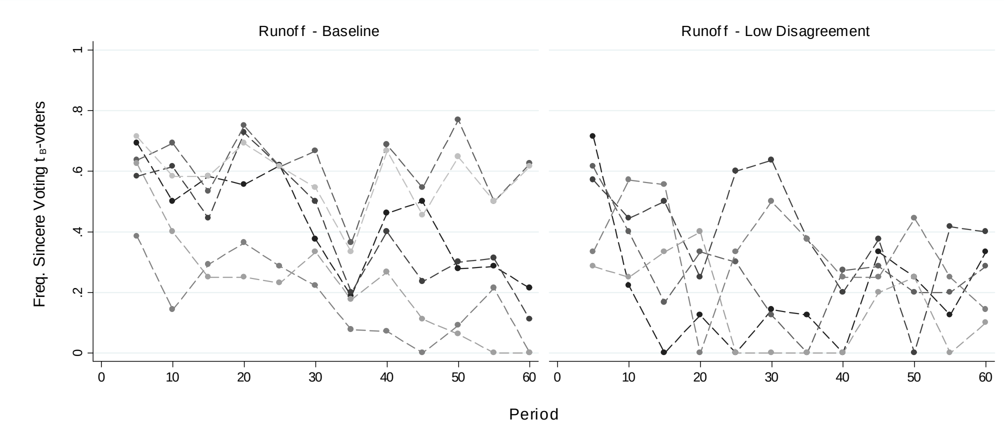


Figure 4: Temporal evolution of sincere voting by t_B-voters in R_B and R_LD treatments. Each point aggregates behaviour in groups of five periods.

predictions (assuming voters anticipate correctly the behavior of other voters). Indeed, voting for A is a best response for a t_B-voter when the other t_B-voters vote for B less than 58.45% of the time.[36] Even in the early periods, t_B-voters in the four converging groups vote for A at a frequency close or lower than 58.45%. And this frequency decreased over time, increasing the incentives of t_B-voters to abandon their favorite alternative.

The case of the two other groups is more puzzling. Figure 4 shows that the frequency of sincere voting by t_B-voters remains somehow stable around 60% over the 60 periods, with drops to as low as 40% in the middle of the experiment. Under the assumption that voters anticipate correctly the behavior of other voters, theory would thus predict that t_B-voters should have abandoned their favorite alternative. Yet, we do not observe the same convergence to the Duvergerian A equilibrium as in the four other groups. Below, we discuss some explanations of why this could be the case.

Overall, the analysis of aggregate voting behavior highlights the presence of strong coordination forces under both plurality and runoff. This is true even in early periods. Yet, the convergence to the Duvergerian A equilibrium is only partial in some groups, especially when majority voters are strongly divided about which alternative is best. Moreover, when a sincere voting equilibrium exists under runoff (but not under plurality), these coordination forces appear stronger under plurality than runoff. Analyzing the voting

[36] Figure D1 in Appendix D displays the utilities of t_B-voters from voting either A or B, for all treatments. Note that this threshold is based on the assumption of risk neutrality. However, laboratory participants typically exhibit moderate risk aversion. As Figure B1 in Appendix B indicates, the threshold of 58.45% increases with risk aversion.

behavior at the individual level allows for a better understanding of this phenomenon.

Individual Behavior. Figure 5 displays a representation of individual behavior, focusing on the behavior of t_B-voters. The horizontal (vertical) axis reports the frequency of sincere voting in the first (second) half of the experiment. The diameter of each hollow circle in the graph corresponds to the number of subjects who played at those frequencies: the larger this number, the larger the diameter.

Figure 5 shows several interesting patterns. First, most circles are below the 45 degree line, indicating that the frequency of sincere voting decreased for most subjects. The proportions of subjects who strictly (weakly) decrease the percentage of sincere voting in the second half are 35% (95%), 33% (88%), 51% (86%), 43% (92%) in treatments P_B, R_B, P_LD and R_LD respectively. Second, we observe two opposite clusters along the 45 degree line: voters who played (almost) always sincere (upper right corner) and voters who (almost) always chose A (lower left corner). The weights of these clusters vary across treatments. The percentage of voters who voted Duvergerian in all periods are 43.33%, 41.67%, 23.61% and 36.67% in treatments P_B, P_LD, R_B and R_LD respectively. The percentages of subjects who voted always sincere are 15.00%, 13.33%, 8.33% and 11.67%, respectively.

The change in weights of these clusters across treatments shows that at least some of these voters react to incentives. One possible explanation is that voters play a mixed strategy equilibrium. Yet, considering mixed strategies equilibria does not seem to be the culprit here. Figure D1 in Appendix D shows that, in the R_B treatment, there is a mixed-strategy equilibrium in which t_B-voters mix between A (with probability 42%) and B (with probability 58%). At first sight, it seems that a purified version of that equilibrium organizes the data reasonably well. Yet, there is a substantial issue with that explanation: we observe similar patterns in the R_LD, P_B, and P_LD treatments, even if such a mixed strategy equilibrium does not exist in those treatments.[37] This issue comes on top of the aforementioned issue of the stability of those mixed strategy equilibria, which makes them unlikely to emerge.

Another possible explanation for those clusters is that there is heterogeneity among

[37] We do not include the equivalent of Figure D1 in the Appendix for the plurality treatments because, in those treatments, as in the R_LD treatment, type-t_B voters' utility of voting for A is always larger than the utility of voting B.

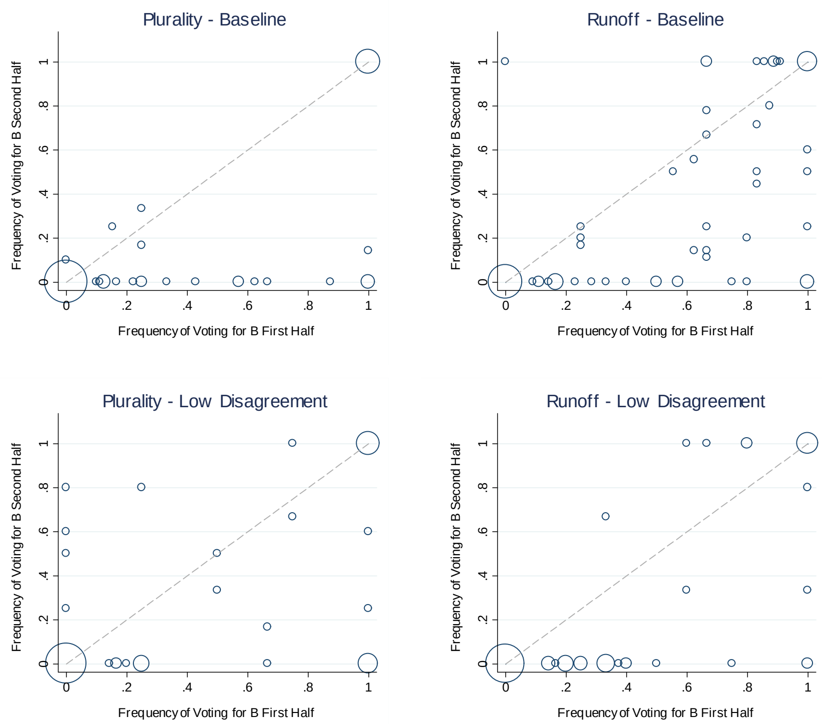


Figure 5: Individual behavior in the main treatments.

voters with respect to their propensity to vote sincere or Duvergerian. Due to some expressive motives (see Schuessler 2000), some voters may be very reluctant to not vote sincerely, while others may be very quick at pulling the Duvergerian trigger. The presence of such heterogeneity would help understand the patterns in the data described above. First, the presence of voters especially prone to sincere voting would weigh against the emergence of a Duverger's law equilibrium. In particular, the presence of those voters would prevent the full convergence to a Duvergerian equilibrium. Second, the presence of prone-to-Duvergerian voters would similarly weigh against the emergence of the sincere voting equilibrium. The issue would again be with convergence.

Yet another possible explanation for the always Duvergerian behavior is a high level of risk aversion. Indeed, as Figure B1 in Appendix B shows, the sincere voting equilibrium exists in none of the treatments if voters are sufficiently risk averse. While the aggregate level of risk aversion that prevents the existence of the sincere voting equilibrium seems unreasonably high, it can still be that some t_B-voters are sufficiently risk averse to prefer voting for A even when all other voters vote sincerely.

(i) Always Duvergerian (ii) Always Sincere

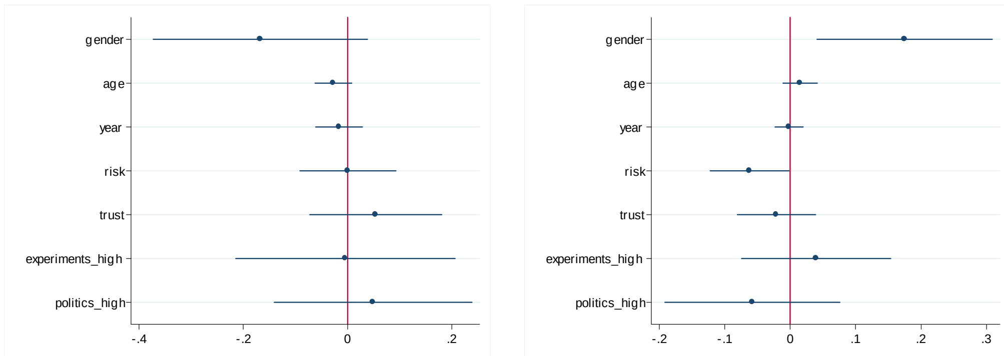


Figure 6: Individual characteristics and "always Duverger" and "always sincere" behavior.

To explore this possibility, for t_B-voters, we estimated a linear regression where the outcome is a dummy indicating if the subject always voted for A, and as regressors we used a battery of self-declared individual-level traits, including age, gender, years in college, self-reported measures of risk aversion and trust, experience participating in experiments and interest in politics. The left panel of Figure 6 shows the coefficients of that regression. Risk aversion does not seems to distinguish always Duvergerian voters from others. In fact, our measure of risk aversion is distributed quite similarly for always Duvergerian voters and the entire population of t_B-voters. The only distinguishing individual characteristic of always Duvergerian voters is gender: women are under-represented in that group. The proportion of women among always Duvergerian voters is 31.67% (while the percentage of women in these treatments is 45.31%).

We can similarly explore whether voters' risk aversion, or other characteristics, are associated with a higher propensity to vote sincerely. The right panel of Figure 6 shows the coefficients of a linear regression where the outcome is a dummy indicating if the subject always voted for B when type t_B, and as regressors we used the same individual-level traits as above. Risk aversion does not significantly distinguish always sincere voters from others. Again, the only distinguishing individual characteristic of always sincere voters is gender: women are over-represented in that group. The proportion of women among always sincere voters is 64.41% (compared to an average of 45.31%).

Finally, it is possible that the always Duvergerian behavior of some voters stems from a misunderstanding of the influence of their ballot on the outcome of the election. Incentives

			Baseline			Low Disaggr.		
			%A	%B	%C	%A	%B	%C
Plurality	Realized	All periods	67.69	0.09	32.22	70.74	0.00	29.26
		Second Half	70.40	0.03	29.57	70.67	0.00	29.32
	Theory	Duv. Eq.	76.30	0.0	23.70	76.27	0.00	23.73
Runoff	Realized	All periods	65.31	1.81	32.88	69.82	0.06	30.11
		Second Half	66.23	1.23	32.53	69.69	0.02	30.29
	Theory	Duv. Eq.	71.24	0.00	28.76	71.51	0.00	28.49
		Sincere	50.12	13.13	36.75	-	-	-

Table 3: Simulated outcomes in main treatments. *All periods* uses the average frequencies for each while *Second Half* uses the averages only for the second half. The equilibrium benchmarks are based on the simulated data.

to coordinate behind the stronger majority candidate are indeed high for voters who wrongly believe that they can prevent an outright victory of C in the first round (as in plurality). But, if C obtains 7 or more votes, then a vote for A or B cannot prevent C's victory. The only effect of coordination in the first round on C's probability of victory is through the decreased likelihood that a second round takes place, in which an upset victory of C can occur. The additional treatments summarized in Section 7 give some credence to this possible explanation: even when the risk of an upset victory of C in the second round is essentially null, some majority voters continue to vote always Duvergerian.

Outcomes and Welfare. As we explained in Section 4.4, electoral outcomes (i.e., which alternative wins) can be substantially different under plurality and runoff. Yet, theory predict that differences should not arise when voters play the same Duvergerian equilibrium under both systems. Due to the (partial) convergence to a Duvergerian A equilibrium under both systems, we find very small differences in outcomes, and hence voters' welfare, across treatments.

In order to produce comparable numbers between treatments and institutions, we present the result of simulations on outcomes based on individual average behavior. That is, for each group, we simulate 10,000 decisions based on the individual behavior of each voter.[38] Table 3 presents both the theoretically and the simulated probabilities of victory

[38]We ran the simulations as follows. First, we computed the individual frequencies with which each subject voted for each of the alternatives for each of their possible types t_A, t_B and t_C (in the second half of the experiment or in all periods). Second, we simulated 10,000 elections for each group. For each of these simulations, we (i) randomly assigned a type to each voter, (ii) simulated a vote for each voter given their type and their observed voting frequencies, and (iii) aggregated the votes and selected the winning alternative. We ran this process independently for each group and then aggregated the results by

of the different alternatives under our treatments. The table shows that outcomes do not differ much between electoral systems and sets of parameters. The realized probabilities of victory are actually close to the theoretical probabilities.

Regression analysis also highlights small differences across treatments. Table C7 in Appendix C reports the results of random effects regressions, in which we regress the probability of victory of each candidate on a runoff indicator and for the different parameter configurations. Columns 1-3 show that in the Baseline treatments, the electoral system has no effect on the probability that A wins. In contrast, a victory of B is 1 percentage point more likely (significant at the 5% level) in runoff elections, while C wins with a probability of 5 percentage points lower (significant at the 10% level) in runoffs. For the LD treatments, the probabilities of A and C winning do not vary significantly across electoral systems.[39]

The previous discussion highlights that, in our laboratory experiments, runoff elections do not differ significantly from plurality elections in terms of outcomes (and hence welfare). Yet, this conclusion overlooks one important difference between the two systems: voters may have to turn out twice in a runoff election. And indeed, we find that a second round was necessary 47.78% of the time in the R_B treatment, and 33% of the time in the R_LD treatment. As Figure 7 shows, under the R_B treatment, the frequency of a second round decreased over time (in concert with the decrease in sincere voting by t_B-voters). Thus, while the coordination on the Duvergerian equilibrium is sufficient to prevent differences in outcomes (and, in particular, the victory of alternative B when there are many t_B-voters in the electorate), it is not sufficient to ensure a systematic outright victory of either A or C in the first round.

7 Additional Treatments

In this section, we summarize additional treatments designed to explore (i) the robustness of the finding that coordination forces are strong under runoff elections, and (ii) the mechanisms underlying the choice of voters to abandon their favorite alternative. We do so by increasing the incentives to vote sincerely along two dimensions: by reducing either

treatment.

[39] We cannot estimate analogous models for the probability of B winning the election, because in LD treatments, B never wins an election.

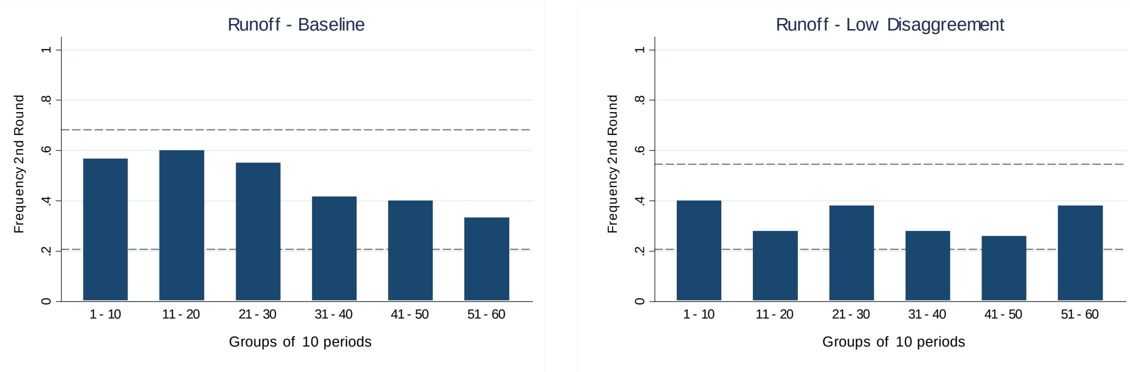


Figure 7: Probability of reaching the second round over time. The upper (lower) gray dashed line indicates the theoretical probability of reaching the second round in the sincere voting (Duverger's Law) equilibrium.

the size of the minority or the risk of upset in the second round. We find that, although the frequency of sincere voting increases, coordination forces are still strong under runoff. We also find that a substantial fraction of voters continue to vote Duvergerian, even when the instrumental motives to vote in such a way are very weak. This reinforces the interpretation that some Duvergerian voters are misunderstanding the influence of their ballot on the outcome of the election.

7.1 Experimental Design

We designed two additional treatments with (i) the same basic structure and procedures as the ones described in Section 5, and (ii) sets of parameters that are related to those of the Baseline treatment. In those treatments, described in Table 4, the incentives of majority voters to vote sincerely are stronger than in the Baseline treatment.

In the *Small Minority* (SM) treatment, we reduced the size of the minority, $r(t_C)$, by 5 percentage points, to 39%. We also increased $r(t_A)$ and $r(t_B)$ as close as possible to proportionality, given the constraint of using integer numbers with the subjects, to 37% and 24%, respectively. The decrease in the expected size of the minority increases the incentives of majority voters to vote sincerely in two ways. First, when majority voters vote sincerely, it is more likely that the second round opposes A and B, the two majority candidates. Second, the probability of an upset victory of C, the minority candidate, in the second round is lower. As a result, the incentives to vote sincerely are stronger than in the Baseline treatment: if t_A and t_C-voters vote sincerely, the minimal percentage of

Treat.	Parameters	Rule	r_A	r_B	r_C	pAB	pAC	pBC	Groups	Sinc	Duv
R_SM	Small Min. (SM)	Runoff	37	24	39	78	86	86	5	✓	✓
P_SM	Small Min. (SM)	Plurality	37	24	39	-	-	-	5	✓	✓
R_NU	No Upset (NU)	Runoff	34	22	44	78	99	99	5	✓	✓

Table 4: Overview of additional treatments.

		Small Minority			No Upset		
		% A	% B	% C	% A	% B	% C
Plurality	t_A	96.59	1.59	1.81	-	-	-
	t_B	66.08	33.33	0.58	-	-	-
	t_C	3.00	0.66	96.34	-	-	-
Runoff	t_A	97.17	1.96	0.87	92.34	5.58	2.08
	t_B	43.02	54.19	2.79	31.92	67.14	0.94
	t_C	2.05	1.47	96.48	2.13	1.35	96.52

Table 5: Aggregate behavior in the additional treatments.

sincere voting by t_B-voters that makes sincere voting a best response for those voters under runoff is 47.87%, compared to a threshold of 58.45% in the Baseline treatment. We conducted sessions with these parameters under both plurality and runoff elections. Unlike the previous treatments, in this case sincere voting is also an equilibrium under plurality.

In the *No Upset* (NU) treatment, we reduced drastically the probability of an upset victory of C in the second round, by setting $\Pr(A|AC) = \Pr(B|BC) = 99\%$. All other parameters remain at the Baseline level. As a result, the incentives to vote sincerely are very strong: if t_A and t_C-voters vote sincerely, sincere voting is the best response for a t_B voters under runoff as long as other t_B-voters vote sincerely with a probability of at least 17.56%. We conducted sessions with these parameters only under runoff elections, because the comparable case under plurality is the baseline treatment.

7.2 Experimental Results

Aggregate Voting Behavior. Voting behavior in the new treatments is summarized in Table 5. In line with the previous treatments, t_A- and t_C-voters vote overwhelmingly for their preferred alternative, and therefore, we focus on the behavior of t_B-voters.

Under R_SM, sincere voting equals 54.19%, as opposed to 40.26% observed in the R_B treatment. However, Table C5 in Appendix Ca shows that this difference is not significant

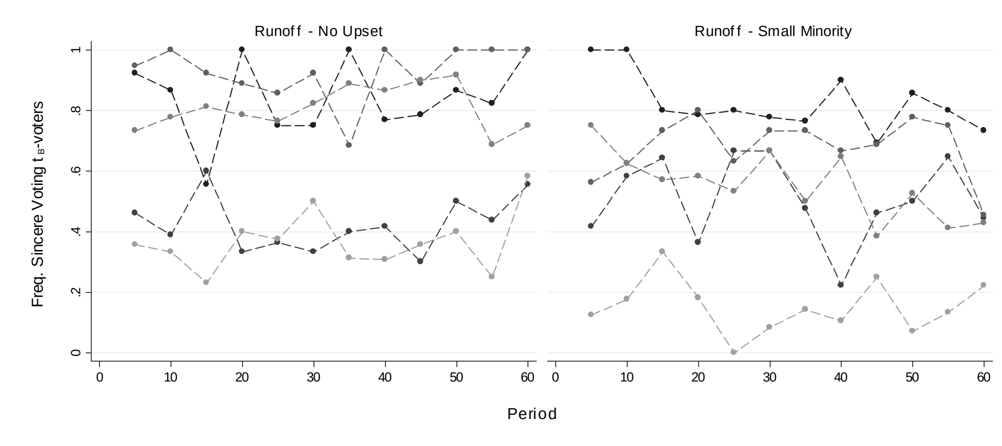


Figure 8: Temporal evolution of sincere voting by t_B-voters in R_NU and R_SM treatments. Each point aggregates behaviour in groups of five periods.

for the random effects model (p-value=0.194 in the model with controls, clustered at the group level). These results are another piece of evidence supporting the fact that Duvergerian forces are strong under runoff: even with a smaller threat of the minority candidate, Duvergerian voting still represents more than 40% of decisions.[40]

As in the Baseline treatment, there is substantial heterogeneity across groups under R_SM: the percentages of sincere voting in the different groups were 15.25%, 50.64%, 56.15%, 68.51%, and 82.39%. This pattern, together with Figure 8 (right-panel), highlight an interesting feature of voters' behavior. Even in groups in which the best response of t_B-voters is clear and stable during the whole session (sincere voting in the last group, and Duvergerian voting in the first group), we still do not observe full convergence. Actually, for the first group, we have that all t_B-voters vote sincerely at the beginning of the experiment, but then some voters move away from that equilibrium.

The R_NU treatment features even stronger incentives to vote sincerely: it is the best response as long as other t_B-voters vote sincerely with a probability of at least 17.56%. And indeed, we observe a significantly higher amount of sincere voting among t_B-voters than in all other runoff treatments: 67.14%. Table C6 in Appendix C shows that, using random effects regressions, compared to the Baseline treatment, t_B-voters are 28 percentage points more likely to vote sincerely (p-value=0.018).[41] This highlights that the threat of an upset

[40] This is also true in the first ten periods of the experiments: the frequency of voting for A by t_B-voters is 38.56%.

[41] The difference is even larger when compared to the P_B: Table C5 in Appendix C shows that the difference of about 49 percentage points of sincere voting between runoff and plurality is highly significant,

victory of the minority candidate is key to explain the importance of Duvergerian forces in the first round of runoff elections. Yet, we find it striking that a non-negligible fraction of these voters continue to abandon their most-preferred candidate, even if incentives to do so are much weaker than in the previous treatments.

As with the other treatments, R_NU also displays some heterogeneity across groups, as shown in the left-panel of Figure 8. There are two sets of groups. Three groups feature behavior in line with the sincere voting equilibrium. The frequency of sincere voting in these groups is over 80%. The two remaining groups feature no convergence to either the sincere or to the Duverger's law equilibrium. In these groups, t_B-voters vote for B around 40% of the time, which means that the best response to observed behavior is unambiguously to vote sincerely. Moreover, we do not find any trend, which excludes even slow convergence towards the sincere voting equilibrium.

Individual Behavior. Figure 9 represents individual behavior of t_B-voters under the SM (plurality and runoff) and NU (runoff) configurations. Consistent with the results in the main treatments, there are two prominent clusters of voters: those who always vote sincerely and those who always vote Duvergerian. Again, the relative size of these two clusters varies across treatments. The percentage of always sincere voters is smaller under plurality than under both runoff treatments. In P_SM, 15% of voters always vote sincere, while this percentage is 38% under both R_SM and R_NU. There are also significantly more such voters in these new runoff treatments, compared to R_B, where the proportion of always sincere voters was just 8.3%. The percentage of always Duvergerian voters are 28.33%, 21.67% and 13.33% in P_SM, R_SM and R_NU, respectively. This compares to 23.67% of always Duvergerian voters in the R_B treatment.

We find these patterns instructive. First, the significant change in the fractions of always sincere and always Duvergerian voters across treatments reinforces the interpretation that there is heterogeneity among voters in their propensity to vote sincere or Duvergerian. Second, we find it striking that, in R_NU, 13% of voters always vote Duvergerian, even if the risk of upset victory of C in the second round is almost null. We interpret this result as suggestive evidence that a substantial fraction of voters misunderstand the influence of their ballot in the first round of a runoff election.

in the random-effects model that includes subject-level covariates.

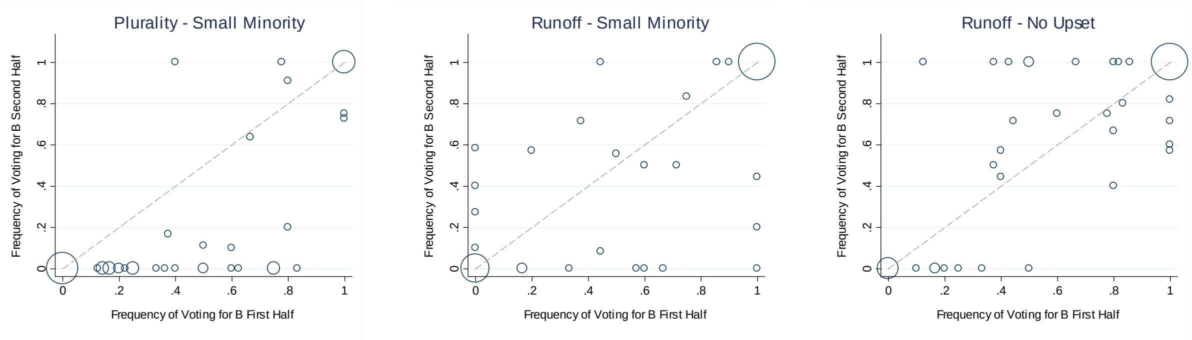


Figure 9: Individual behavior in the new treatments.

As in the main treatments, the only characteristic of voters that appears to play a role in explaining why some t_B-voters always vote sincere or Duvergerian is gender: female participants are less likely to be always Duvergerian and more likely to be always sincere. See Figure E1 in Appendix E.

Outcomes and Welfare. We now turn to the realized electoral outcomes for the SM and NU treatments. The change of behavior observed in these new treatments raises two questions: (i) are there differences in outcomes under runoff compared to the Baseline treatment? and (ii) are there bigger institutional differences than the ones observed in Section 6? Table 6 presents both the theoretically and the simulated probabilities of victory of the different alternatives based on individual average behavior for the new treatments.[42] The analysis clearly suggests that the answer to both questions is positive, which makes sense given the increase in sincere voting in the new treatment.

Let us first focus on the differences in outcomes under runoff rule across treatments. For this purpose, we compare the probability of victory of each alternative under the new treatments versus the Baseline treatment. Columns 1-3 of Table C8 (in Appendix C) show that in the case of Small Minority, the reduction in the expected proportion of t_C-voters has a negative effect on the probability that C wins the election. Although, the combined likelihood that either A or B wins the election increases with the same magnitude, the coefficients for these two outcomes are positive but not individually significant. For the No Upset treatment, there is a sharp and significant higher probability that B wins the election (of about 12 percentage points), while the probability that C wins diminishes by

[42] The simulations were produced exactly as in the previous section (see Footnote 38 for details).

			Small Minority			No Upset		
			%A	%B	%C	%A	%B	%C
Plurality	Realized	All periods	76.21	0.30	23.50	-	-	-
		Second Half	78.71	0.04	21.25	-	-	-
	Theory	Duv. Eq.	86.02	0.00	13.98	-	-	-
		Sincere	48.25	11.10	40.64	-	-	-
Runoff	Realized	All periods	74.99	5.35	19.66	67.23	9.84	22.93
		Second Half	74.83	3.90	21.27	67.69	9.44	22.87
	Theory	Duv. Eq.	83.46	0.00	16.54	75.83	0.00	24.18
		Sincere	61.80	16.55	21.65	59.41	16.08	24.51

Table 6: Simulated outcomes in the new treatments. *All periods* uses the average frequencies for each while *Second Half* uses the averages only for the second half. The equilibrium benchmarks are based on the simulated data.

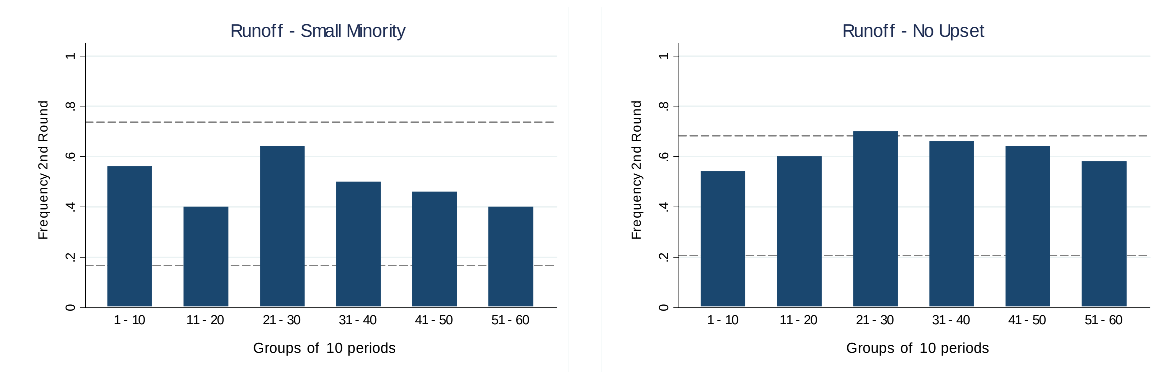


Figure 10: Probability of reaching the second round over time. The upper (lower) gray dashed line indicates the theoretical probability of reaching the second round in the sincere voting (Duverger's Law) equilibrium. Note that the theoretical predictions in $R_S M$ are the same that in R_B.

6.6 percentage points. Alternative A also exhibits a reduction in its probability of victory, albeit not statistically significant. Hence, making extremely likely a defeat of C in the second round increases the chances of victory of B substantially.

Let us now analyze the differences in outcomes for the new runoff treatments compared to their plurality counterparts. For the Small Minority treatment, we compare R_SM with P_SM, while in the case of No Upset, R_NU is compared with P_B. In the Small Minority treatment (Table C9, columns 1-3), runoff elections increases B's chances of victory by 4.3 percentage points (p-value = 0.074), at the expense of A and C. In the No Upset treatment (Table C9, columns 4-6), the increase in the probability that B wins the election is even higher (13 percentage points and p-value = 0.014) and is explained by a decrease in C's winning probability.

The benefits of better preference revelation and aggregation under runoff is associated with a potential increase in the likelihood of having a second round. Remember that the frequency of a second round in the R_B treatment was 47.78%. Under R_SM, we observed only a marginal and not significant increase to 49.33% (p-value = 0.895). Under R_NU, however, there was a significant increase to 62.00% (p-value = 0.054). As Figure 10 shows, there is no strong temporal change over time. The figure also shows that, while under R_SM, the probability of a second round lies between the predictions of the sincere voting and Duvergerian equilibria, the frequency of a second round under R_NU is close to the prediction under the sincere voting equilibrium.

8 Conclusion

One believed advantage of runoff electoral systems is that, compared to plurality rule, they favor preference revelation by voters: Duvergerian incentives are supposedly weaker. Such preference revelation is seen as advantageous: it can lead to the selection of a better alternatives, and also inform both voters and candidates as to the distribution of preferences in the electorate.

However, recent theoretical analyses show that Duvergerian incentives also exist in runoff systems. Voters may vote strategically as they are predicted to do under plurality rule: some voters abandon their most preferred alternative when it is unlikely to be in contention for victory. In fact, in some situations, the only equilibria under runoff electoral systems involve such strategic behavior.

In this paper, we investigate to which extent voting is more sincere in runoff systems than it is in plurality elections using laboratory experiments. Our results point towards the existence of strong coordination forces under both systems. Yet, we also find evidence that, in line with the theoretical predictions, there are situations in which these forces are weaker under runoff. Only then are runoff elections more effective than plurality elections in allowing voters to better reveal their preferences.

The differences in voting behavior under plurality and runoff elections are not substantial enough to lead to large differences in terms of electoral outcomes and voters' welfare. In our laboratory experiments, the only substantial difference between the two systems is that, even with the observed high levels of strategic voting, runoff elections required a second round in many instances.

Our experimental results suggest that, when one takes the behavior of candidates and parties as given, plurality and runoff elections are not fundamentally different. The higher cost of the runoff system would then tilt the balance in favor of plurality. Yet, there is evidence that candidates and parties behave differently under these two systems. For instance, Chin (2019) finds that candidates appeal to a broader voter base in runoff elections. These differences may justify the use of runoff instead of plurality. Future work should explore further the different incentives of candidates and parties under plurality and runoff, taking into account that the behavior of voters under the two systems only differ in specific situations.

References

[1] Ballard, Mark. (2016). 'In rarely seen twist, last-minute candidate Taylor Barras selected Louisiana House speaker', *The Advocate,* Jan. 14, 2016, https://www.theadvocate.com/baton_rouge/news/politics/legislature/article_1e1b57db-13d3-5822-b86f-62ed800f1ce3.html, accessed May 19, 2019.

[2] Bendor, J., Diermeier, D., Siegel, D. A., & Ting, M. (2011). A behavioral theory of elections. Princeton University Press.

[3] Blais, A. (2004a). Strategic voting in the 2002 French presidential election. In The French Voter: Before and After the 2002 Election (pp. 93-109). Palgrave Macmillan, London.

[4] Blais, A. (2004b). Y a-t-il un vote stratégique en France?. In Le nouveau désordre électoral (pp. 279-301). Presses de Sciences Po (PFNSP).

[5] Bock, O., Baetge, I., & Nicklisch, A. (2014). 'hroot: Hamburg registration and organization online tool', *European Economic Review*, vol. 71, pp. 117-120.

[6] Bordignon, M., Nannicini, T., & Tabellini, G. (2016). 'Moderating political extremism: single round versus runoff elections under plurality rule', *American Economic Review*, vol. 106(8), pp. 2349-70.

[7] Bormann, N. C., & Golder, M. (2013). 'Democratic electoral systems around the world, 1946–2011', *Electoral Studies*, vol. 32(2), pp. 360-369.

[8] Bouton, L. (2013). 'A theory of strategic voting in runoff elections', *American Economic Review*, vol. 103(4), pp. 1248-88.

[9] Bouton, L., & Castanheira, M. (2012). 'One person, many votes: Divided majority and information aggregation', *Econometrica*, vol. 80(1), pp. 43-87.

[10] Bouton, L., Castanheira, M., & Llorente-Saguer, A. (2016). 'Divided majority and information aggregation: Theory and experiment', *Journal of Public Economics*, vol. 134, pp. 114-128.

[11] Bouton, L., Castanheira, M., & Llorente-Saguer, A. (2017). 'Multicandidate elections: Aggregate uncertainty in the laboratory', *Games and economic behavior*, vol. 101, pp. 132-150.

[12] Bouton, L., & Gratton, G. (2015). 'Majority runoff elections: strategic voting and Duverger's hypothesis', *Theoretical Economics*, vol. 10(2), pp. 283-314.

[13] Bouton, L., & Ogden, B. G. (2021). 'Group-based voting in multicandidate elections', *Journal of Politics*, vol. 83(2), pp. 468-482.

[14] Brusco, S., Dziubiński, M., & Roy, J. (2012). 'The Hotelling–Downs model with runoff voting', *Games and Economic Behavior*, vol. 74(2), pp. 447-469.

[15] Buisseret, P. (2017). 'Electoral competition with entry under non-majoritarian run-off rules', *Games and Economic Behavior*, vol. 104, pp. 494-506.

[16] Callander, S. (2005). 'Duverger's hypothesis, the run-off rule, and electoral competition', *Political Analysis*, vol. 13(3), pp. 209-232.

[17] Castanheira, M. (2003). 'Why Vote for Losers?', *Journal of the European Economic Association*, vol. 1, pp. 1207-1238.

[18] Chin, M. (2019). 'When Do Politicians Appeal Broadly? The Economic Consequences of Electoral Rules in Brazil.', Working Paper, Harvard University.

[19] Colomer, J. M. (2007). 'On the origins of electoral systems and political parties: the role of elections in multi-member districts', *Electoral Studies*, vol. 26, pp. 262-273.

[20] Cox, G. W. (1997). Making votes count: strategic coordination in the world's electoral systems. Cambridge University Press.

[21] Dewan, T., and Myatt, D. (2007). 'Leading the Party: Coordination, Direction, and Communication', *American Political Science Review*, vol. 101 (4), pp. 825-843.

[22] Duverger, M. (1959). Political parties: Their organization and activity in the modern state. Methuen.

[23] Fey, M. (1997). 'Stability and coordination in Duverger's law: A formal model of preelection polls and strategic voting', *American Political Science Review*, vol. 91(1), pp. 135-147.

[24] Fischbacher, U. (2007). 'z-Tree: Zurich toolbox for ready-made economic experiments', *Experimental economics*, vol. 10(2), pp. 171-178.

[25] Forsythe, R., Myerson, R., Rietz, T., and Weber, R. (1993). 'An experiment on coordination in multi-candidate elections: The importance of polls and election histories', *Social Choice and Welfare*, vol. 10(3), pp. 223-247.

[26] Forsythe, R., Rietz, T., Myerson, R., and Weber, R. 1996. 'An experimental study of voting rules and polls in three-candidate elections', *International Journal of Game Theory*, vol. 25(3), pp. 355-383.

[27] Fujiwara, T. (2011). 'A regression discontinuity test of strategic voting and Duverger's Law', *Quarterly Journal of Political Science*, vol. 6(3–4), pp. 197-233.

[28] Lizzeri, A., and Persico, N. (2005). 'The Drawbacks of Electoral Competition', *Journal of the European Economic Association*, vol. 3(6), pp. 1318–1348.

[29] Martinelli, C. (2002). 'Simple plurality versus plurality runoff with privately informed voters', *Social Choice and Welfare*, vol. 19(4), pp. 901-919.

[30] McClintock, C. (2018). *Electoral Rules and Democracy in Latin America*, Oxford: Oxford University Press.

[31] Morton, R. B., and Rietz, T. A. (2007). 'Majority requirements and minority representation', *New York University Annual Survey of American Law*, vol. 63, p.691.

[32] Myatt, D. P. (2007). 'On the theory of strategic voting', *The Review of Economic Studies*, vol. 74(1), pp. 255-281.

[33] Myerson, R. (2002). 'Comparison of Scoring Rules in Poisson Voting Games', *Journal of Economic Theory*, vol. 103, pp. 219-251.

[34] Myerson, R., and Weber R. (1993). 'A Theory of Voting Equilibria', *American Political Science Review*, vol. 87, pp. 102-114.

[35] Osborne, M., and Slivinski A. (1996). 'A Model of Political Competition with Citizen-Candidates'. *Quarterly Journal of Economics*, vol. 111, pp. 65-96.

[36] Palfrey, T. R. (1989). 'A Mathematical Proof of Duverger's Law', in Models of Strategic Choice in Politics, ed. Peter C. Ordeshook. Ann Arbor: University of Michigan Press.

[37] Perez-Linan, A. (2006). 'Evaluating Presidential Runoff Elections', *Electoral Studies*, vol. 25, pp. 129-146.

[38] Piketty, T. (2000). Voting as communicating. *The Review of Economic Studies*, vol. 67(1), pp. 169-191.

[39] Pons, V., and Tricaud, C. (2018). 'Expressive voting and its cost: Evidence from runoffs with two or three candidates', *Econometrica*, vol. 86(5), pp. 1621-1649.

[40] Riker, W. H. (1982). 'The two-party system and Duverger's law: an essay on the history of political science', *American Political Science Review*, vol. 76(4), pp. 753-766.

[41] Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., & Webb, M. D. (2019). 'Fast and wild: Bootstrap inference in Stata using boottest', *The Stata Journal*, vol. 19(1), pp. 4-60.

[42] Saari, D. G. 2003. 'Unsettling aspects of voting theory', *Economic Theory*, vol. 22(3), pp. 529-555.

[43] Schuessler, A. (2000). A logic of expressive choice. Princeton University Press.

[44] Shin, E. (2019). 'A model of pre-electoral coalition formation', *Games and Economic Behavior*, vol. 118, pp. 463-485.

[45] Shugart, M. S., and Taagepera, R. (1994). 'Plurality versus majority election of presidents: A proposal for a "double complement rule"', *Comparative Political Studies*, vol. 27(3), pp. 323-348.

[46] Smith, J. H. 1973. 'Aggregation of preferences with variable electorate', *Econometrica*, vol. 41(6), pp. 1027-1041.

[47] Solow, B. (2018). 'Aggregate Uncertainty in Runoff Elections', Working Paper, University of Iowa.

[48] Stewart, C. III, and Ansolabehere, S. (2015). 'Waiting to vote', *Election Law Journal,* vol. 14(1), pp. 47-53.

[49] Van der Straeten, K., Laslier, J.-F., Sauger, N., and Blais, A. (2010). 'Strategic, sincere, and heuristic voting under four election rules: an experimental study', *Social Choice and Welfare*, vol. 35(3), pp. 435-472.

[50] Van der Straeten, K., Laslier, J.-F., and Blais, A. (2016). 'Patterns of strategic voting in run-off elections', in Voting Experiments (pp. 215-236). Springer.

[51] Wright, S. G., and Riker, W. H. (1989). 'Plurality and runoff systems and numbers of candidates', *Public Choice*, vol. 60(2), pp. 155-175.

Appendix

Appendix A

Proof of Proposition 1. Part (i): direct from Proposition 2 in Bouton, Castanheira, and Llorente-Saguer (2017).

Part (ii): Let us consider the case with an even number of voters (the proof is similar for an odd number of voters). We prove that $\sigma_t^R(A) = 1 \; \forall t \in \{t_A, t_B\}$ is an equilibrium. In that case, the only possible pivotal events under runoff (i.e., the only ones that are not zero probability events) are the threshold pivotability A/AC and C/AC. Indeed, when all majority voters are voting for A, and minority voters are voting for C, either one alternative obtains strictly more than $\frac{n}{2}$ with a lead of two votes or more, or the two alternatives tie. Thus, from the standpoint of a given voter, the only situation in which a ballot can change the final outcome is when either A is lacking one vote to win outright or C does. Majority voters of both types thus strictly prefer voting for A because that eliminates the risk of an upset victory of C in the second round. ■

Appendix B: Indifference Threshold with Risk Aversion

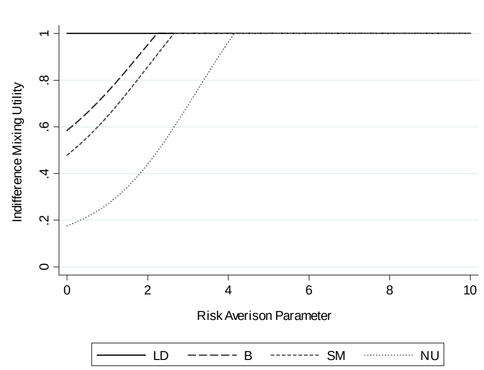


Figure B1: Probability of sincere voting of other t_B-voters that makes a t_B-voters indifferent between voting sincerely and voting strategically, as a function of the CRRA risk aversion parameter. We assume that t_A-voters and t_C-voters vote sincerely. The threshold equal to one indicates either indifference when all t_B-voters vote sincerely or that sincere voting is never a best response.

Appendix C: Regression Analyses

This section presents the basic regression analysis of the experiment. For this purpose, a panel dataset at the subject/period level was constructed. Random effects regression models are estimated for t_B-voters and the outcome of interest is whether the subject votes sincerely or not (i.e. votes for alternative B). In every case, standard errors are clustered at the group level. Throughout Tables C1-C4, we compare the effect of having the runoff rule on the probability of voting sincerely. Table C1 compares runoff and plurality under the Baseline configuration; Table C2 corresponds to the Low Disagreement case; Table C3 does the same but under the small minority configuration; while Table C4 compares runoff and plurality under the No Upset rule.

For each treatment configuration, six different models are estimated. In each table, column 1 corresponds to a model in which we do not control for subject-level covariates. Column 2 corresponds to models with covariates, which include subjects' gender, age, years in school, level of risk, level of trust, and dummies for experience participating in experiments and interest in politics. Column 3 includes a dummy indicating whether the minority candidate (C) won the previous election or not. Column 4 includes a count variable indicating the number of elections previously won by the minority candidate.[43] Column 5 includes the interaction between the runoff condition and the indicator of whether C won the previous election, to test for heterogeneous effects at such level. Column 6 does the same, but for the the number of previous elections won by C.

Tables C5 and C6 compare the different runoff configurations. Table C5 compares the runoff treatments across each other. Columns 1 and 2 compare the small minority and baseline conditions; columns 3 and 4 the low disagreement and baseline treatments; while columns 5 and 6 compare our small minority and low disagreement configurations. Table C6 compares each of these treatments with the No Upset configuration. Finally, Tables C7-C9 compare the election outcomes across the different treatment configurations.

[43] We also estimate models where we include period fixed-effects (results not shown). The results are robust to this alternative specification.

Table C1: Runoff Elections and B-Voters' Sincerity under Baseline

	(1) Vote B	(2) Vote B	(3) Vote B	(4) Vote B	(5) Vote B	(6) Vote B
Runoff	0.164*	0.180**	0.180**	0.152*	0.168*	0.209**
	(0.0964)	(0.0890)	(0.0891)	(0.0900)	(0.0872)	(0.0905)
Lag C Victory			-0.00526		-0.0288	
			(0.0210)		(0.0201)	
No. of C Victories				-0.0174***		-0.0145***
				(0.00288)		(0.00376)
Runoff×Lag C Victory					0.0435	
					(0.0385)	
Runoff×No. of C Victories						-0.00635
						(0.00505)
Constant	0.239***	-0.419	-0.418	-0.221	-0.412	-0.242
	(0.0700)	(0.337)	(0.337)	(0.354)	(0.339)	(0.363)
Subject-level Covariates	N	Y	Y	Y	Y	Y
N	1798	1798	1798	1798	1798	1798

Notes: Standard errors clustered at group level in parentheses. $Vote\ B$ indicates if the subject votes for candidate B or not. $Runoff$ indicates whether the treatment corresponds to R_B or P_B. $Lag\ C\ Victory$ indicates if C won the previous election. $No.\ of\ C\ Victories$ is the number of previous victories of C. Subject-level covariates include: gender, age, years in school, level of risk, level of trust, and dummies for experience participating in experiments and interest in politics. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table C2: Runoff Elections and B-Voters' Sincerity under Low Disagreement

	(1) Vote B	(2) Vote B	(3) Vote B	(4) Vote B	(5) Vote B	(6) Vote B
Runoff	-0.00597	0.0174	0.0173	0.0360	0.0492	0.0217
	(0.0568)	(0.0497)	(0.0502)	(0.0544)	(0.0564)	(0.0608)
Lag C Victory			0.000705		0.0596***	
			(0.0312)		(0.0171)	
No. of C Victories				-0.0152***		-0.0161***
				(0.00150)		(0.00233)
Runoff×Lag C Victory					-0.117**	
					(0.0481)	
Runoff×No. of C Victories						0.00173
						(0.00290)
Constant	0.293***	-0.562**	-0.562**	-0.510**	-0.568**	-0.505*
	(0.0317)	(0.256)	(0.256)	(0.257)	(0.259)	(0.258)
Subject-level Covariates	N	Y	Y	Y	Y	Y
N	977	977	977	977	977	977

Notes: Standard errors clustered at group level in parentheses. $Vote\ B$ indicates if the subject votes for candidate B or not. $Runoff$ indicates whether the treatment corresponds to R_LD or P_LD. $Lag\ C\ Victory$ indicates if C won the previous election. $No.\ of\ C\ Victories$ is the number of previous victories of C. Subject-level covariates include: gender, age, years in school, level of risk, level of trust, and dummies for experience participating in experiments and interest in politics. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table C3: Runoff Elections and B-Voters' Sincerity under Small Minority

	(1) Vote B	(2) Vote B	(3) Vote B	(4) Vote B	(5) Vote B	(6) Vote B
Runoff	0.226*	0.231**	0.232**	0.202*	0.231**	0.0876
	(0.128)	(0.105)	(0.105)	(0.106)	(0.109)	(0.102)
Lag C Victory			0.0270		0.0251	
			(0.0278)		(0.0492)	
No. of C Victories				-0.0160***		-0.0251***
				(0.00522)		(0.00593)
Runoff×Lag C Victory					0.00398	
					(0.0536)	
Runoff×No. of C Victories						0.0192**
						(0.00805)
Constant	0.325***	-1.145***	-1.146***	-1.068***	-1.146***	-0.995***
	(0.0769)	(0.186)	(0.186)	(0.186)	(0.186)	(0.181)
Subject-level Covariates	N	Y	Y	Y	Y	Y
N	1715	1715	1715	1715	1715	1715

Notes: Standard errors clustered at group level in parentheses. $Vote\ B$ indicates if the subject votes for candidate B or not. $Runoff$ indicates whether the treatment corresponds to R_SM or P_SM. $Lag\ C\ Victory$ indicates if C won the previous election. $No.\ of\ C\ Victories$ is the number of previous victories of C. Subject-level covariates include: gender, age, years in school, level of risk, level of trust, and dummies for experience participating in experiments and interest in politics. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table C4: Runoff Elections and B-Voters' Sincerity under No Upset

	(1) Vote B	(2) Vote B	(3) Vote B	(4) Vote B	(5) Vote B	(6) Vote B
Runoff	0.423***	0.485***	0.486***	0.448***	0.468***	0.352***
	(0.131)	(0.0941)	(0.0933)	(0.0942)	(0.0977)	(0.103)
Lag C Victory			0.00363		-0.0286	
			(0.0198)		(0.0203)	
No. of C Victories				-0.0109***		-0.0146***
				(0.00324)		(0.00378)
Runoff×Lag C Victory					0.0712**	
					(0.0329)	
Runoff×No. of C Victories						0.0132**
						(0.00554)
Constant	0.239***	-0.0885	-0.0893	0.0352	-0.0829	0.0529
	(0.0703)	(0.355)	(0.355)	(0.375)	(0.356)	(0.376)
Subject-level Covariates	N	Y	Y	Y	Y	Y
N	1661	1661	1661	1661	1661	1661

Notes: Standard errors clustered at group level in parentheses. $Vote\ B$ indicates if the subject votes for candidate B or not. $Runoff$ indicates whether the treatment corresponds to R_NU or P_B. $Lag\ C\ Victory$ indicates if C won the previous election. $No.\ of\ C\ Victories$ is the number of previous victories of C. Subject-level covariates include: gender, age, years in school, level of risk, level of trust, and dummies for experience participating in experiments and interest in politics. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table C5: Comparison of Runoff Treatments

	(1)	(2)	(3)	(4)	(5)	(6)
	Vote B	Vote B	Vote B	Vote B	Vote B	Vote B
Small Minority	0.148	0.140			0.264**	0.249***
	(0.121)	(0.108)			(0.112)	(0.0898)
Low Disagreement			-0.116	-0.122		
			(0.0812)	(0.0915)		
Constant	0.403***	-0.658	0.403***	-0.181	0.287***	-1.019***
	(0.0662)	(0.475)	(0.0663)	(0.455)	(0.0472)	(0.276)
Control Group	Baseline	Baseline	Baseline	Baseline	Low Dis	Low Dis
Covariates	N	Y	N	Y	N	Y
N	1846	1846	1458	1458	1332	1332

Notes: Standard errors clustered at group level in parentheses. *Vote B* indicates if the subject votes for candidate B or not. Subject-level covariates include: gender, age, years in school, level of risk, level of trust, and dummies for experience participating in experiments and interest in politics. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table C6: Comparison of Runoff Treatments vs. No Upset

	(1)	(2)	(3)	(4)	(5)	(6)
	Vote B	Vote B	Vote B	Vote B	Vote B	Vote B
Baseline	-0.258**	-0.279**				
	(0.129)	(0.118)				
Small Minority			-0.111	-0.107		
			(0.151)	(0.127)		
Low Disagreement					-0.375***	-0.392***
					(0.121)	(0.106)
Constant	0.662***	0.487	0.662***	-0.0948	0.662***	0.546
	(0.110)	(0.400)	(0.111)	(0.464)	(0.111)	(0.478)
Control Group	R_NU	R_NU	R_NU	R_NU	R_NU	R_NU
Covariates	N	Y	N	Y	N	Y
N	1835	1835	1709	1709	1321	1321

Notes: Standard errors clustered at group level in parentheses. *Vote B* indicates if the subject votes for candidate B or not. Subject-level covariates include: gender, age, years in school, level of risk, level of trust, and dummies for experience participating in experiments and interest in politics. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table C7: Runoff Condition and Election Outcomes (Baseline and Low Disagreement)

	(1)	(2)	(3)	(4)	(5)	(6)
	A Victory	B Victory	C Victory	A Victory	B Victory	C Victory
Baseline	0.0433	0.0111**	-0.0544*			
	(0.0334)	(0.00532)	(0.0328)			
Low Disagreement				-0.0267	.	0.0267
				(0.0306)	.	(0.0306)
Control Group	P_B	P_B	P_B	P_LD	P_LD	P_LD
N	660	660	660	600	600	600

Notes: Standard errors clustered at group level in parentheses. *A Victory*, *B Victory*, and *C Victory* indicate if A, B, or C win the election. *Baseline* indicates if the treatment is *R_B*. *Low Disagreement* indicates if the treatment is *R_LD*. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table C8: Election Outcomes in Small Minority and No Upset vs. Runoff Baseline

	(1) A Victory	(2) B Victory	(3) C Victory	(4) A Victory	(5) B Victory	(6) C Victory
Small Minority	0.0567	0.0356	-0.0922***			
	(0.0424)	(0.0245)	(0.0287)			
No Upset				-0.0533	0.119**	-0.0656*
				(0.0592)	(0.0530)	(0.0346)
Control Group	R_B	R_B	R_B	R_B	R_B	R_B
N	660	660	660	660	660	660

Notes: Standard errors clustered at group level in parentheses. *A Victory*, *B Victory*, and *C Victory* indicate if *A*, *B*, or *C* win the election. *Small Minority* indicates if the treatment is *R_SM*. *No Upset* indicates if the treatment is *R_NU*. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Table C9: Runoff Condition and Election Outcomes (Small Minority and No Upset)

	(1) A Victory	(2) B Victory	(3) C Victory	(4) A Victory	(5) B Victory	(6) C Victory
Small Minority	-0.00667	0.0433*	-0.0367			
	(0.0436)	(0.0243)	(0.0300)			
No Upset				-0.01000	0.130**	-0.120***
				(0.0623)	(0.0530)	(0.0398)
Control Group	P_SM	P_SM	P_SM	P_B	P_B	P_B
N	600	600	600	600	600	600

Notes: Standard errors clustered at group level in parentheses. *A Victory*, *B Victory*, and *C Victory* indicate if *A*, *B*, or *C* win the election. *Small Minority* indicates if the treatment is *R_SM*. *No Upset* indicates if the treatment is *R_NU*. * is significant at the 10% level, ** is significant at the 5% level, *** is significant at the 1% level.

Appendix D: Utilities of t_B voters

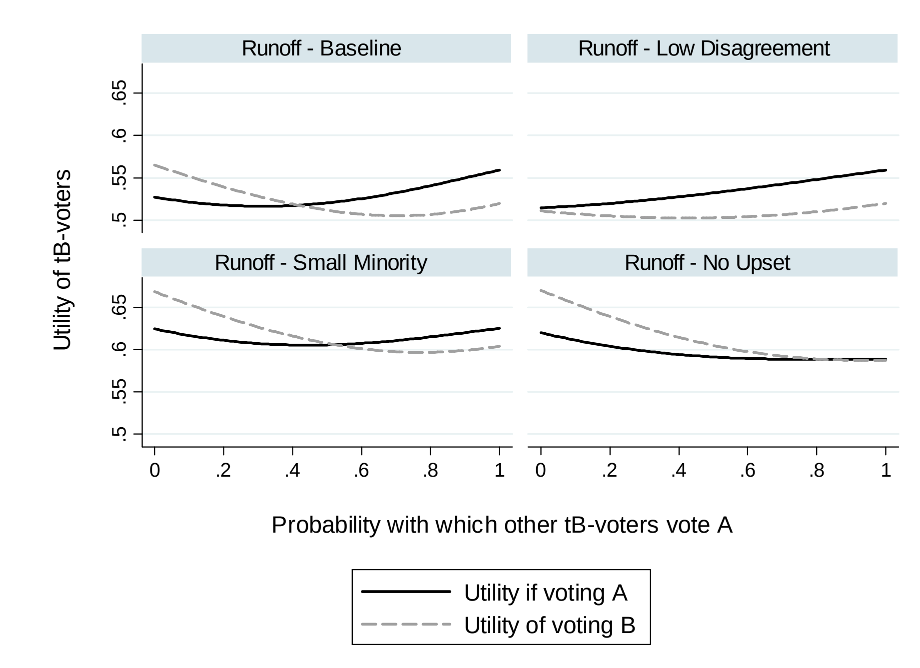


Figure D1: Utilities of t_B-voters depending on the action they take and the mixing probabilities of the other t_B-voters. To compute these utilities, we assume that t_A and t_C voters vote their preferred option with probability one.

Appendix E: Individual Characteristics Figure for Additional Treatments

(i) Always Duvergerian (ii) Always Sincere

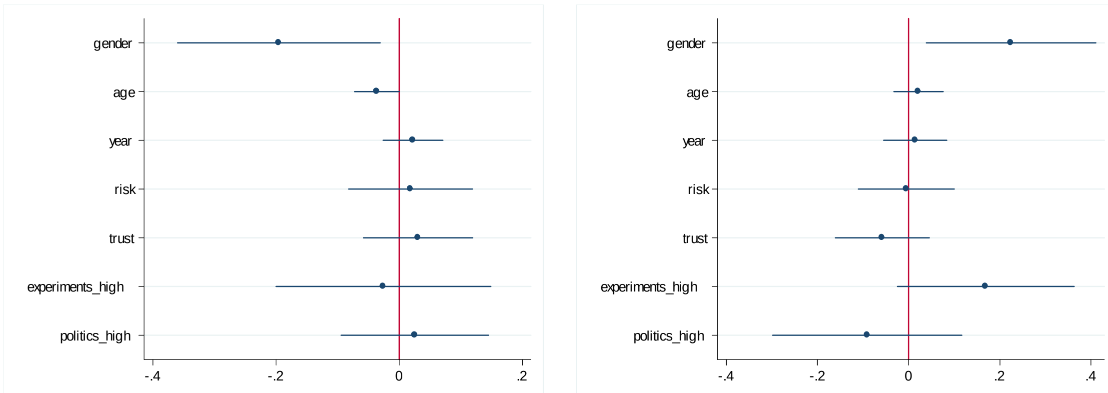


Figure E1: Individual characteristics and "always Duverger" and "always sincere" behavior.

Appendix F: Questionnaire Data

Variable	Description
Gender	Female $= 1$; Male $= 0$.
Age	Age in years.
Year	Years of studies.
Risk	Tendency to take risks. Likert scalefrom 1 to 5.
Trust	Tendency to trust people. Likert scale from 1 to 5.
Experiments	$= 1$ if the subject has participated in 4 or more experiments. Originally, this was a categorical variable about participation in previous experiments: “Never”, “1-3”, “4-6”, and “More than 6”.
Politics	$= 1$ if politics is very important or extremely important in the subject’s life. Originally Likert scale from 1 to 4.
Religiosity	Degree of religiosity. Likert scale from 1 to 4. Not included in the regressions.
Religion	Categorical variable: Christian (22.69%), Hinduist (41.44%), Judaism (2.55%), Muslim (4.17%), No religion (18.98%), Other Religion (5.79%), Prefer not to answer (4.40%). Not included in the regressions.
Major	Categorical variable with the options: “Law" (0.93%), “Economics" (12.96%), “Literature" (0.69%), “Physics/Chemistry/Biology" (5.79%), “Engineering" (37.50%), “History" (0.93%), “Politics" (3.01%), “Mathematics" (0.28%), “Others" (34.49%). Not included in the regressions.
Party	Categorical variable: Democrat (43.75%), Republican (7.18%), Other (10.65%), NA (38.43%). Not included in the regressions.
Siblings	Number of siblings. Not included in the regressions.

Table F1: Description of variables in the questionnaire data.

Appendix G: The Nature of the Second Round

One open question is what was the effect of a computerized second round, and how results would differ from an experiment in which decisions in the second round were taken by participants. In order to shed some light on this question, we use the data from a pilot (run on 2013), in which the second round was not computerized. The qualitative findings are similar to the ones reported in the experiment. But the challenge is how to quantitatively compare the two sets of data given that the setups differ in a number of ways. In order to overcome this challenge, we propose a measure of the incentives to vote strategically in our setting that allows for comparability across treatments. We then show that there is no significant difference in the level of strategic voting between computerized and non-computerized second rounds.

We structure this appendix in four parts. First, we present the experimental design of the pilot. Second, we show aggregate data of the pilot. Third, we propose the measure of incentives to vote strategically. And finally, we report the estimated differences.

The pilot design. The main difference in the designs of the pilot and the experiment was the fact that in the runoff sessions of the pilot, there were two groups: one that would vote in the first round, and another one that would vote in a second round (if it was reached). The other differences were that in the pilot: (i) abstention was allowed; (ii) the payoffs of the expected majority voters were 12, 10 and 2; (iii) the distribution of types t_A, t_B, and t_C, was 0.32, 0.22, and 0.46 respectively; (iv) the group size was $n = 9$; (v) the threshold for outright victory in the first round of the runoff election was 45%.[44] Under these parameters, sincere voting is an equilibrium only under runoff, while the Duvergerian equilibria exist under both systems. Overall, we have 4 pairs of groups under runoff (four groups of nine subjects voting in the first round, and four groups of nine subjects (potentially) voting in the second round), and two groups on nine under plurality.

The pilot data. The aggregate behaviour under plurality and the first round of the runoff elections in the pilot is reported in Table G1. There are some interesting patterns. First, abstention is minimal both across types and voting rules. This is in fact the reason

[44] Note that if nobody abstains, a 45% threshold is equivalent to a 50% threshold in the case under consideration because strictly passing the threshold implies obtaining 5 (out of 9) votes in both cases.

		% A	% B	% C	% Abst
Plurality	t_A	97.4	0.58	1.73	0.29
	t_B	72.8	23.3	3.88	0
	t_C	4.38	0.2	95	0.4
Runoff	t_A	97.1	2.17	0.58	0.14
(first round)	t_B	49.8	49.4	0.79	0
	t_C	1.04	0.21	98.6	0.21

Table G1: Aggregate behavior in the first round in the pilot.

	A vs B			A vs C			B vs C		
	% A	% B	% Abst	% A	% C	% Abst	% B	% C	% Abst
t_A	100	0	0	98.7	1.31	0	97	3.03	0
t_B	5.71	94.3	0	92.9	3.06	4.08	93.8	3.13	3.13
t_C	19.3	22.8	57.9	3.02	97	0	1.64	98.4	0

Table G2: Aggregate behavior in the second round in the pilot.

why we did not allow for abstention in the experiments reported in the paper. Second, and in line with the results of our experiments, t_A and t_C voters overwhelmingly vote for their preferred alternative. In contrast, t_B voters vote strategically with high frequency, and this frequency is higher under plurality. This is also consistent with our findings in the experiments reported in the paper.

Table G2 represents the voting behaviour in the second round under runoff (whenever it was reached). The table shows two clear patterns. First, when voters have a strict preference for one of the two candidates participating in the second round, they overwhelmingly vote for their preferred candidate. This is, in fact, their dominant strategy. Second, when the two majoritarian parties go to the second round, t_C voters tend to abstain, but they vote for alternatives A and B with relatively high probabilities (19.3% and 22.8% respectively). This departure from the hypothesized behaviour does not have a significant effect on the probability that A wins in a second round between A and B. The conditional probability that A wins when voters vote as hypothesized is equal to 65.19% while it is equal to 64.85% when we use the frequencies of vote observed in the pilot.

The measure of strategic incentives. In order to compare the differential incentives to vote strategically across treatments, and in particular, between the pilot and the different treatments in the experiment, we generated a measure of strategic incentives. Given that t_A and t_C voters overwhelmingly vote for their preferred candidates, we focus on the

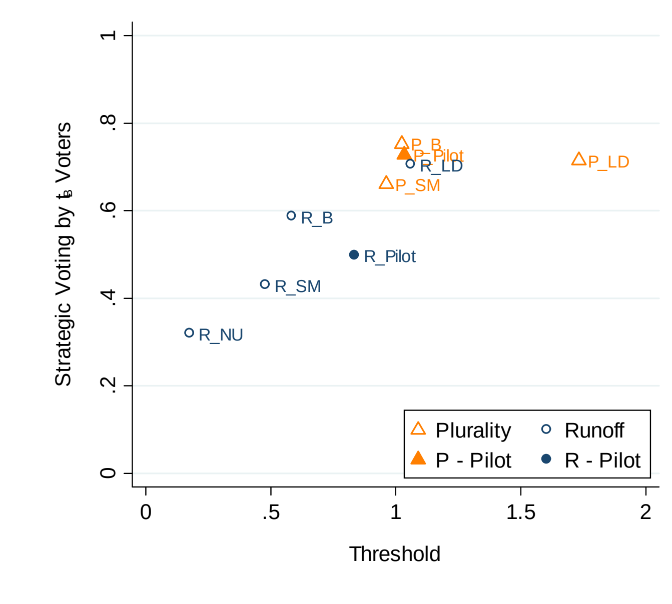


Figure G1: Thresholds that define the incentives to vote strategically and actual strategic voting in each different treatment.

strategic behaviour of t_B voters. To build our measure, we start by assuming that t_A and t_C voters vote for their favorite candidates and that t_B voters mix between voting sincerely for B and voting strategically for A. The measure that we use is the mixing probability of sincere voting among t_B voters that makes a t_B voter indifferent between voting for B and voting for A. This threshold determines the best responses of t_B voters. In treatment R_B, for instance, this threshold is equal to 58.45%. This means that if other t_B voters vote sincerely for B with a probability higher (lower) than 58.45%, the unique best response is to vote sincerely for B (strategically for A). We allow this measure to go beyond 100%, but the results are not substantially different if we impose a bound at 100%. The interpretation of a measure of 120% would be that in order to be indifferent between voting sincerely and voting strategically, a t_B voter would need all t_B voters to vote sincerely and enough t_A voters voting for B, to add an extra 20%.

Table G3 reports the thresholds for all treatments, including the ones from the pilot. We see that, according to our measure, incentives to coordinate in the pilot are stronger than in the Baseline, Small Minority, and No Upset treatments, but weaker than in the Low Disagreement treatment. Figure G1 shows that this measure of strategic incentives (horizontal axis) correlates well with the observed amount of strategic voting by t_B voters

	Plurality	Runoff
P_B	1.0236	0.5845
P_LD	1.7322	1.0607
P_SM	0.9616	0.4786
R_NU	0.1756	-
Pilot	1.034	0.8354

Table G3: Thresholds in the different treatments.

	(1) Vote for A	(2) Vote for A
Threshold	0.433***	0.224**
	(0.125)	(0.0896)
Pilot	-0.123	0.0759
	(0.139)	(0.0538)
Runoff		-0.0461
		(0.0776)
Pilot * Runoff		-0.144
		(0.140)
Constant	0.267***	0.432***
	(0.100)	(0.123)
N	3671	6075
Data	Only Runoff	All Data

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table G4: Regressions on the likelihood that t_B voters vote strategically for alternative A.

(vertical axis). Therefore, this measure does not only give a measure of the theoretical incentives to coordinate, but it has some explanatory power in terms of observed behaviour.

Comparing the data. We hypothesized that if the second round was not computerized, voters would have higher incentives to vote strategically in order to save time. In order to test this, we estimate a random effects linear probability model of the probability that a t_B-voter vote strategically for alternative A on the value of the threshold, the dummy variables Pilot and Runoff and their interaction. We cluster standard errors at the matching group level. We use two different specifications. In the first one, we only consider observations under runoff elections (see column (1) in Table G4). In the second one, we take a differences in differences approach (see column (2) in Table G4) and use all the data of the pilot and the experiment. Both specifications lead to the same conclusions. First, the coefficient of the threshold is positive and significantly different from zero. Second, none of the coefficients of the other variables (except for the constant) are significantly different from zero. In particular, the different level of strategic voting in the pilot compared to the experiment (controlling for the different incentives) is negative but not significantly different from zero. In the first specification, this is measured by the coefficient on the variable Pilot, which has a value of -0.12 and a p-value = 0.377. In the second, it is measured by the sum of coefficients on Pilot and Pilot*Runoff, which has a value of -0.0681 and a p-value = 0.625.

Appendix H: Instructions

Below are the instructions for the Baseline treatments. Instructions for the other treatments were exactly the same except for parameters in blue, which determine probabilities of types and winners in the second round.

General Indications. Hello and welcome to the experiment. We appreciate your time and disposition for participating in this academic exercise. We would like to remind you that your answers and choices are anonymous, confidential, and that they will be used strictly for academic purposes. Also, remember that in addition to the show-up fee, at the end of the experiment you will be paid according to your performance in the experiment. Throughout the exercise we will use dollars to express the benefits you obtain from your decisions.

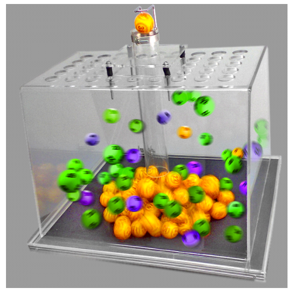
Please feel free to ask any questions about the experiment's instructions before we start. Remember that during the experiment you cannot talk or communicate with any other subject in this room, unless otherwise stated. In your desk you will find a consent form and a pen. Please read and sign the consent form before we begin. You will also find a piece of paper, which can be used during the experiment in case you need to do any calculations. Any other things, such as cell phones or other devices, cannot be used during the experiment.

Experiment's Instructions. The experiment is composed of 60 periods. The rules are the same for all periods. At the beginning of the experiment, you will be randomly assigned to a group of 12 (including yourself). However, you will not know the identity of them, and they will not know your identity. You will belong to the same group throughout the whole experiment.

Your earnings will depend partly on your decisions, partly on the decisions of the other participants in your group and partly on chance. After each period, you will earn certain amount of dollars. At the end of the experiment, the computer will randomly select three periods, and you will earn the sum of dollars you get in these three periods. Each of the 60 periods has the same chance of being selected. Additionally, you will get a show-up fee of $7.

Your Role. At the beginning of every period, the first thing the computer will do is to randomly assign you a role out of three possibilities: you can be a green, a purple, or a yellow voter. Allocation of roles is as follows: before each period starts, you will see the urn depicted in the next figure.

This urn contains 100 balls: 34 are green, 22 are purple, and 44 are yellow. After you click OK, one ball will be selected at random by the computer. If it is a green ball, you will be a green voter; if it is purple you will be a purple voter; and if it is yellow you will be a yellow voter. **Therefore, you have a 34 percent chance of being a green voter; 22 percent chance of being a purple voter; and 44 percent chance of being a yellow voter**.



Your decision. After your role is assigned, an election is held between three candidates: the green, the purple, and the yellow candidates. You have to decide for which candidate you cast your vote. That is, **you have to decide whether to vote for green, vote for purple, or vote for yellow**.

Winning Candidate [If Runoff]. When all participants have taken their decision, the votes of all participants will be added up. The winning candidate will depend on the final amount of votes that each color receives.

• If the candidate who obtains the largest number of votes in the first round receives 7 votes or more, this candidate is declared the winner.

• If the candidate with the largest number of votes in the first round obtains less than 7 votes, then the two top vote receivers face each other in a runoff election, which will be called the second round. If two candidates are tied in the second place, the tie is broken by alphabetical order. Similarly, if there is a triple tie, green and purple go to the second round.

The winner of the second round is randomly chosen by the computer by drawing a ball from an urn that contains 100 balls. The winning candidate is the color of the random ball.

• If the **green** and the **purple** candidates face in the second round, the urn contains **78** **green** balls and **22** **purple** balls.

• If the **green** and the **yellow** candidates face in the second round, the urn contains **76** **green** balls and **24** **yellow** balls.

• If the **purple** and the **yellow** candidates face in the second round, the urn contains **76** **purple** balls and **24** **yellow** balls.

Winning Candidate [If Plurality]. When all participants have taken their decision, the votes of all participants will be added up. The candidate who obtains the largest number of votes is declared the winner. If there is a tie for the first place, the tie is broken by alphabetical order.

Your payoff. Your final payoff for each period depends on your role and on who wins the election in that period. The next table summarizes the payoff that you get depending on your role and on the winning candidate:

Winning Candidate

		Green	Purple	Yellow
	Green	11	8	1
Your Role	Purple	8	11	1
	Yellow	1	1	11

The top part of the table indicates the winner of the election and the left part of the table indicates your role.

- If you are a green voter and the winning candidate is the green one, your payoff is $11.
- ____________________________________ purple __, your payoff is $8.
- ____________________________________ yellow __, your payoff is $1.
- If you are a purple voter and the winning candidate is the green one, your payoff is $8.
- ____________________________________ purple __, your payoff is $11.
- ____________________________________ yellow __, your payoff is $1.
- If you are a yellow voter and the winning candidate is the green one, your payoff is $1.
- ____________________________________ purple __, your payoff is $1.
- ____________________________________ yellow __, your payoff is $11.

Information at the end of each Round

Once you and all the other participants have made your choices, the winning candidate will be determined and the period will be over. At the end of each round, you will receive information about the vote shares of candidates on every previous period, whether a second round was necessary or not, and the winning candidate.

Control Questions. Before starting the experiment, you will have to answer some control questions in the computer terminal. Click Ok after you have answered all the questions of a page. Please feel free to ask any questions about the experiment's instructions. Once you and all the other participants have answered all the control questions, the experiment will start.