How Do People Vote Under Instant Runoff Voting? An Experiment on Complexity and Voting Behavior*

Arnaud Dellis, Université du Québec à Montréal†

Sabine Kröger, Université Laval‡

June 2024

Abstract: Instant Runoff Voting is a voting procedure that requires voters to rank candidates. It is currently used in several countries (Australia, USA, Canada), and various places worldwide are debating its adoption for political elections. A primary argument in support of its adoption is that it is such a complex voting procedure that people would be unable to vote strategically and would then choose to resort to voting sincerely. We conducted a laboratory experiment to assess the validity of this claim. More generally, we investigate how complexity affects voting behavior. Our findings confirm that complexity does impede strategic voting. However, we also observe that rather than resorting to sincere voting, voters tend to respond to complexity by adopting a voting heuristic, which we call *Lifting*, that consists of reversing the ranks of their two most preferred candidates. Additionally, we find that the complexity of Instant Runoff Voting adversely affects voters and makes it more difficult for them to learn from experience.

JEL codes: C23, C72, C91, C92, D72, D91.

Keywords: Strategic Voting; Sincere Voting; Complexity; Instant Runoff Voting; Borda Count; Positional Voting Rules; Voting Experiment.

* We thank Yacine Yaddaden and Philippe Chabot for their valuable research assistance. We are grateful to Garret Binding, Sebastian Ebert, Anastasia Leontiou, Pierre-Guillaume Méon, and Stefan Trautmann for their comments and suggestions. We also thank seminar and conference participants at the ESA global conference, Meeting of the SCSE, C-DEM/CSDC Forum on Democratic Citizenship, World Congress of the IAPSS, EPSA conference, PEDD conference at Universität Münster, ETH Workshop on Democracy, Universität Würzburg, the Berlin Behavioral Economics Seminar, Florida State University, Carleton University, McMaster University, Universität Heidelberg, and City-University of London. We acknowledge financial support from Canada's Social Sciences and Humanities Research Council (SSHRC). Arnaud Dellis also thanks faculty members at the Alfred Weber Institute of Universität Heidelberg for their hospitality.

† dellis.arnaud@uqam.ca; ‡ sabine.kroger@ecn.ulaval.ca

"*[Instant Runoff] voting takes the simple concept of plurality voting – each voter selects his preferred candidate, the candidate with the most votes wins – and adds complexity*" Saltsman and Paxton (2021)

1. Introduction

Instant Runoff Voting (henceforth, *IRV*) is receiving a lot of attention in the public electoral reform debate.[1] Several advocacy groups, such as *FairVote* in the United States and *the Electoral Reform Society* in Great Britain, are actively campaigning for its adoption in political elections. Political parties such as the Liberal Democrats in Great Britain and the Liberal Party of Canada have championed the adoption of IRV for legislative elections. While advocacy for the adoption of IRV is underway in multiple places, this voting procedure is already utilized in various locations worldwide. For example, Australia has been employing IRV for its parliamentary elections since 1918, and Ireland employs it for electing its president. In the USA, some states and cities, including Maine, Alaska, San Francisco, and New York City, have adopted IRV for federal and local elections, and several political parties in Canada employ IRV for selecting their leader. IRV is also used outside the political arena (e.g., to designate the winner of the Academy Award for Best Picture).

The primary appeal of IRV is to allow voters to rank candidates rather than restrict them to vote for only one candidate. Indeed, under IRV, voters rank candidates, and the candidate who is ranked first on a majority of ballots is elected. If no candidate achieves a majority of first-place rankings, then the candidate with the fewest first-place rankings is eliminated from every ballot and the next-ranked candidate takes the place of the eliminated candidate. This elimination process is iterated until one of the remaining candidates secures a majority of first-place rankings.

In the public electoral reform debate, a prominent argument in support of IRV is that it discourages people from voting strategically, and induces them to vote sincerely, meaning to reveal their true preference ranking of the candidates. Quoting *FairVote*, an advocacy group in the United States, "*voters can sincerely rank candidates in order of preference. Voters know that if their first choice doesn't win, their vote automatically counts for their next choice instead. This frees voters from worrying about how others will vote and which candidates are more or less likely to win.*"[2] Similarly, the *Electoral Reform Society* in Great Britain states that the adoption of IRV will result in "*no more tactical voting. Supporters of parties large and small can vote sincerely for their preferred party in the knowledge their vote can still help decide the winner.*"[3] In our paper, we verify the validity of this claim. Are people indeed deterred from voting strategically under IRV?[4] If so, do they choose to resort to sincere voting, as claimed? Or

[1] IRV is also referred to as Ranked-Choice Voting (RCV), Preferential Voting, the Alternative Vote, Single Transferable Vote, and the Hare method of voting.

[2] Retrieved from https://www.fairvote.org on June 6th, 2024.

[3] https://www.electoral-reform.org.uk/wp-content/uploads/2017/06/introducing-the-alternative-vote.pdf.

[4] We adopt the game-theoretic definition of strategic voting, which characterizes strategic voting as best-responding to the

do they adopt other voting heuristics and, if so, which ones? A key novelty of our paper lies in controlling voters' preferences together with their beliefs about others' voting choices. This approach brings novel insights into IRV and, more generally, voting behavior by enabling us to 1) identify when a vote is consistent with strategic voting, 2) manipulate voters' beliefs to encompass various forms of strategic voting behavior, and 3) identify heuristics that voters choose to adopt.

Some scholars have put forth various arguments in favor of a voting procedure that deters strategic voting behavior and encourages sincere voting. One argument is associated with elections' goal of aggregating individual preferences so that collective choices reflect the general will. For this goal to be met, voters must arguably cast ballots that reveal their true preferences for the various candidates. As Riker (1981, p.110) states: "*Even if a society agrees on a method of voting and even if it produces a coherent outcome, we still do not know whether this outcome truly reflects the values of the voters or whether it is the result of some kind of manipulation.*" A voting procedure that induces people to vote sincerely increases the likelihood that the outcome of the vote reflects people's preferences. Another argument in favor of a voting procedure that discourages strategic voting is to prevent special interests from manipulating the outcome of the vote by convincing voters that some candidates are trailing and, therefore, that voters should not waste their votes on these candidates. A voting procedure that induces sincere voting insulates the election outcome from such manipulations.[5] One more argument in favor of a voting procedure that discourages strategic voting behavior is based on empirical evidence showing that some groups of voters (e.g., wealthy people) are better able than others to behave strategically and 'manipulate' the outcome of the vote to their advantage (Eggers and Vivyan, 2020; Loewen et al., 2015). For this reason, a voting procedure that induces sincere voting behavior is arguably more equitable.

However, the claim that IRV prevents people from voting strategically is at odds with the Gibbard-Satterthwaite theorem (Gibbard, 1973; Satterthwaite, 1975). This theorem establishes that, in voting situations with three or more candidates, any non-dictatorial voting procedure, including IRV, is subject to (non-sincere) strategic voting behavior.[6] As Cox (1997, p.93) writes "*...it would be erroneous to conclude, as is sometimes hinted in the literature, that [IRV] produces no incentives to vote strategically. This conclusion would of course run afoul of the Gibbard-Satterthwaite Theorem's general guarantee that any democratic voting procedure can generate incentives to vote strategically.*"

believed strategy profile of the other voters, meaning voting in a way that maximizes the voter's expected payoff given her beliefs about the voting strategies of the other voters. With this definition, strategic voting can involve casting a non-sincere ballot (*non-sincere strategic voting*) or a sincere ballot (*sincere strategic voting*).

[5] Aligned with this argument, Myerson (1993a) suggests that a voting procedure is more effective at fighting corruption if it induces sincere voting behavior. In the same vein, strategic voting can erect entry barriers to new candidates who, despite being preferred by a majority of voters, may choose not to run because of the fear that voters will believe those new candidates have no chance of winning and, therefore, will not want to waste their vote on them (see, e.g., Bol et al., 2016).

[6] For instances of the forms that strategic voting can take under IRV, see Laslier (2016).

An argument has been put forward to reconcile the Gibbard-Satterthwaite theorem with the claim that IRV prevents people from voting strategically. The argument relies on the complexity of IRV and the assumption of unbounded cognitive abilities underlying the Gibbard-Satterthwaite theorem. Specifically, while the Gibbard-Satterthwaite theorem establishes the possibility of strategic voting under IRV, in practice, voters' bounded cognitive abilities make it difficult for them to vote strategically under such a complex voting procedure. Quoting Farrell and McAllister (2006, p.126-7): "*The general consensus is that the scope for strategic voting declines as an electoral system becomes more complex ... the reason being that the computational and information requirements in more complex systems make it all but impossible for voters to act in a strategic manner.*" The argument goes on by suggesting that people respond to the complexity of voting strategically under IRV by choosing to resort to sincere voting.[7]

The complexity of voting strategically under IRV takes several forms. Firstly, there is the *computational complexity*, which refers to the difficulty a voter faces in identifying an optimal ballot, meaning a ballot that maximizes the voter's expected payoff given her beliefs about the voting decisions of the other voters (Bartholdi and Orlin, 1991). In the case of IRV, this form of complexity is primarily associated with the iterated elimination process, which requires multiple steps for the computation of an optimal ballot.[8] Secondly, there is the *strategic complexity*, which refers to the difficulty of accurately predicting the voting behavior of the other voters or, in game-theoretic parlance, the strategy profile to which the voter has to best-respond. As Cox (1997, p.94) writes, "*voters need more information in order to cast a strategic vote under [IRV] than under ordinary plurality.*" In our paper, we isolate computational complexity from strategic complexity and study the effect of the former on voting behavior.

We designed and conducted a laboratory experiment to explore how the computational complexity of IRV affects voters' behavior, whether it prevents people from voting strategically, and whether it induces them to vote sincerely.[9] To separate computational complexity from strategic complexity, we adopt an

[7] In addition to the complexity of voting strategically, IRV offers fewer possibilities (Chamberlin, 1985; Chamberlin et al., 1984) and weaker incentives (Eggers and Nowacki, 2024) for non-sincere strategic voting behavior compared to other voting procedures such as Plurality Voting.

[8] The computational complexity of IRV is further compounded by the violation of the *monotonicity* property by IRV (Doron and Kronick, 1977), which signifies that ranking higher a candidate may hurt his winning prospects and, vice versa, ranking lower a candidate may improve his winning prospects. The violation of the monotonicity property by IRV complicates the search for an optimal ballot as a voter who seeks to improve the winning prospects of a candidate must consider not only moving this candidate to a higher rank, as it is sufficient to do under voting procedures satisfying the monotonicity property (like Plurality Voting), but must also consider moving this candidate to a lower rank.

[9] The experimental laboratory methodology exhibits several advantages for studying how complexity affects voters' behavior. Firstly, it allows varying the complexity of the voting procedure used in a voting situation and see how this affects the voting behavior of the same voters. Secondly, this methodology allows us to endow people with specific preferences and beliefs, which enables us to better identify when a person casts an optimal ballot, votes sincerely, or adopts another voting heuristic.

approach where a voter (a participant in the experiment) observes the votes of the other, computerized, voters before making his own voting decision. This means that, instead of eliciting participants' beliefs about others' voting decisions, we induce those beliefs by letting participants know the votes of the other, computerized, voters.[10] Moreover, to capture various forms of optimal voting behavior, we provide participants with a series of predefined vote profiles for the computerized voters. We present the same vote profile twice consecutively to allow for learning. To further give participants opportunities to learn, we provide them with information about the outcome of their vote once they have submitted their ballot.

Our experimental design captures complexity in two ways. Firstly, we vary the number of candidates available for selection, what we call *environmental complexity*. Some voting situations involve three candidates, while others, more computationally complex, involve four candidates. Secondly, every participant casts ballots under two voting procedures: IRV and the Borda Count (henceforth, *Borda*), what we call *procedural complexity*.[11] Under both IRV and Borda, voters rank candidates.[12] The difference between the two voting procedures lies in their allocation rules, which determine the election winner from voters' ballots.[13] In the case of Borda, a voter gives each candidate a number of votes that is equal to the number of candidates she ranks below.[14] The election winner is the candidate with the highest vote count. Hence, while Borda and IRV share the same ballot structure, their different allocation rules imply that voting strategically is less complex under Borda than under IRV (Conitzer and Walsh, 2016; Durand, 2023).[15,16] Quoting Black (1976, p.15): "*Even for the unsophisticated voter the Borda count is an invitation to strategic voting.*"

It is important to highlight that our design separates complexity from confusion. This is done by requiring that participants answer every question of a comprehension test on the voting procedure correctly

[10] Esponda and Vespa (2014) use a similar approach to distinguish between information extraction and hypothetical thinking.

[11] Borda has received considerable attention from scholars and is used to determine the winners of various awards, including the annual Eurosong contest, the MLB's Most Valuable Player, and the Heisman Trophy for college football in the USA.

[12] Considering two voting procedures that share the same ballot structure allows us to separate the complexity of the voting procedure from the complexity of expressing preferences for multiple candidates, something which would not be feasible if we were to compare IRV with Plurality Voting.

[13] Rae (1967) defines a voting procedure as the combination of three components: district magnitude, ballot structure, and allocation rule. The *district magnitude* specifies the number of candidates to elect, which in our experiment is equal to one. The *ballot structure* specifies the set of admissible ballots, meaning all the ballots that a voter is allowed to cast, which in our experiment is a ranking of the candidates. The *allocation rule* specifies how voters' ballots are aggregated to determine the election winner.

[14] For example, in a voting situation with three candidates, a voter gives two votes to the candidate she ranks first (since she ranks the two other candidates below), one vote to the candidate she ranks second (since she ranks only one candidate below), and no vote to the candidate she ranks last.

[15] In particular, unlike IRV, Borda satisfies the *monotonicity* property, meaning that a voter does not hurt the winning prospects of a candidate by ranking him or her higher on his ballot.

[16] Anecdotally, when Pierre-Simon Laplace pointed out to Jean-Charles de Borda that his voting procedure encourages (non-sincere) strategic voting behavior, Borda reportedly responded that his voting procedure is intended only for honest men.

before casting votes. This enables us to remove confusion as a potential explanation for our findings and concentrate on the impact of complexity.[17]

Comparing participants' voting behavior under IRV and Borda, we find a lower frequency of optimal ballots, meaning ballots that are consistent with strategic voting, cast under IRV compared to Borda, with the casting of an optimal ballot to be less frequently the result of strategic voting behavior under IRV than under Borda. The lower frequency of optimal ballots under IRV is accompanied by a higher frequency of sincere votes. We also observe that participants with stronger numeracy skills tend to be better able to cast an optimal ballot under Borda. Numeracy skills have no effect under IRV, a voting procedure for which strategic voting seems to require greater cognitive abilities. For example, we observe that participants with a graduate degree are better able to cast an optimal ballot under IRV. Moreover, we find that the probability with which a participant cast an optimal ballot increases under Borda with 1) the amount of time that the participant devoted to making a voting decision, as well as 2) the participant's experience. We find no such relations under IRV. All these findings provide support for the claim that the complexity of IRV makes it difficult for people to vote strategically. Similarly, we observe that an increase in environmental complexity, induced by an increase in the number of candidates, decreases the frequency with which participants cast an optimal ballot. Voters resort to sincere voting and other heuristics when procedural complexity increases in elections involving three candidates and when environmental complexity increases under the simple Borda voting procedure. When procedural complexity increases within an already complex environment or vice versa when environmental complexity increases under the more complex IRV procedure, voters tend to turn towards other heuristics than sincere voting.

Nevertheless, the support for the claim that IRV discourages strategic voting and induces people to vote sincerely requires several qualifications. Firstly, we observe in the experiment that only a few of the votes cast under IRV are sincere. Interestingly, participants responded to the complexity of IRV more often by adopting voting heuristics other than sincere voting. Among the other voting heuristics that participants used under IRV, two stick out. In one heuristic, which we refer to as *Lifting*, voters invert the ranking of their top two preferred candidates. In the other heuristic, which we refer to as *Conforming*, voters submit a ballot similar to those of the computerized voters. The adoption of the *Lifting* heuristic is related to complexity, being used more frequently under IRV than under Borda. By contrast, the adoption of the Conforming heuristic is unrelated to complexity, being used as frequently under the two voting procedures.

Secondly, we observe that the complexity of IRV adversely affected participants in the experiment

[17] It is interesting to note that surveys conducted in US jurisdictions that utilize IRV for political elections show that a relatively small proportion of voters express confusion regarding IRV (see, e.g., Donovan et al. (2022)).

and impaired their capacity to learn from experience. Indeed, participants' payoffs were on average lower in elections conducted using IRV as opposed to Borda. More than half of the payoff difference between IRV and Borda elections can be directly attributed to differences in participants' voting behavior, while the remaining portion is attributable to the mechanics of the two voting procedures. Additionally, contrary to Borda, IRV prevented participants from learning and increasing their earnings with experience, despite an observed higher frequency of vote changes under IRV compared to Borda. Even more concerning is the finding that participants would have obtained higher payoffs under IRV if, under this voting procedure, they had cast the same ballot as the one they submitted under Borda.[18] This illustrates the difficulty that participants faced when trying to figure out how best to vote under a complex voting procedure like IRV.

Finally, we find that IRV is more equitable than Borda, as it results in more equal payoffs among participants.

Similar to our findings on the complexity of the voting procedure, we observe that, under both IRV and Borda, participants are less likely to cast an optimal ballot in elections involving four candidates compared to elections involving only three candidates. Furthermore, as we find that participants tend to change their vote between two consecutive elections more often under IRV than under Borda, we also find a higher frequency of vote changes in voting situations with four candidates compared to the ones with three candidates.

The remainder of the paper is organized as follows. In the next section, we discuss the related literature and point out the contributions of our paper. We outline the experimental design in Section 3. We analyze voting behavior at the aggregate level in Section 4, and the determinants of individual voting behavior in Section 5. Finally, we conclude in Section 6. An appendix contains additional material. A supplementary online appendix contains the instructions and screenshots.

2. Related literature

Our paper contributes mainly to three branches of literature.

Literature on IRV. Previous research shows that, compared to several other voting procedures, IRV provides fewer opportunities and weaker incentives to vote strategically. In particular, Chamberlin et al. (1984) and Chamberlin (1985) run computational simulations to compare the incentives for non-sincere strategic voting under IRV and Borda (and several other voting procedures).[19] Building upon these works, we look at actual, instead of simulated, voting behavior. Therefore, our focus is less on the incentives to vote strategically and more on voters' ability to do so. Furthermore, we identify voting

[18] It is worth noting that the reverse is not true: participants would have obtained lower payoffs under Borda if, under this voting procedure, they had cast the same ballot as the one they submitted under IRV.

[19] Eggers and Nowacki (2024) do so as well, comparing IRV and Plurality Voting.

heuristics that voters choose to adopt when they are unable, or unwilling, to vote strategically.

Van der Straeten et al. (2010) conduct a laboratory experiment to compare voters' behavior under four voting procedures, including IRV (but not Borda), in a spatial voting setting with single-peaked preferences and a large number of voters (21 or 63). Similar to our study, they find that complexity hinders voters' ability to vote strategically, albeit in a different setting and with a different set of voting procedures. We build upon their work in several ways. Firstly, we isolate computational complexity from strategic complexity by presenting participants with the votes of the other, computerized, voters before they make their own voting decisions. This enables us to induce and, therefore, control the beliefs of each participant regarding the voting behavior of others. Because we control beliefs, we can determine whether participants best respond to their beliefs. Secondly, we compare voters' behavior under IRV and Borda, two voting procedures that differ in complexity while sharing the same ballot structure. This enables us, among other things, to isolate the effect of the complexity of the voting procedure from the complexity of ranking multiple candidates, and to obtain interesting findings relative to the effect of voting procedure complexity on voters' payoffs. Thirdly, we capture various forms of strategic voting behavior and identify when a heuristic, and which one, is adopted in response to complexity. More precisely, we identify *Lifting* as the primary heuristic that voters adopt as a response to the complexity of voting strategically under IRV.[20] Fourthly, we identify determinants of individual voting behavior and how strategic behavior varies with individual characteristics and the complexity of the voting procedure.

Other works on IRV investigate candidates' incentives for targeting a subset of voters versus appealing to the entire electorate (Buisseret and Prato, 2023; Myerson, 1993b), and candidates' incentives for entering the race (Callander, 2005; Dellis et al., 2017). In contrast to these studies, we look at the actual behavior of human voters instead of looking theoretically at candidates' behavior under IRV.

Literature on strategic voting. Our paper contributes to a large literature on strategic voting.[21] Of primary relevance are laboratory experimental works that look at voters' behavior under either IRV or Borda. We have already mentioned Van der Straeten et al. (2010) that look at voters' behavior under IRV. Several other works investigate voters' behavior under Borda. For instance, Forsythe et al. (1996) study the effect of pre-election polls on vote coordination in a divided-majority setting. Kube and Puppe (2009) and Granic (2017) examine the effect of information on voters' behavior. And Bassi (2015) looks at voters' strategic behavior. Like us, she finds that a substantial portion of votes cast under Borda is neither sincere nor consistent with strategic voting.

We make several contributions to this literature. Firstly, as mentioned above, we separate computa-

[20] Interestingly, Van der Straeten et al. (2010) find that under IRV, their participants inverted on average the ranks of one pair of candidates, which is consistent with participants adopting the *Lifting* heuristic.

[21] See, among others, Myatt (2007), Bouton and Castanheira (2012), and Kawai and Watanabe (2013). For a review of this literature, see Bol and Verthé (2021).

tional complexity from strategic complexity. As previously noted, this approach is based on controlling beliefs regarding others' voting behavior by assigning beliefs to participants. We can then identify, for various forms of strategic voting behavior and variously complex voting procedures, when participants cast a ballot consistent with strategic voting, when they vote sincerely, when they adopt a voting heuristic, and, in the latter case, which voting heuristic they adopt. Secondly, we compare voters' behavior under IRV and Borda, two voting procedures that differ in the complexity of voting strategically while sharing the same ballot structure. Thirdly, we identify factors that influence voters' behavior and explore how the impact of complexity on voters' behavior relates to individual characteristics like age, education, numeracy skills, social preferences, risk aversion, and lie aversion.

Literature on voting complexity. Our work is also related to a small literature that uses the laboratory experimental approach to investigate the effect of complexity in voting situations.[22] Herzberg and Wilson (1988) study strategic voting in a setting of sequential voting over an agenda. They capture computational complexity by varying the length of the agenda, meaning the number of amendments on the agenda. Like us, they design an experiment where participants know in advance the votes that the other, computerized, voters will cast. However, contrary to us, they do not find that the frequency of strategic voting decreases with computational complexity. Harrison and McDaniel (2008) investigate the conjecture that computational complexity induces people to vote sincerely. They do so by running a laboratory experiment in which they capture computational complexity by varying how much detail about the voting procedure (in their case, Young's Condorcet consistent voting procedure) they provide to participants.

Our work contributes to this literature in several ways. Firstly, we provide an experimental measure of voting complexity using how much time participants took to make their voting decisions.[23] Secondly, we account for participants' characteristics (e.g., age, education, numeracy skills) in our regression analysis. Thirdly, we go beyond examining strategic and sincere voting, identifying voting heuristics employed by participants in response to complexity. Fourthly, we consider voting procedures relevant to real-world elections and whose adoption is advocated in the public electoral reform debate.

[22] There is also a series of formal works in Economics and Computer Science that look at the relationship between computational complexity and the possibility of vote manipulation (e.g., Bartholdi and Orlin, 1991; Bartholdi III et al., 1989; Elkind et al., 2020). For a review of this literature, see, among others, Conitzer and Walsh (2016). At a more general level, our paper is also related to works on Computational Complexity Theory (e.g., Bossaerts and Murawski, 2017; Murawski and Bossaerts, 2016) and experimental works on complexity (e.g., Grimm and Mengel, 2012; Oprea, 2020).

[23] To the best of our knowledge, we are the first to use response time to measure complexity in the context of voting, although other laboratory experiments have already employed response time for various purposes (e.g., Nielsen and Rehbeck, 2022; Rubinstein, 2016; Wilcox, 1993).

3. Experiment

3.1. Experimental design

In the experiment, a group must select one among either three or four candidates.[24] The group consists of four voters: one experimental participant (hereafter, the *voter*), and three computerized voters. Following standard practice in the experimental voting literature (Forsythe et al., 1993), candidates are designated to participants by colors – Blue, Green, Orange, and Grey – and appear in random order on participants' screens. To simplify exposition, we shall from now on refer to candidates as A, B, C, and D. The voter has the following preference ordering over candidates: $A \succ B \succ C \succ D$, meaning the voter strictly prefers A to B, B to C, and, when there are four candidates, C to D. The voter has to rank all candidates (as is required, for instance, in elections to the Australian House of Representatives). Before casting her ballot, the voter observes the votes of the three computerized voters.[25] The winning candidate is selected using either IRV or Borda.

The two important features of the design are, first, the controlled variation of complexity and, second, the calibration of the vote profile of the three computerized voters. We explain both below.

3.1.1. Complexity of the voting situation

We vary different aspects of the complexity of the voting situation. Firstly, we compare voters' behavior under IRV to that under the less complex Borda voting procedure (procedural complexity). Secondly, we vary the number of candidates available for selection, considering voting situations involving three candidates and other, arguably more complex, involving four candidates (environmental complexity).

Description of IRV and Borda

We start by describing in more detail the functioning of IRV and Borda.

Under IRV, the winner of an election is the candidate whom a majority of voters ranks first. If neither candidate receives a majority of first rankings, then the candidate with the fewest first rankings is eliminated. In case of a tie, a random draw decides which of the candidates with the fewest first rankings is eliminated.[26] Every ballot is then updated, with the candidate ranked next on the ballot taking the place

[24] In the experimental instructions, we use neutral language, referring to candidates as "alternatives".

[25] The votes of the three computerized voters can be interpreted as the voter's beliefs about the ballots that the other three voters are about to cast. Alternatively, this setting corresponds to a roll-call vote, where voters cast their ballots sequentially and publicly, and our voter is the last member to vote.

[26] This corresponds, for example, to the tie-breaking rule used under IRV in Alaska: "In the event of a tie between two candidates with the fewest votes in a round, Alaska law resolves the tie "by lot" to determine which candidate is eliminated and which candidate advances to the next round. Believe it or not, "by lot" means the division's director will flip a coin, draw a name or draw straws. The same applies if there is a tie for the last two remaining candidates." (Retrieved from https://www.elections.alaska.gov/election-information/#RankedChoice on June 6th, 2024).

of the eliminated candidate. If one of the remaining candidates receives a majority of first rankings on the updated ballots, then this candidate is declared the winner and the process stops. Otherwise, the process continues with the elimination of the remaining candidate who receives the fewest first rankings on the updated ballots. This process is iterated until one candidate receives a majority of the first rankings.

Under Borda, each candidate receives from a voter a number of votes that varies with his position on the voter's ballot. Specifically, the voter gives each candidate a number of votes equal to the number of candidates she ranks below. The winner is the candidate with the highest vote total. As under IRV, ties are broken equiprobably.

The following example illustrates the functioning of the two voting procedures.

EXAMPLE 1. Consider a group composed of four voters (called 1, 2, 3, and 4) that must select one among three candidates (called A, B, and C). Voters 2 through 4 cast the following ballots: BAC, BAC, and CAB, that is, each of Voters 2 and 3 ranks B first, A second, and C third, while Voter 4 ranks C first, A second, and B third. Suppose that Voter 1 were to submit the sincere vote ABC.

Rank	Voter 1	Voter 2	Voter 3	Voter 4
1	A	B	B	C
2	B	A	A	A
3	C	C	C	B

Under IRV, a candidate needs at least three first rankings to be the winner. At the first count, no candidate receives a majority of the first rankings, while each of A and C receives only one first ranking. Thus, neither candidate is elected, and one among A and C is eliminated by a random draw. If A is eliminated, which happens with probability 1/2, the updated ballots are then BC, BC, BC, and CB. Candidate B is now ranked first three times and is thus the winner. If instead, C is the eliminated candidate, which happens with probability 1/2, the updated ballots are AB, BA, BA, and AB. Candidates A and B are now ranked first twice, meaning that neither has yet a majority of the first rankings. One of them is then eliminated by a random draw. Thus, either A or B is the winner, each with an equal probability. To sum up, the submitted ballots result in the following winning probabilities under IRV: $Pr(\mathbf{A}) = 1/4; Pr(\mathbf{B}) = 3/4; Pr(\mathbf{C}) = 0$. Note that given the ballots cast by the other three voters, submitting a sincere vote is the best that Voter 1 can do. Hence, in this example, the strategic vote under IRV is sincere.

Under Borda, A and B receive five votes each (A : two from Voter 1, plus one from every other voter; B : two from Voter 2 and from Voter 3, one from Voter 1, and zero from Voter 4), and C receives two votes (from Voter 4, and zero from every other voter). Candidates A and B receive both the most votes and one of them is selected randomly as the winner. To sum up, the submitted ballots result in the following winning probabilities under Borda: $Pr(\mathbf{A}) = Pr(\mathbf{B}) = 1/2$ and $Pr(\mathbf{C}) = 0$. However, Voter 1 could do better under Borda by submitting the ballot ACB, ensuring a certain win for A (with five votes, compared to only four votes for B and three for C). The winning probabilities are then: $Pr(\mathbf{A}) = 1$ and $Pr(\mathbf{B}) = Pr(\mathbf{C}) = 0$. Hence, in this example, the strategic vote under Borda is non-sincere.

Computational complexity of the voting procedure

Why is IRV a more complex voting procedure compared to Borda? Firstly, IRV is cognitively more challenging. This is notably the case when no candidate immediately receives a majority of the first rankings. In this case, a voter needs to anticipate the unraveling of the whole elimination process, while Borda just requires counting and adding up the votes for each candidate. Secondly, while Borda satisfies the *monotonicity* property, meaning that raising the rank of a candidate does not lessen its winning probability, IRV violates this property. This implies that increasing the rank of a candidate under IRV may lessen his winning probability. Such reasoning is not only counter-intuitive but needs considerable strategic thinking. Thirdly, outcomes under IRV appear to be more probabilistic, what Santucci (2021) refers to as the *lottery effect* of IRV.[27] As a result, voters need to consider more often uncertain outcomes under IRV even when they have resolved the strategic uncertainty about other voters' behavior. Decision-making under uncertainty is complex. Most violations of the rationality axioms have been shown to occur for decisions under uncertainty (Grether and Plott, 1979; Lichtenstein and Slovic, 1971).

At the same time, IRV might be seen as simpler than Borda when it comes to the number of optimal ballots. Indeed, often more than one ballot is payoff maximizing under IRV. This occurs because once one of the candidates reaches a majority of the first rankings, the remaining ranks on the ballots are not considered and make no difference. For example, when a candidate immediately obtains a majority of the first rankings, IRV ignores how voters have ranked candidates down on their ballots. By contrast, under Borda, all ranks count for the final score and, most often, only a single ballot is optimal.[28] An absent-minded voter, who chooses a ballot randomly, or a trembling voter, who makes mistakes, will therefore more likely submit an optimal ballot under IRV than under Borda.

3.1.2. Vote profiles

We prepared in advance the ballots of the three computerized voters, which we shall refer to as the *vote profile*. It is important to mention that each participant knows that the three computerized voters are not humans participating in the experimental session. Before deciding which ballot to cast, the voter is informed about the vote profile. This approach has several advantages. Firstly, it simplifies the calculus

[27] In our experiment, the *lottery effect* of IRV is captured with a greater likelihood of ties under IRV than under Borda. This is a general feature of elections involving four voters and three or four candidates, where more vote profiles result in a tie under IRV than under Borda.

[28] This feature is not specific to the vote profiles selected for this experiment but is rather a general feature of elections involving four voters and three or four candidates. There are indeed more vote profiles for which IRV admits a multiplicity of optimal ballots than there are in the case of Borda.

of the voter by removing the strategic complexity of anticipating others' voting behavior. This allows us to isolate computational complexity from strategic complexity. Secondly, it endows the voter with beliefs about the voting behavior of the other voters. This allows us to control the voter's beliefs and, together with the induced preferences, identify optimal ballots and ballots that are consistent with voting heuristics. Preparing the ballots of the other voters in advance has the additional advantage of allowing us to control the voter's incentives as it is possible to calibrate meaningful and interesting vote profiles that induce a representative sample of the different forms of strategic voting behavior under IRV and Borda.

From all possible vote profiles (216 for elections involving three candidates, and 13,824 for elections involving four candidates), we selected a total of six based on two criteria. The first criterion is that the voter's ballot is decisive under both IRV and Borda, meaning the voter's ballot matters in determining the winning candidate.[29] The second criterion is that the vote profiles capture various forms of strategic voting behavior under IRV and Borda.[30] Half of the retained vote profiles involve elections with three candidates, while the other half involve elections with four candidates.

The retained vote profiles (Table 1) capture several forms of strategic voting behavior: *Lifting, Burying, Overstating,* and *Sincere*. We name a vote profile by the strategic incentives it provides to the voter. If the voting strategies differ between the two voting procedures, we name the incentives for IRV before those for Borda. The number at the end of a vote profile indicates the number of candidates. The six vote profiles are *Lift-3, Lift-4, Lift-Overstate-4, Sincere-Bury-3, Sincere-Bury-4,* and *Sincere-3*. In the following, we explain the voting strategies and the purpose of the retained vote profiles.[31]

The first form of strategic voting behavior involves **lifting** a candidate, meaning moving the rank of a candidate up to improve his winning prospects. We consider three vote profiles where lifting is optimal. The first two vote profiles, (*Lift-3* and *Lift-4*), concern voting situations involving three and four candidates with the three computerized voters casting ballots BCA, CAB, CBA and $BCDA, DBAC, DACB$, respectively. For both vote profiles, lifting the second preferred candidate (B) to improve his winning prospects is optimal under both IRV and Borda. The third vote profile, *Lift-Overstate-4*, concerns a voting situation involving four candidates with the computerized voters casting ballots $DBCA, CBAD$,

[29] For elections involving three candidates, 80.6% of the possible vote profiles satisfy this criterion. For elections involving four candidates, 89.1% of the possible vote profiles satisfy this criterion.

[30] For almost every possible vote profile, one of the voting strategies selected based on this criterion is optimal. Indeed, the types of voting strategies selected in the case of elections involving three candidates are optimal in 100% of the possible vote profiles, under both IRV and Borda. The types of voting strategies selected in the case of elections involving four candidates are optimal in 98.7% of the possible vote profiles (that satisfy the first selection criterion, that is, where the voter's ballot is decisive) under IRV, and in 79% of them under Borda.

[31] Table A2 in the Appendix reports, for each vote profile, every candidate's winning probability for every possible ballot under each voting procedure.

(1)	(2) **Vote profiles (ballots cast by the three computerized voters)**	(3) **Optimal ballots IRV**	(4) **Optimal ballots Borda**
1a) ***Lift-3***	*BCA*,*CAB*,*CBA*	*B*− (*BAC*, *BCA*)	*BAC*
1b) ***Lift-4***	*BCDA*,*DBAC*,*DACB*	*B*− (*BACD*, *BADC*, *BCAD*, *BCDA*, *BDAC*, *BDCA*)	*B* − *D* (*BACD*, *BCAD*)
2) ***Lift-Overstate-4***	*DBCA*,*CBAD*,*CABD*	*BACD*, *BCAD*, *BCDA*	*B* − *C* (*BADC*, *BDAC*)
3a) ***Sincere-Bury-3***	*BAC*,*BAC*,*CAB*	*ABC*	*ACB*
3b) ***Sincere-Bury-4***	*BADC*,*CDAB*,*DBAC*	*AB*− (*ABCD*, *ABDC*)	*ACBD*
4) ***Sincere-3***	*BAC*,*CAB*,*CAB*	*ABC*	*ABC*

Table 1: Vote profiles and optimal ballots.

Each vote profile (column (2)), is named (column (1)) after its strategic incentives. If they differ across voting procedures, incentives for IRV are named first. The number at the end of a vote profile indicates the number of candidates. Optimal ballots under IRV and Borda appear in columns (3) and (4).

and *CABD*. Here lifting *B* is optimal under both IRV and Borda, but is only part of the optimal voting strategy under Borda (as we explain below). We get a greater multiplicity of optimal ballots under IRV than under Borda (2 vs. 1 with *Lift-3*; 6 vs. 2 with *Lift-4*; 3 vs. 2 with *Lift-Overstate-4*).

We use these vote profiles to study 1) whether people move a candidate up on their ballot, compared to their actual preference ordering, when it is optimal to do so, and 2) how much complexity matters for people to adopt this voting strategy.

Another form of strategic voting behavior involves **burying** a candidate, that is, moving the rank of a candidate down to weaken his chances of winning. We consider two vote profiles where burying is optimal under Borda. The first vote profile (*Sincere-Bury-3*) concerns a voting situation involving three candidates with the computerized voters casting ballots *BAC*,*BAC*, and *CAB*. Under Borda, the unique optimal ballot *ACB* consists of ranking *B* last to prevent it from winning. The second vote profile

(*Sincere-Bury-4*) applies to a voting situation involving four candidates with the computerized voters casting ballots *BADC*,*CDAB*, and *DBAC*. Under Borda, the unique optimal ballot *ACBD* consists of moving *B* down by only one rank to prevent him from winning without triggering the election of *D*. Under IRV, voting sincerely is optimal for both vote profiles. More precisely, the sincere ballot (*ABC*) is the unique optimal ballot with *Sincere-Bury-3*, and any ballot that ranks *A* first and *B* second is optimal with *Sincere-Bury-4*. Hence, there are two optimal ballots under IRV for *Sincere-Bury-4*: the sincere ballot *ABCD*, and the ballot *ABDC*.

We use these two vote profiles to study whether people cast different votes under IRV and Borda when their sets of optimal ballots are different. We also use them to learn whether the sincere vote is focal when it is one among several optimal ballots, as is the case under IRV in *Sincere-Bury-4*. Finally, we can observe whether people pull a candidate down on their ballot when it is optimal to do so.

The next form of strategic voting behavior that we consider is **overstating** the ranking gap between two candidates. This form of strategic voting behavior implies lifting one candidate while, at the same time, burying another. The vote profile *Lift-Overstate-4* concerns a voting situation involving four candidates with computerized voters casting ballots *DBCA*,*CBAD*, and *CABD*. Here, the optimal ballots *BADC* and *BDAC* under Borda imply overstating the ranking gap between *B* and *C*, that is, burying *C* to prevent him from winning and lifting *B* to ensure his victory. By contrast, strategic voting under IRV involves only lifting. Here, increasing the rank of *B* maximizes the voter's expected payoff, as under Borda, but without having to bury *C*.

Again, we use this vote profile to study whether people cast different ballots under IRV and Borda when their sets of optimal ballots differ. We also use this vote profile to observe whether people widen the ranking gap between candidates when it is optimal to do so.

Finally, we consider a voting situation where **Sincere** voting, meaning submitting a ballot that corresponds to one's genuine preference ranking of the candidates, maximizes expected payoffs. The vote profile *Sincere-3* concerns a voting situation involving three candidates with one computerized voter casting ballot *BAC* and the other two casting ballot *CAB*. The sincere vote *ABC* is the unique optimal ballot for the voter under both IRV and Borda. We use this vote profile as a benchmark and to study how much complexity matters when strategic voting coincides with sincere voting.

3.1.3. Sampling and repetition

We adopt a within-subject design, where each participant encounters all six vote profiles and casts ballots under both IRV and Borda. To control for the possibility of order effects, around half of the participants started by casting all their votes under IRV, and continued with Borda, whereas the rest of the participants faced the inverse ordering of voting procedures, starting with Borda and continuing with IRV. The order

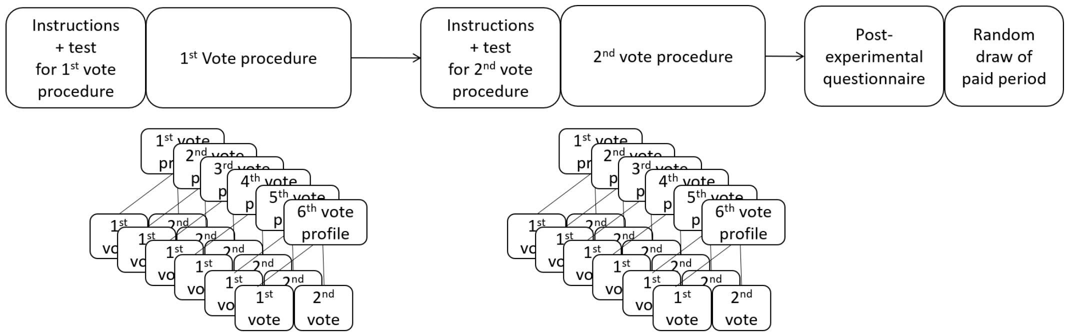


Fig. 1: Experimental design. Random order for vote procedures and vote profiles.

in which the six vote profiles were presented was randomized across participants, while the order in which vote profiles were presented to a participant was similar under the two voting procedures. To allow for learning, every participant repeats each voting situation twice consecutively. Hence, we observe a total of 24 voting decisions for every participant (2 voting procedures $\times$ 6 vote profiles $\times$ 2 votes per vote profile). Figure 1 presents the experimental design.

In summary, the experimental design controls for voters' preferences, their beliefs about others' voting behavior, the forms of strategic voting behavior, and the complexity of voting strategically associated with the voting procedure and the number of candidates. We observe the voting behavior of every participant under the two voting procedures and for the same six vote profiles, presented to the participant in the same order under both voting procedures. However, the ordering of the voting procedures and the vote profiles vary across participants.

3.2. Experimental procedure

We conducted 15 sessions, five of them at the LEEL (Laval Experimental Economics Laboratory) at Université Laval in Quebec City and the other ten at the Claude Montmarquette Laboratory at CIRANO in Montréal, for a total of 104 participants. Participants were recruited from each laboratory's mailing list containing persons who are generally interested in participating in experiments.[32]

Upon arrival, participants were randomly assigned to individual computer terminals separated by opaque panels. After having collected all signed consent forms, the experimenter started the session. Participants voted sequentially in two series of elections, each held under a different voting procedure. We used neutral language, labeling voting procedures as the first and second voting procedures. Both series started with participants reading the computerized instructions explaining the corresponding voting procedure. Each election series started after having correctly answered a set of comprehension questions designed to test their understanding of the voting procedure and ensure that our findings are not the

[32] For recruiting, the CIRANO used the ORSEE software (Greiner, 2015). LEEL used an in-house recruitment software.

result of confusion on the part of participants.[33]

In each election, the voter's payoff was 40$ (resp. 30$, 20$, or 10$) if A (resp. B, C, or D) was the winning candidate. During an election, the participant was presented first with a *ballot screen* containing (i) the vote profile, (ii) the participant's payoff from each candidate, and (iii) a ballot to fill. To be valid, a vote required a ranking of all candidates. Once the participant had cast her vote, a *result screen* appeared reporting the outcome and the participant's payoff in the election.[34] In the repeat election, a sentence appeared on the computer screen to notify the participant that the vote profile of the computerized voters was the same as in the previous election.

After the end of the first series of elections, the second series started using the other voting procedure. Out of the 104 participants, 56 randomly selected participants started with IRV and 48 with Borda. Upon completion of the second series of elections, the participants were invited to answer a post-experimental questionnaire on their socio-demographic characteristics (gender, age, education) as well as a series of questions designed to evaluate the participant's 1) attitude towards risk, 2) lying aversion, 3) social value orientation, and 4) numeracy skills using the Berlin Numeracy Test (Cokely et al., 2012).[35]

At the end of the experiment, one of the 24 elections was randomly drawn to determine each participant's payoff. On top of that amount, each participant received a 15$ show-up fee. Total payments ranged from 25$ to 55$, with an average of 39.62$. Participants were paid in cash and in private. The session lasted on average slightly less than an hour.

4. Aggregate outcomes

4.1. Complexity

We start by checking that our design effectively captures that voting strategically is more complex under IRV than under Borda, as well as in voting situations involving four candidates compared to those with only three candidates. To do so, we use an empirical measure of 'revealed' complexity.

Complex situations require more consideration and attention on the part of the decision-maker because they are cognitively more challenging and, as a consequence, take more time. Therefore, we expect complexity to be reflected in the amount of time that voters take to make their voting decisions.

We start with the two participants who, under both IRV and Borda, always submitted an optimal ballot, meaning a ballot consistent with strategic voting behavior, during the first attempt (that is, the

[33] See the Supplementary online appendix for the transcript of instructions and comprehension tests.

[34] Under Borda, the result screen reported each candidate's total vote score. Under IRV, the result screen reported the elimination sequence, specifying for each elimination round the number of first rankings for each of the remaining candidates as well as the candidate eliminated in that round. Under both IRV and Borda, the result screen also reported the candidate selected as the election winner and the participant's payoff in that election. Screenshots of the ballot and result screens are displayed in the Supplementary online appendix.

[35] See the Supplementary online appendix for transcripts of the post-experimental questionnaires.

first of the two consecutive votes for a given vote profile and voting procedure).[36]

	Decision time		
Vote profile	**IRV**	**Borda**	**All**
Lift-3	76	60	68
Sincere-Bury-3	63	54	58
Sincere-3	197	67	132
Mean - 3	112	60	86
Lift-4	293	111	202
Lift-Overstate-4	229	73	151
Sincere-Bury-4	164	121	143
Mean - 4	229	101	165
Mean - all	170	81	125

Table 2: Average decision time in seconds by voting procedure, vote profile, and number of candidates for arriving at an optimal ballot for voters who submitted an optimal ballot at all 12 first-attempt votes. ($N = 2$)

Table 2 reports average voting decision times separately by vote profile, number of candidates, and voting procedure. We find empirical support for strategic voting being more complex under IRV than Borda, with an average voting decision time of 170 seconds under IRV versus 81 seconds under Borda. This pattern of longer voting decision time under IRV repeats for each of the six vote profiles. We also observe that arriving at an optimal ballot took more time in voting situations involving four candidates compared to those with only three candidates, both in general (165 sec vs. 86 sec on average) and for each voting procedure (IRV: 229 sec vs. 112 sec; Borda: 101 sec vs. 60 sec).

As those voting decision times are based on only two participants, Appendix B presents the same tables but for the participants who submitted at least 11 ($N = 5$), 10 ($N = 9$), and 9 ($N = 17$) optimal ballots during the twelve first attempts, respectively. These tables show that the general pattern observed in Table 2 is robust and that a decrease in the submitted number of optimal ballots usually goes hand in hand with a decrease in the average voting decision time.

To sum up, IRV is more complex than Borda in the framework of our experiment, as participants took on average more time to submit an optimal ballot under IRV than under Borda. Similarly, when it comes to voting situations with four candidates compared to those with only three candidates, participants took

[36] We focus on decision time during the first attempt since it is the first time that the voter is confronted with this vote profile under the considered voting procedure. Not surprisingly, decision time tends to be shorter during the second attempt.

more time to submit an optimal ballot in the former situations compared to the latter ones.

4.2. *Complexity and voting behavior*

We now study how complexity affects voters' behavior at the aggregate level. We do so by comparing voting decisions under IRV and Borda.

In Table 3, we report descriptive statistics consistent with different categories of voting behavior, comprising strategic voting (Column *Optimal*), sincere voting (Column *Sincere*), and two voting heuristics that we explore in the next section (Columns *Lifting* and *Conforming*).[37,38]

We observe that complexity appears to make it hard for voters to vote strategically. Column *Optimal* of Table 3 reports the frequency of optimal votes, meaning votes that are consistent with strategic voting behavior. Only 35.4% of the ballots cast under IRV are optimal compared to 49.2% under Borda. The difference between the two voting procedures is substantial and statistically significant ($p = 0.0005$ for a Wilcoxon signed-rank test, hereafter *WSR test*).[39] Moreover, for every vote profile, except the two *Lift* profiles, the frequency of optimal votes is lower under IRV than under Borda and the difference is statistically significant (using McNemar χ^2 test, hereafter *McNemar test*). Interestingly, the lower frequency of optimal votes under IRV compared to Borda occurs even though IRV admits, for several vote profiles, a greater multiplicity of *Optimal* ballots than Borda.[40]

Not only are fewer optimal votes cast under IRV than under Borda, but fewer of these votes seem to be the result of strategic voting behavior. To see this, we compare (1) the frequency with which ballots are cast in a vote profile where they are optimal with (2) the frequency with which the same ballots are cast

[37] In Table 3 we report only the votes submitted in the first attempt. We analyze the second-attempt votes in Section 4.2.2.

[38] It is worth observing that under both IRV and Borda, the frequencies of submitted ballots consistent with the four categories of voting behavior are substantially different from the frequencies we would have observed if voters had chosen their ballots randomly (see Table A6 in Appendix C). We can therefore reject the hypothesis that the observed election outcomes are the result of random voting behavior, and accept the alternative hypothesis that they are the result of purposeful voting behavior.

[39] We use participants as independent units of observation for statistical tests. All reported tests are two-sided. We use two non-parametric statistical tests: McNemar χ^2 test and Wilcoxon signed-rank test. McNemar χ^2 test can be used in within-subjects designs to assess a change in binary dependent variables, in particular the consistency with a category of voting behavior of the ballots cast under the two voting procedures for the same vote profile. Wilcoxon signed-rank test can be used in within-subjects designs to assess a difference in (multinomial) dependent variables, in particular, the frequencies with which the ballots cast over the six vote profiles and under the two voting procedures are consistent with a category of voting behavior.

[40] It is worth observing that when multiple ballots are optimal, the ballot corresponding to the strategy identifying the vote profile (*Lifting, Sincere, Burying*, or *Overstating*) is played overwhelmingly among the optimal ballots: between 84% (in the *Lift-Overstate-4* vote profile) and 86% (in the *Lift-4* vote profile) of the time under Borda, and between 64% (in the *Lift-Overstate-4* vote profile) and 96% (in the *Sincere-Bury-4* vote profile) of the time under IRV. In particular, the frequency with which the sincere ballot is cast (96%) under IRV among the two optimal ballots in the *Sincere-Bury-4* vote profile illustrates the focal nature of the sincere ballot when it is one among several optimal ballots.

Vote profile	Voting procedure	Voting behavior *Optimal*	*Sincere*	*Lifting*	*Conforming*
All profiles	**IRV**	35.4%	20.0%	38.5%	23.4%
	Borda	49.2%	18.6%	28.0%	22.4%
	p-value	*0.0005*	*0.7145*	*0.0002*	*0.4177*
Lift-3	**IRV**	49.0%	18.3%	39.4%	18.3%
	Borda	51.9%	5.8%	51.9%	29.8%
	p-value	*0.631*	*0.0067*	*0.0236*	*0.0233*
Lift-4	**IRV**	62.6%	9.6%	46.2%	7.7%
	Borda	62.5%	9.6%	53.8%	9.6%
	p-value	*1*	*1*	*0.217*	*0.593*
Lift-Overstate-4	**IRV**	26.9%	19.2%	17.3%	28.8%
	Borda	41.3%	7.7%	9.6%	27.9%
	p-value	*0.0222*	*0.0186*	*0.0594*	*0.8474*
Sincere-Bury-3	**IRV**	25.0%	25.0%	56.7%	56.7%
	Borda	48.1%	18.3%	30.8%	30.8%
	p-value	*0.0003*	*0.2498*	*0.0002*	*0.0002*
Sincere-Bury-4	**IRV**	27.0%	26.0%	35.6%	6.7%
	Borda	36.5%	15.4%	12.5%	14.4%
	p-value	*0.0864*	*0.0706*	*<0.001*	*0.0325*
Sincere-3	**IRV**	22.1%	22.1%	35.6%	22.1%
	Borda	54.8%	54.8%	9.6%	22.1%
	p-value	*<0.001*	*<0.001*	*<0.001*	*1*

Table 3: Frequencies of first-attempt votes consistent with each category of voting behavior and p-values from a Wilcoxon signed rank test/McNemar test that a category of voting behavior is used the same way under both voting procedures. Note that *Sincere* and *Lifting* may contain optimal votes in the Sincere and Lift profiles, and that *Conforming* may contain *Lifting* votes (and vice versa) in the *Sincere-Bury-3* profile.

in vote profiles where they are *not* optimal. If the submission of an optimal ballot is the result of strategic voting behavior, we should observe that this ballot is not cast in vote profiles where it is not optimal. This is essentially what we observe under Borda, but not under IRV. Indeed, 35.4% of the ballots cast under IRV are optimal, and the same ballots are submitted 34.4% of the time in vote profiles where they are not optimal. The difference is negligible and not statistically significant ($p = 0.3978$, WSR test).[41] In other words, under IRV, a ballot is cast almost as often when it is optimal to cast this ballot as when it is not. By contrast, under Borda, 49.2% of the ballots cast are optimal and the same ballots are cast only 20.4% of the time in vote profiles where they are not optimal. This time, the difference is substantial and statistically significant ($p < 0.001$, WSR test).[42]

4.2.1. Voting heuristics

We now identify the voting heuristics that voters adopt when they do not vote strategically.

We find some, although limited, support for the argument that voters react to the complexity of IRV by resorting to sincere voting. Overall, the frequency of sincere votes (Column *Sincere* of Table 3) is slightly higher under IRV (20%) than under Borda (18.6%), but the difference is not statistically significant ($p = 0.7145$, WSR test). However, this overall observation hides variations across vote profiles. Specifically, for the *Sincere-3* profile, where voting sincerely is optimal under both IRV and Borda, the frequency of sincere votes is significantly larger under Borda than under IRV (54.8% vs. 22.1%; $p < 0.001$, McNemar test). The reverse holds for the other vote profiles, where voting sincerely is not optimal under Borda. Indeed, for each of these vote profiles, the frequency of sincere votes is (weakly) larger under IRV than under Borda, and the difference is statistically significant for three of the five vote profiles. In other words, except when voting sincerely is optimal under Borda, our observations are (partially) in agreement with the argument that voters react to the complexity of IRV by resorting to sincere voting.

While we find partial support for the claim that the complexity of IRV prevents voters from voting strategically and induces them to vote sincerely, this support needs qualifications. Indeed, only 43.3% of all the ballots cast under IRV (compared to 58.7% under Borda) are consistent with either strategic voting or sincere voting (or both). This means that many voters did not vote strategically and did not react to the complexity of voting strategically under IRV by voting sincerely. This raises the question of whether complexity induces voters to adopt other voting heuristics than *Sincerity*. We identify two voting heuristics, which we refer to as *Lifting* and *Conforming*, with which a sizable fraction of the votes cast under IRV is consistent.

The *Lifting* heuristic consists of inverting the ranks of the top two preferred candidates. Specifically, we say that a voter's behavior is consistent with the Lifting heuristic if the voter casts ballot *BAC* in voting situations with three candidates, meaning ranks *B* first, *A* second, and *C* third, and submits ballot

[41] Table A7 in Appendix D reports detailed frequencies and McNemar tests by vote profile.

[42] Furthermore, if the submission of an optimal ballot in the first attempt is the result of strategic voting behavior, the participant should not change her vote in the second attempt. We show in Section 4.2.2 that this is the case under Borda, but not under IRV. Specifically, participants who submit an optimal ballot in the first attempt still submit an optimal ballot in the second attempt 92% of the time under Borda versus only 61% of the time under IRV.

BACD in voting situations with four candidates.[43]

Column *Lifting* of Table 3 reports the frequencies of lifting votes. We observe that a substantial fraction of the votes cast under IRV (38.5%) are lifting votes. Interestingly, this fraction is almost twice as large as the fraction of sincere votes (20%) and tends to be relatively similar in vote profiles where lifting is not optimal under IRV (namely, *Sincere-Bury-3, Sincere-Bury-4* and *Sincere-3*) as in vote profiles where it is optimal (namely, *Lift-3, Lift-4* and *Lift-Overstate-4*). Furthermore, we observe that the frequency of lifting votes is substantially larger under IRV than under Borda (38.5% vs. 28%; $p = 0.0002$, WSR test). Interestingly, we observe important variations across vote profiles. Indeed, for *Lift-3* and *Lift-4* vote profiles, where lifting is optimal under both voting procedures, the frequency of lifting votes is larger under Borda than under IRV, which is consistent with our previous observation that voters tend to cast an optimal ballot more often under Borda than under IRV. The reverse holds for each of the other four vote profiles, where lifting is not optimal under Borda. For each of these four vote profiles, the frequency of lifting votes is significantly larger under IRV than under Borda.

To sum up, these observations suggest that voters react to the complexity of IRV partly by adopting the Lifting heuristic and that they do so almost twice as often as voting sincerely.

We identify a second voting heuristic, referred to as *Conforming*, with which a sizable fraction of the votes cast under IRV and Borda is consistent. This heuristic consists of submitting a ballot that is similar to the ones cast by the three computerized voters. We measure similarity using the Hamming distance.[44] Specifically, we say that a voter's behavior is consistent with the Conforming heuristic if the voter casts the ballot that has the smallest total Hamming distance with the ballots of the three computerized voters. The Hamming distance between two ballots represents the minimum number of pair inversions required to transform one ballot into another.[45] Table A1 in Appendix A reports the conforming ballot for each of the six vote profiles.

We find that 23.4% of the ballots cast under IRV are consistent with the Conforming heuristic (Column *Conforming* of Table 3). This frequency is slightly larger than the 20% of sincere votes cast under IRV, even though, contrary to a sincere vote, a conforming vote is never optimal. We also observe that the frequency of conforming votes is similar under IRV and Borda (23.4% vs. 22.4%; $p = 0.4177$, WSR test).[46] The overall absence of a difference in the frequencies of conforming votes under IRV and Borda

[43] It is worth observing that the Lifting heuristic is reminiscent of strategic voting behavior under Plurality Voting when a voter's most preferred candidate is trailing while the voter's second most preferred candidate is one of the serious contenders.

[44] The Hamming distance measures the difference between two strings of equal length (Hamming, 1950). It originates from information theory and computer science, and is sometimes used in voting theory (Elkind et al., 2012).

[45] For example, the Hamming distance between ballots *ABC* and *CAB* is equal to 2, as the pairs *AC* and *BC* must be inverted into *CA* and *CB* to transform ballot *ABC* into ballot *CAB*, and vice versa. The Hamming distance can take values between 0 and 3 (resp. 6) in elections involving three (resp. four) candidates.

[46] There is an exception for the *Sincere-Bury-3* vote profile, where the frequency of conforming votes is significantly larger

suggests that the adoption of the Conforming heuristic is not related to the complexity of the voting procedure.

Summing up:

RESULT 1. We have observed that:

1. The complexity of IRV does impede strategic voting behavior.

2. There is limited support for the argument that voters react to complexity by voting sincerely.

3. As an additional response to the complexity of IRV, voters tend to rely on the *Lifting* heuristic, and do so more often than voting sincerely.

4. Voters sometimes adopt the *Conforming* heuristic, but they appear to do so independently of the complexity of the voting procedure.

4.2.2. Learning from experience

Our experimental design employs and controls for learning opportunities, enabling us to investigate the effect of experience on voters' behavior and the extent to which this effect depends on the complexity of the voting procedure. There are three sources of learning in our experimental design: 1) learning from experience through the repetition of the same vote profile under the same voting procedure (two consecutive vote attempts); 2) learning from experience with other vote profiles, through the sequential ordering of vote profiles under the same voting procedure (six consecutive vote profiles); and 3) learning from experience under the other voting procedure, through the sequential ordering of voting procedures (two consecutive voting procedures). In this section, we study the first source of learning by looking at individual vote changes between the first and second vote attempts with the same vote profile and voting procedure. We study the other two sources of learning in Section 5, where we identify the determinants of individual voting behavior.

Consistent with the claim that complexity makes it hard for voters to figure out how they should vote, we find a higher frequency of vote changes between the two attempts with the same vote profile under IRV compared to Borda. More specifically, voters changed their vote between the two attempts 47% of the time under IRV compared to only 26% of the time under Borda (Table 4). The difference is statistically significant ($p < 0.001$, WSR test) and repeats for each of the six vote profiles. Likewise, we find a substantially higher frequency of vote changes in elections involving four candidates than in elections involving only three candidates under both IRV (57% vs. 37%; $p < 0.001$, WSR test) and

under IRV than under Borda (56.7% vs. 30.8%; $p = 0.0002$, McNemar test). This difference is explained by the coincidence of the conforming and lifting votes in this vote profile (Table A1 in Appendix A) together with a higher frequency of lifting votes under IRV than under Borda when a lifting vote is not optimal.

Borda (36% vs. 17%; $p < 0.001$, WSR test). These observations suggest that voters are more inclined to think that they can improve upon their original, first-attempt vote when the voting situation is more complex.

Vote profile	Frequency of vote change			Intensity of vote change		
	IRV	Borda	*p-value*	IRV	Borda	*p-value*
All profiles	47%	26%	*<0.001*	1.69	1.67	*0.7625*
Lift-3	43%	22%	*0.0005*	1.27	1.43	*0.1094*
Lift-4	55%	42%	*0.0687*	2.18	1.84	*0.2231*
Lift-Overstate-4	59%	35%	*0.0002*	1.95	1.47	*0.2539*
Sincere-Bury-3	32%	9%	*<0.001*	1.42	1.44	*1*
Sincere-Bury-4	57%	30%	*<0.001*	1.59	2.00	*0.6931*
Sincere-3	37%	19%	*0.0044*	1.45	1.55	*1*

Table 4: Vote changes between the two attempts with the same vote profile and voting procedure.

At the same time, we find that the *intensity* of vote changes is independent of complexity. We measure the intensity of vote changes by computing the Hamming distance between the ballot submitted on the first attempt and the ballot submitted on the second attempt. The higher the Hamming distance, the greater the intensity of the vote change as the voter makes more changes between the original ballot submitted in the first attempt and the ballot submitted in the second attempt.

Conditional on the vote being changed between the two attempts (that is, restricting attention to Hamming distances different from zero), we find that the Hamming distance between the ballots submitted in the two attempts is on average slightly larger under IRV compared to Borda (Table 4: 1.69 vs 1.67). However, the difference is not statistically significant ($p = 0.7625$, WSR test). In other words, the effect of complexity lies on the extensive margin (occurrence of a vote change), not on the intensive margin (intensity of the vote change).

Although complexity induces more vote changes, it hinders learning from experience. Indeed, the rise between the first and second attempts in the frequency of optimal votes (Table 5) is larger under Borda (+4.3%, from 49.2% to 53.5%) than under IRV (+3.4%, from 35.4% to 38.8%). Moreover, while the increase is statistically significant under Borda ($p = 0.0003$, WSR test), it is not statistically significant at the 5%-threshold under IRV ($p = 0.0682$). More generally, while under IRV there is no statistically significant change in the winning probabilities of the four candidates, we find under Borda significant increases in the winning probabilities of candidates A and B, and significant drops in the winning probabilities of candidates C and D (see Table A8 in Appendix E). Hence, although voters changed their

		IRV	Borda	*p-value*
Optimal	**1st attempt**	35.4%	49.2%	*0.0005*
	2nd attempt	38.8%	53.5%	*0.0001*
	p-value	*0.0682*	*0.0003*	
Sincere	**1st attempt**	20.0%	18.6%	*0.7145*
	2nd attempt	21.2%	18.3%	*0.3279*
	p-value	*0.6851*	*0.8875*	
Lifting	**1st attempt**	38.5%	28.0%	*0.0002*
	2nd attempt	34.9%	29.3%	*0.0303*
	p-value	*0.0842*	*0.4335*	
Conforming	**1st attempt**	23.4%	22.4%	*0.4177*
	2nd attempt	20.2%	18.8%	*0.3503*
	p-value	*0.034*	*0.0067*	

Table 5: Frequencies of votes consistent with each type of voting behavior: by attempt.

votes more frequently under IRV than under Borda, they seem to have had difficulties learning from experience under IRV, as vote changes have no significant effects on the election outcome, contrary to what happens under Borda.

To further validate the argument that complexity makes it hard for voters to figure out how best to vote and hinders learning from experience, we find that conditional on submitting an optimal vote in the first attempt, a voter cast the same vote in the second attempt 91.9% of the time under Borda compared to only 61.5% of the time under IRV. This observation strengthens the notion that while the optimal ballots submitted under Borda were the result of strategic voting behavior, this was not the case for a substantial fraction of the optimal ballots submitted under IRV.

Summing up:

RESULT 2. We have observed that:

1. Consistent with the argument that complexity makes it hard for people to vote strategically, voters change their vote between the two attempts more frequently 1) under IRV than under Borda, and 2) in elections involving four candidates than in elections involving only three candidates.

2. Complexity hinders learning from experience.

4.2.3. Payoffs

We have observed that the complexity of IRV impacts voters' behavior. We now investigate how this translates into voters' payoffs.

We can measure voters' payoffs in two ways. One is the *expected payoff* at the time when the voter submits his ballot, and thus before any potential tie between the candidates is broken. Another way is the *realized payoff* after the election winner is designated, and thus after any potential tie between the candidates is broken. The two measures are identical when there is no tie. In this section, we focus on the *expected payoff* (henceforth, *payoff*). It is worth mentioning that all the results obtained with the expected payoffs are robust if we consider instead the realized payoffs (see Table A9 in Appendix F).

We find that voters get overall lower payoffs under IRV than under Borda. In the first attempt, voters' payoffs are on average 23.21$ under IRV, compared to 28.33$ under Borda (Table 6).[47] The payoff difference between the two voting procedures is statistically significant ($p < 0.001$, WSR test) and repeats for each of the six vote profiles, except for *Lift-3* where the difference is not statistically significant ($p = 0.1978$, WSR test).

The payoff difference between IRV and Borda is even bigger in the second attempt (5.68$, compared to 5.12$ in the first attempt), as average payoffs increase between the two attempts less under IRV (+0.11$, from 23.21$ to 23.32$) than under Borda (+0.67$, from 28.33$ to 29.00$). Again, the payoff difference between IRV and Borda in the second attempt is statistically significant ($p < 0.001$, WSR test) and repeats for each of the six vote profiles except *Lift-3*. This increase between the two attempts in the payoff difference between IRV and Borda is consistent with our previous observation that the complexity of the voting procedure hinders learning from experience. Indeed, the average payoff under IRV remains stable between the two attempts (+0.11$; $p = 0.525$, WSR test), while the increase under Borda is statistically significant (+0.67$; $p = 0.0015$, WSR test).[48]

Two effects can explain the payoff difference between IRV and Borda.

- **Behavioral effect**, which is associated with voters' behavior. Voters may choose to cast different ballots under different voting procedures. This may happen, for instance, as a response to the differential complexity of voting procedures.

- **Mechanical effect**, which is associated with the determination process of the election winner

[47] The difference of 5.12$ is associated with a lower winning probability for candidate A under IRV compared to Borda (4.0% versus 27.6%; $p < 0.001$, WSR test) and higher winning probabilities for each of the other three candidates (Table A8 in Appendix E).

[48] The difference in payoff increase between IRV and Borda is associated with a differential change in the winning probabilities of the various candidates. Under Borda, the winning probabilities of A and B increase between the first and second attempts, while the winning probabilities of C and D decrease. By contrast, under IRV, there is no statistically significant change between the two attempts in candidates' winning probabilities. See Table A8 in Appendix E

from the set of submitted ballots. Specifically, different voting procedures may designate different election winners from the same set of ballots.

Vote profile	Voting procedure	1^{st} attempt	2^{nd} attempt	*p-value*
All profiles	**IRV**	23.21$	23.32$	*0.525*
	Borda	28.33$	29.00$	*0.002*
	p-value	*<0.001*	*<0.001*	
Lift-3	**IRV**	22.91$	23.10$	*0.251*
	Borda	22.60$	22.93$	*0.065*
	p-value	*0.1978*	*0.3779*	
Lift-4	**IRV**	17.00$	17.26$	*0.507*
	Borda	23.85$	24.33$	*0.420*
	p-value	*<0.001*	*<0.001*	
Lift-Overstate-4	**IRV**	20.53$	20.34$	*0.254*
	Borda	24.71$	25.14$	*0.095*
	p-value	*<0.001*	*<0.001*	
Sincere-Bury-3	**IRV**	30.48$	30.55$	*0.772*
	Borda	35.72$	36.01$	*0.172*
	p-value	*<0.001*	*<0.001*	
Sincere-Bury-4	**IRV**	24.10$	24.34$	*0.554*
	Borda	30.24$	31.75$	*0.026*
	p-value	*<0.001*	*<0.001*	
Sincere-3	**IRV**	24.25$	24.30$	*0.923*
	Borda	32.88$	33.85$	*0.055*
	p-value	*<0.001*	*<0.001*	

Table 6: Average expected payoffs.

We now isolate the behavioral effect by decomposing the payoff difference between IRV and Borda into the behavioral and mechanical effects. To do so, we first need to introduce extra notation. We denote by $\pi_r(v_{r'})$ the voter's payoff under voting procedure r from the ballot cast under voting procedure r', where $r, r' \in$ {IRV, Borda}. The payoff difference between Borda and IRV can then be written as

$$\Delta\pi = \pi_{Borda}(v_{Borda}) - \pi_{IRV}(v_{IRV}).$$

Adding and subtracting the payoff that the voter would have obtained under Borda if he had cast the ballot that he submitted under IRV, $\pi_{Borda}(v_{IRV})$, we obtain

$$\Delta\pi = \Big[\underbrace{\pi_{Borda}(v_{Borda}) - \pi_{Borda}(v_{IRV})}_{\text{behavioral effect}}\Big] + \Big[\underbrace{\pi_{Borda}(v_{IRV}) - \pi_{IRV}(v_{IRV})}_{\text{mechanical effect}}\Big].$$

The first bracketed expression, $\pi_{Borda}(v_{Borda}) - \pi_{Borda}(v_{IRV})$, measures the *behavioral effect* since the voting procedure (Borda) is the same in both terms, but the ballots are different (the one cast under Borda in the first term, and the ballot cast under IRV in the second term). Hence, this expression measures the payoff difference that results from the voter submitting potentially different ballots under the two voting procedures. The second bracketed expression, $\pi_{Borda}(v_{IRV}) - \pi_{IRV}(v_{IRV})$, measures the *mechanical effect* since the ballot is the one submitted under IRV in both terms, but the voting procedure used to determine the election winner is different (Borda in the first term, and IRV in the second term). Hence, this expression measures the payoff difference that results from the application of different voting procedures to determine the election winner from the same set of ballots.[49]

We find that 52% of the 5.12$ payoff difference between Borda and IRV during the first attempt is due to the behavioral effect, and the remaining 48% to the mechanical effect. We find furthermore that the rise in the payoff difference between the first and second attempts is entirely due to the behavioral effect, which explains a larger share (58%) of the 5.68$ payoff difference during the second attempt compared to the 5.12$ payoff difference during the second attempt. Indeed, the payoff difference resulting from the mechanical effect remains unchanged at 2.40$ between the two attempts ($p = 0.818$, WSR test), while the payoff difference resulting from the behavioral effect rises between the two attempts ($p = 0.024$, WSR test).

[49] To isolate the behavioral effect, we have added and subtracted the payoff that the voter would have obtained under Borda if he had cast the ballot that he submitted under IRV. Alternatively, we could have added and subtracted the payoff that the voter would have obtained under IRV if he had cast the ballot that he submitted under Borda. It is important to mention that this would have generated different values for the mechanical and behavioral effects. This happens because Borda captures more information about the ordering of candidates on the ballot than IRV does since Borda utilizes the full ordering of candidates to determine the election winner, which is not always the case for IRV (e.g., when a candidate immediately obtains a majority of the first rankings, IRV ignores how voters have ranked candidates down on their ballots). For this reason, we choose Borda as the voting procedure to isolate the behavioral effect since, contrary to IRV, Borda incorporates information about differences at every rank on the ballots submitted under IRV and Borda.

To better understand how complexity limits voters' ability to vote strategically, we compare a voter's payoff with the payoff that she would have gotten if, instead, she had cast the same ballot as the one that she submitted under the other voting procedure. If a voter behaves strategically, she should earn more, or at least as much, with the ballot that she submitted than with the ballot that she had cast under the other voting procedure. This is what we find in the case of Borda. Indeed, a voter's payoff at the first attempt is on average 28.33$ under Borda, but would have been only 25.65$ if the voter had cast the same ballot as the one that she submitted under IRV (Table 7). The gap is even larger at the second attempt (29$ vs. 25.72%). Differences are statistically significant at both attempts ($p < 0.001$, WSR test).

	IRV			**Borda**		
	Actual vote $\pi_{IRV}(v_{IRV})$	**Vote under Borda** $\pi_{IRV}(v_{Borda})$	*p-value*	**Actual vote** $\pi_{Borda}(v_{Borda})$	**Vote under IRV** $\pi_{Borda}(v_{IRV})$	*p-value*
1st attempt	23.21$	23.34$	*0.162*	28.33$	25.65$	*<0.001*
2nd attempt	23.32$	23.66$	*0.008*	29.00$	25.72$	*<0.001*
p-value	*0.525*	*0.003*		*0.002*	*0.738*	

Table 7: Average expected payoffs with the votes submitted under IRV and Borda.

However, we find the opposite in the case of IRV. A voter would, on average, have earned under IRV slightly more, not less, if she had submitted the same ballot as the one she submitted under Borda (23.34$ vs. 23.21$ at the first attempt, and 23.66$ vs. 23.32$ at the second attempt). While the difference is small and not statistically significant at the first attempt ($p = 0.162$, WSR test), it is slightly bigger and statistically significant at the second attempt ($p = 0.008$, WSR test). These observations further show that the complexity of the voting procedure makes it difficult for voters to figure out how best to vote.

If payoffs are, on average, lower under IRV compared to Borda, they are also more evenly distributed. Indeed, the standard deviation of voters' payoffs (defined for each voter as his average payoff over all six vote profiles) is substantially lower under IRV than under Borda (1.88 vs. 5.54 at the first attempt; 1.77 vs. 5.45 at the second attempt). Likewise, the Gini coefficient is lower under IRV than under Borda (0.04 vs. 0.11 at the first attempt; 0.04 vs. 0.10 at the second attempt). These differences repeat for each of the six vote profiles.

Summing up:

RESULT 3. We have observed that:

1. Payoffs are on average lower, but more equally distributed, under IRV than under Borda.

2. More than half of the payoff difference between IRV and Borda is accounted for by differences in

voters' behavior between the two voting procedures (*behavioral effect*). The rest is accounted for by the difference in the way IRV and Borda determine the election winner (*mechanical effect*).

3. While voters learn and increase their payoffs as they gain experience under Borda, they have difficulties doing so under IRV.

4. Voters would have received, on average, higher payoffs under IRV if they had cast the ballot that they submitted under Borda instead of the ballot they actually submitted. By contrast, voters would have received, on average, lower payoffs under Borda if they had cast the ballot they submitted under IRV instead of the ballot they actually submitted. This observation further highlights the challenge that arises from the complexity of a voting procedure, making it difficult for voters to comprehend how best to vote.

5. Determinants of voting behavior

In this section, we propose an empirical model that investigates the observed voting behavior, taking into account not only the previously mentioned factors but also voters' observable socio-demographic characteristics and their unobserved heterogeneity. The identification of the model parameters benefits from the controlled variation of the experimental within-subjects design. More precisely, election complexity varies because, firstly, half of the decisions are observed under either the IRV or the Borda voting procedures, and, secondly, within each voting procedure, half of the elections involve three candidates and the other half four candidates. Therefore, each of the 104 participants experienced throughout a series of 24 elections the different levels of complexity in a balanced way. Furthermore, the empirical model accounts for the panel structure of the data and will also look at the three sources of learning: firstly from voting a second time under the same vote situation, secondly, within the same voting procedure, and, thirdly, across voting procedures.

5.1. Empirical model

Our empirical model supposes that a voter selects a ballot to maximize his utility. He can choose one out of at most five unordered categories of voting behavior, referred to as category of voting behavior, or for short voting category, representing *Optimal, Sincere, Lifting*, *Conforming* behavior, and a last alternative that contains all other remaining *Unclassified* ballots. The literature on multinomial models refers to the categories a decision maker can choose from as *alternatives*. Given that we are in a voting context, to avoid confusion, we will refer to them in general as "voting categories," or "heuristics" when we refer to all other (non-optimal) vote categories except the *Optimal* ballot category. More formally, for each voter i, we associate with each possible ballot h in election t an indirect utility, U_{iht}, with $t \in \{1,\dots,24\}, h \in$ $\{$*Optimal, Sincere, Lifting, Conforming, Unclassified*$\}$. The utility comprises a deterministic component,

V_{iht}, and an error component, u_{iht}:

$$U_{iht} = V_{iht} + u_{iht}.$$

Voter i chooses in election t a ballot $y_{it} = h$ representing voting category h if this ballot provides the highest utility. Therefore,

$$\begin{aligned} Pr(y_{it} = h) &= Pr(U_{iht} \geq U_{igt}) \ \forall\, g \neq h \\ &= Pr(u_{igt} - u_{iht} \leq V_{iht} - V_{igt}) \ \forall\, g \neq h. \end{aligned}$$

The deterministic component,

$$V_{iht} = \mathbf{b}'_{iht}\rho_i + \mathbf{w}'_{it}\delta_h,$$

consists of time-variant and time-invariant variables. Those variables include voting category (alternative) specific regressors $\mathbf{b}_{iht}$ (e.g., the ballot characteristics), which effects are allowed to vary across voters with voter-specific random coefficients ρ_i. The variables also include voter (case) specific regressors $\mathbf{w}_{it}$ (e.g., their characteristics and behavior) and their particular situation (e.g., vote profile and voting procedure), which vary across voters and elections, but not across voting categories for the same election. The coefficients δ_h are fixed vote categories-specific coefficients. The determinants are explained in more detail in the following section.

The errors u_{iht} follow a type I extreme value distribution and are independent over voters, voting categories, and time. The voter-specific random effects relax the independence of irrelevant alternatives assumption of standard logit models as they allow unobserved heterogeneity and choices to be dependent over time. More precisely, the heuristic-specific coefficients $\rho_i = \rho + \zeta\iota_\mathbf{i}$ are composed of the vector of means ρ to be estimated along with the variance-covariance matrix ($\Sigma = \zeta\zeta'$). The voter individual random effects $\iota_i \sim \mathbf{N}(0, \mathbf{I})$ follow a multivariate normal distribution. Allowing for correlation across the random effects enables us to capture trade-offs that voters make between the characteristics of heuristics, e.g., voters might be more likely to accept (non-optimal) sincere or conforming ballots when they are more likely to be close to the expected gain maximizing ballot, or on the contrary, they might be insensitive to the expected gain of a particular ballot and blindly follow one heuristic.

The probability that a voter chooses a particular voting strategy in a particular election has no closed-form solution. Its integral over the mixing distribution of the random coefficients ρ_i needs to be calculated by simulation. For the integration by simulation, different random draws are taken from the multi-variate normal density of ρ_i for each voter and election. For the simulation, we use integration sequences of Halton draws, because they are deterministic and provide a better equidistant coverage over the domain of the integration space for each observation compared to standard random draws. Bhat (2001) and Train (2000) found Halton sequences in mixed logit models to reduce the simulation error in the estimated parameters and this even with a lower number of draws, which considerably increases the speed of the estimation compared to random draws.[50]

Variable names	Mean	Definitions
		Category specific ballot characteristics
Expected gains	25.70	The ballots ex-ante expected gain under the vote profile and voting procedure. Between 10 and 40 with a total of 13 distinct values.
HD Sincere	1.47	The ballot's Hamming distance to the sincere vote.
HD Conforming	1.88	The ballot's Hamming distance to the conforming vote.
		Case specific voting situation characteristics
Voting procedure	0.50	$= 1$ if IRV; $= 0$ if Borda.
4 candidates	0.50	$= 1$ if vote profile involves 4 candidates (*Lift-4, Sincere-Bury-4, Lift-Overstate-4*); $= 0$ if 3 candidates (*Lift-3, Sincere-Bury-3, Sincere-3*).
Second attempt	0.50	$= 1$ if ballot was cast the second time under the same vote profile; $= 0$ if first time.
Period: 1 to 6	3.5	Integer values (1 to 6) indicating the order of vote profiles under the same voting procedure.
Order	0.50	$= 1$ ballot was cast under second voting procedure;$= 0$ under first.
Multiple opt ballots		$= 1$ more than one ballot is optimal; $= 0$ otherwise.
LEEL	0.24	$= 1$ if participated in a session at the LEEL; $= 0$ at the CIRANO.

Table 8: Ballot characteristics and voting situation – Variable definitions and mean values for observed choices ($N = 2,496$).

5.2. *Determinants*

Ballot characteristics, $\mathbf{b}_{iht}$, vary over time, that is, by vote profile and voting procedure. They comprise the expected gain associated with the ballot ('Expected gains') and the ballot's closeness to two reference ballots that voters might take as starting points when making their voting choice, namely, the sincere ('HD Sincere') and the conforming ('HD Conforming') ballots, measured by the respective Hamming distance (see footnote 44). Table 8 summarizes the definition and mean values of the ballot character-

[50] One important advantage of Halton sequences is that they cover the domain of the mixing distribution relatively evenly, because of their deterministic nature, resulting in less variation of simulated probabilities over observations compared to when using standard random draws. A second advantage is that because of the way they are constructed, they reduce the variance in the log-likelihood function (Train, 2000).

Variable names	Mean	Definitions
		Case specific voter characteristics
BNT: 1	0.38	= 1 if Berlin Numeracy Test score = 1 (low numeracy).
BNT: 2	0.31	= 1 if Berlin Numeracy Test score = 2 (medium numeracy).
BNT: 3	0.32	= 1 if Berlin Numeracy Test score = 3 or 4 (high numeracy).
DT, DT^2	0.94	Linear decision time (in minutes) that a voter took to submit a ballot and its quadratic polynomial.
Attention	5.69	Instructions reading time (in minutes).
Female	0.47	= 1 if women; = 0 otherwise.
Education	0.41	= 1 if graduate degree (masters or PhD); = 0 otherwise.
Age 1: < 27	0.22	= 1 if less than 27 years of age.
Age 2: $[27-34]$ years	0.27	= 1 if between 27 and 34 years of age.
Age 3: $[35-44]$ years	0.25	= 1 if between 35 and 44 years of age.
Age 4: $[45+]$ years	0.26	= 1 if 45 years of age or older.
Giving to others (Giving)		Amount out of 100$ given to another person as dictator. Cost of giving vary.
Giving efficient	28.42	cost of giving 1$: 0.50
Giving	33.75	cost of giving 1$: 1.00
Giving inefficient	38.74	cost of giving 1$: 2.00.
Lie aversion		Is lying in general justified? measured on a 10 point Likert scale (Not at all= 1; Yes= 10).
Lying not acceptable	0.36	= 1 if "no" (on Likert scale= 1)
Lying slightly acceptable	0.26	= 1 if "slightly" (on Likert scale = 2).
Lying very acceptable	0.38	= 1 if "Yes" (on Likert scale = 3 and more).

Table 9: Voter characteristics – Variable definitions and mean values for voter-participants ($N = 104$).

istics. The vector containing voter (case)-specific variables, $\mathbf{w}'_{it}$, includes characteristics describing the voting situation and the voter's characteristics. The indicator variable 'Voting procedure' enters separately (with 1 for IRV and 0 for Borda) and in interaction with most voter-specific variables, permitting the effect of those variables to vary by the voting procedure. The indicator variable '4 candidates' is equal to 1 for vote profiles involving four candidates and 0 for those with three candidates. This variable captures the effect of an increase in environmental complexity due to an increase in the number of candidates. The experimental design allows and can control for learning as voters make two consecutive decisions for the same vote profile, and this for six consecutive vote profiles under the same voting procedure. The indicator variable 'Second attempt' designates the second decision under the same vote profile and same voting procedure. It captures learning within a vote profile when a voter is confronted with the same vote profile a second time. The variable 'Period: 1 to 6' is an integer vector counting from 1 to 6 the order in which vote profiles were presented to the voter, controlling for learning across vote profiles under the same voting procedure. The variable 'Order' indicates vote decisions under the second voting procedure, meaning the one used for the last 12 votes of the session. It controls for potential effects on the ballot choice originating from the order in which the voting procedures were presented to a participant. This variable also captures the impact of learning transfer from one voting procedure to the other. We include an indicator variable 'Multiple opt ballots' that captures if a vote profile under a certain voting procedure has more than one optimal ballot, which varies across vote profiles and might affect the probability of choosing an optimal ballot.[51] The last variable that we consider concerning the voting situation is the indicator variable 'LEEL' specifying the experimental laboratory in which the session was held (with 1 for LEEL and 0 for CIRANO).

Other voter-specific variables in $\mathbf{w}'_{it}$ are related to their individual characteristics and behavior (Table 9). The latter focuses on how voters adapt to the complexity of the voting procedure and the complexity of the environment. Voters can adapt to the complexity in two ways: first, by using their numeracy skills, and second, by taking more time to make a decision. A voter's numeracy skills are measured with the Berlin Numeracy test score ('BNT') as three indicator variables, one for each of the following scores: 1 (= low numeracy), 2 (= medium numeracy), and 3 or 4 (= high numeracy). The variable 'DT' is a continuous variable that measures the time in minutes that a voter took to make a decision. 'DT^2' is the square of the variable 'DT' allowing its effect to enter non-linearly. Likewise, the 'Attention' voters pay to an activity might affect their ballot choice, e.g., because they are more conscientious. We use the time in minutes that voters took to read the instructions as a measure of attention.

Individual time-invariant characteristics of voters include socio-demographic information and (social)

[51] Under Borda, two out of the six vote profiles, *Lift-4* and *Lift-Overstate-4*, have two optimal ballots each, whereas the other four have one. Under IRV, *Lift-4* and *Sincere-Bury-4* have two optimal ballots each, *Lift-Overstate-4* has three, and *Lift-4* six. See Table 1 for more details.

attitudes that were elicited in a post-experimental questionnaire. More precisely, the former comprise indicator variables for the gender of the voter ('Female') that is equal to one ($= 1$) if the voter identifies herself as a woman and zero ($= 0$) if the voter identifies himself as a man or prefers not to answer, four age groups (below 27; between 27 and 34; between 35 and 44; and 45 and older) and education, (master's degree or PhD $= 1$, bachelor degree or lower $= 0$). Eliciting social attitudes allows us to account for voting as a social activity, that might be associated with social norms, such as truth-telling or social preferences (Kube and Puppe, 2009; Messer et al., 2010). The variables 'Giving to others (inefficient),' 'Giving to others,' and 'Giving to others (efficient)' are the (hypothetical) share of 100$ that a voter would give to another person and that vary in their efficiency of giving implied by the price of giving. The first variable measures giving when giving is costly – specifically, giving 1$ costs 2$, the second when giving 1$ costs 1$, and the third when giving 1$ costs 0.50$. Another social attitude that might affect the ballot choice is honesty. In other words, wanting to adhere to a social norm of truth-telling or being averse to lying might be an obstacle to voting strategically. Even voters who understand strategic voting and can identify optimal ballots might refrain from casting an optimal ballot if it is not sincere because of the social norm to be honest and that condemns lying. The indicator variables 'Lying' measure the attitudes towards lying in general and whether it can be acceptable (no, slightly, yes).[52]

5.3. Voting categories

In our empirical analysis, we consider a voter's ballot choice as the choice of the particular category of voting behavior to which the ballot belongs. The first columns of Table 10 lists the observed proportions of the different voting behavior categories over all the vote profiles and attempts. Each category of voting behavior comprises more than 15% of all submitted ballots. More precisely, *Optimal*: 44%, *Sincere*: 20%, *Lifting*: 33%, *Conforming*: 21%, and *Unclassified*: 18%.

Note that *Optimal* includes 40% of all sincere ballots that are observed in the *Sincere-3* vote profile, for both voting procedures, in the *Sincere-Bury-3/4* vote profiles for IRV, and 54% of lifting ballots in the *Lift-3/4* vote profiles for both voting procedures and in *Lift-Overstate-4* for IRV. However, sincere and lifting votes are not optimal in the other vote profiles. *Unclassified* and *conforming* ballots are never optimal. The conforming and lifting ballots coincide in the vote profile *Sincere-Bury-3*. There is, by definition, no overlap for unclassified ballots with other categories of voting behavior.

For a given vote profile, categories are singular with the exceptions of *Optimal* and *Unclassified*. There is more than one optimal ballot in *Lift-3* and *Sincere-Bury-3* for IRV and in *Lift-4* and *Lift-Overstate-4* for both voting procedures. The category containing all *Unclassified* ballots comprises multiple ballots

[52] The question in the post-experimental survey measuring Lying aversion is based on the World Value Survey question on the acceptability of cheating and lying (Inglehart et al., 2014). More precisely, voters indicated on a scale of 1 to 10 whether lying in general is justified. We re-code indicator variables as "No" = 1, "Slightly" = 2, and "Yes" = 3 or more.

Voting categories	Ballots (%)					
	Each category			Mutually exclusive categories		
	IRV	Borda	All	IRV	Borda	All
Optimal	37.10	51.36	44.23	37.10	51.36	44.23
Sincere	20.59	18.43	19.51	7.61	8.89	8.25
Lifting	63.70	28.69	32.69	19.71	10.26	14.98
Conforming	21.79	20.59	21.19	13.14	15.70	14.42
Unclassified	22.44	13.78	18.11	22.44	13.78	18.11
Total				100	100	100
N. obs.	1,248	1,248	2,496	1,248	1,248	2,496

Table 10: Classification of the observed ballot choices. ($N = 2,496$: 6 vote profiles x 2 attempts x 2 voting procedures x 104 participants).

Columns "*Each category*," count all ballots that correspond to a particular vote category. Columns "*Mutually exclusive categories*," count optimal sincere and optimal lifting ballots as *Optimal*, and lifting conforming ballots as *Lifting* in the "*Mutually exclusive categories*" column.

for every vote profile.

The first set of columns in Table 10 presents the proportions of the ballots that are associated with a particular voting category. Because, for certain vote profiles, some ballots belong to more than one category, they are accounted for twice. For example, *Sincere* and *Lifting* include optimal sincere and lifting votes as well as ballots that correspond at the same time to the *Lifting* and *Conforming* heuristic. The second set of columns reports proportions for mutually exclusive voting categories. In this column, *Sincere* and *Lifting* heuristics do not include optimal votes and ballots that also correspond to the *Lifting* or *Conforming* heuristic .

Due to the overlap of some voting categories under certain vote profiles and voting procedures, our data set has, by construction, an unbalanced number of voting categories from which a voter can choose. For example, voters who cast a ballot under the vote profile *Lift-Overstate-4* can choose under Borda one out of all five categories, whereas under IRV *lifting* is optimal and therefore, there are only four categories (*Optimal, Sincere, Conforming, Unclassified*) out of which the voter can choose. Table 11 lists the number of possible categories and the resulting number of observations for each voting procedure and vote profile. As for each of the 2,496 actual voting choices, our empirical model takes into account the other categories that were available at the moment of choice, resulting in a total of 10,192 data entries.

Profiles	Possible categories of voting behavior IRV	Borda	Total
Lift-3	4	4	
Lift-4	4	4	
Lift-Overstate-4	4	5	
Sincere-Bury-3	3	4	
Sincere-Bury-4	4	5	
Sincere-3	4	4	
Sum of categories	23	26	49
Total # of obs (x2 attempts x104 participants)	4,784	5,408	10,192

Table 11: Total of possible numbers of vote categories by vote profile and voting procedure. Unbalanced categories by construction of the experimental design.

5.4. Estimation results

Our model, fitting data from 104 voters across 24 voting decisions, has produced precise estimation results. As explained in detail in Section 5.3, our panel is unbalanced, with voters choosing from between three and five voting categories in each voting decision, resulting in a total of 10,192 observations. The integral of the choice probability allowing for voter-specific random effects was approximated using 200 Halton draws. The estimation results are presented in Tables 12, 13, and 14.

The estimated parameters cannot be directly interpreted because the actual probabilities of choosing a particular category of voting behavior are a nonlinear function of the estimated parameters. We will discuss first some impressions from the parameter estimates, highlighting the direction of their influence on the choice of voting heuristics compared to the baseline category, which in our case is the *Optimal* vote category. Subsequently, we will present a marginal effects analysis based on the model estimates, allowing us to draw precise conclusions about the average marginal effect of certain variables.

Table 12 presents the parameter estimates for heterogeneous individual responses to the characteristics of ballot categories. The estimated mean of the random coefficients on the expected gain of a ballot is 0.216 with a p-value below 0.001%, indicating that the probability of choosing a ballot is significantly positively linked to its expected gain. The estimated standard deviation of these coefficients is 0.234, indicating some heterogeneity in how individuals react to incentives in an election context. The estimated means of the reactions towards the Hamming distance to the sincere or conforming ballot are

Expected gains	0.216***
HD sincere	-0.036
HD conforming	-0.098
σ(Expected gains)	0.234
σ(HD sincere)	0.382
σ(HD conforming)	0.346
ρ(Expected gains,HD sincere)	0.062
ρ(Expected gains,HD conforming)	0.722***
ρ(HD sincere,HD conforming)	0.005

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 12: Estimation results – Voting behavior parameters, standard deviations and correlations.

negative, suggesting that, on average, voters are less likely to choose a ballot that is further away from the sincere or conforming ballot and corroborating that those two heuristics might serve as a reference point. However, these effects are not significant at conventional levels. We observe a strong and positive correlation between the random effects of the expected gains and the Hamming distance of a conforming ballot, indicating that people who are sensitive to the expected gains of a ballot are less likely to use the conforming ballot as a reference point.

Tables 13 and 14 present the parameter estimates on voter-specific variables for each heuristic that need to be interpreted with respect to the baseline category, *Optimal*. (Non-optimal) lifting ballots are more likely under the more complex voting procedure (IRV). There are no other general effects of the procedural complexity on the other heuristics. Not surprisingly, when there is more than one optimal ballot (Multiple opt. ballots), voters are more likely to choose an optimal ballot, significantly reducing the probability of choosing any heuristic and confirming the importance of control in the empirical model for situations with multiple optimal ballots.

A larger environmental complexity due to more candidates to choose from, increases the probability of observing a lifting ballot or one from the *Unclassified* category with respect to the optimal vote. Further, the likelihood of sincere and conforming votes increases with environmental complexity, but only under Borda. We will explore the effects of complexity on the choice of voting categories in greater detail below.

We observe learning, but mostly under the less complex Borda procedure. Notably, the probability of choosing any heuristic decreases with repetition (Attempt) and over time within the same voting procedure (Period). We observe mixed effects for learning across voting procedures (Order), where having experienced the Borda procedure first, reduced significantly the usage of conforming and unclassified

	Sincere		Lifting	
IRV	1.307		1.793***	
Multiple opt. ballots	-1.302***		-1.715***	
Attention	0.001		0.001	
Interaction effects	IRV	Borda	IRV	Borda
4 candidates	0.039	1.582***	1.313**	1.010**
Attempt	-0.165	-0.876***	-0.303	-0.466
Period	0.134	-0.188**	-0.145*	-0.204**
Order	0.002	-0.754*	-0.308	-0.080
LEEL	0.169	-0.436	0.252	-1.153*
BNT=2	-0.032	-0.325	0.668	-0.246
BNT=3	-0.578	0.372	0.128	0.079
DT	0.215	-2.199***	0.099	-0.992*
DT^2	-0.098	0.292***	-0.004	0.158*
Lying slightly acceptable	-0.030	0.264	-0.368	0.082
Lying very acceptable	-0.880*	-0.145	-0.200	-0.563
Giving efficient	-0.001	0.008	-0.010	-0.005
Giving	-0.011	-0.008	-0.010	0.022
Giving inefficient	-0.011	0.008	0.005	-0.004
Female	-0.667	0.633	-0.045	0.571
Age 2: $27-34$	-0.693	1.184*	0.402	1.186*
Age 3: $35-44$	-0.601	1.426*	0.023	1.703**
Age 4: >45	-0.404	2.307***	0.462	2.377***
Education	-0.517	0.325	0.017	0.270

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 13: Estimation results – Voting situation and voter characteristics (case based) parameters for the following voting categories: *Sincere* and *Lifting* with *Optimal* as baseline category. ($N = 10,192$)

	Conforming		Unclassified	
IRV	-0.163		0.744	
Multiple opt. ballots	-1.092***		-1.960***	
Attention	0.002		0.000	
Interaction effects	IRV	Borda	IRV	Borda
4 candidates	-0.252	1.192**	1.669***	3.151***
Attempt	-0.220	-1.036***	0.180	-0.453*
Period	-0.031	-0.486***	0.054	-0.376***
Order	-0.897**	-0.166	-0.702**	-0.129
LEEL	0.325	-0.536	-0.196	-0.520
BNT=2	0.361	0.627	-0.218	-0.419
BNT=3	0.472	0.138	-0.378	-0.320
DT	0.082	-2.034***	0.915***	-0.710*
DT^2	-0.004	0.276***	-0.110**	0.087
Lying slightly acceptable	-0.280	0.296	-0.314	-0.345
Lying very acceptable	-0.254	-0.424	-0.153	-0.745*
Giving efficient	-0.005	-0.002	-0.012	-0.001
Giving	0.031*	0.038*	0.006	0.024
Giving inefficient	-0.022	-0.007	0.001	-0.003
Female	0.722*	1.139**	0.111	0.868*
Age 2: $27-34$	0.367	0.060	0.178	0.450
Age 3: $35-44$	0.802	2.278**	0.365	2.110***
Age 4: >45	1.188*	3.218***	0.898***	1.973***
Education	-0.915**	-0.242	-0.873***	-0.220
Nobs	**10,192**			

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Table 14: Estimation results – Voting situation and voter characteristics (case based) parameters for the following voting categories: *Conforming* and *Unclassified* with *Optimal* as baseline category. ($N = 10,192$)

ballots compared to the optimal ballots. Having experienced IRV first reduces the sincere ballots in favor of the optimal ballots.

Under Borda, taking more time (DT and DT^2) to submit a ballot increases the likelihood of submitting an optimal ballot, reducing the chances of submitting any other non-optimal vote. Under IRV, on the contrary, thinking longer does not at all affect the use of any non-optimal heuristic (*Sincere, Lifting, Conforming*).

We observe some subgroups of the population to be more likely to use certain vote categories. Notably, older voters tend to be more likely to use non-optimal vote categories, and women are more likely to submit conforming and unclassified ballots instead of an optimal ballot. However, we find these effects of background characteristics on vote decisions only under the less complex Borda voting procedure, suggesting that the deliberate choice of optimal ballots was less easy for young voters under the more complex IRV procedure. This suggestion seems to be corroborated by the observation that voters with higher education are more likely to choose an optimal vote by reducing the chance of choosing a conforming or unclassified ballot.

Finally, we find relations between the heuristics and voters' social attitudes. People who consider lying in general to be acceptable submit less often sincere (under IRV) and unclassified (under Borda) ballots in favor of the optimal ballot. Voters who give more to another person in a hypothetical standard dictator game are more likely to submit a conforming ballot than an optimal ballot, suggesting a relation between choosing this category and social preferences.

In the following, we examine the marginal effects of certain category- and voter-specific variables on the probability of selecting a specific category of voting behavior. This analysis allows us to further explore their overall effects.

Incentives

We are first interested in understanding how incentives affect the choice across categories of voting behavior. Table 15 reports how an increase in expected gains of a (row) category by 1 unit affects the probability that another (column) category will be chosen. For example, an increase of the expected payoff from choosing an optimal ballot by 1$, increases the probability that an optimal ballot is cast by 3.4%. The table also tells us where this increase comes from, namely from a large decrease in the Sincere and Lifting heuristics by 0.9% and 1.1%, respectively, followed by a smaller decrease in the Conforming heuristics (0.3%) and a larger decrease in the Unclassified category (1.1%).[53] We also observe an increase in the likelihood of choosing any of the non-optimal sincere or lifting ballots, albeit smaller, when the expected gain of those ballots increases. This can be explained by the fact that even

[53] Remember that contrary to the other categories that are singleton sets, the *Unclassified* category contains multiple ballots. This makes it mechanically more likely to observe a decrease in this category.

if their expected gain increases, there is still a better option to choose, the optimal ballot. Furthermore, those results indicate an even lower sensitivity to an increase in expected gains for the conforming ballot and complete insensitivity for unclassified ballots. We interpret the increase in the likelihood of choosing the *Sincere*, *Lifting*, or *Conforming* heuristic when its respective expected gain increases, to suggest ulterior motives, such as an aversion to not revealing their preferences truthfully, reducing cognitive costs, or conforming to a potential social norm. We explore these possible explanations separately.

	Optimal	*Sincere*	*Lifting*	*Conforming*	*Unclassified*
Optimal	0.034***	-0.009***	-0.011***	-0.003***	-0.011***
Sincere		0.012***	-0.001***	0.000	-0.002***
Lifting			0.015***	0.000***	-0.004***
Conforming				0.005***	-0.001
Unclassified					0.017

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 15: Marginal effects of expected gains.

In summary, we observe that voters react to incentives when making a vote decision. They increase the probability of choosing a ballot by decreasing that of other heuristics when the expected gain of this particular category increases. We observe more sensitivity to a change in the expected gain of optimal ballots and less sensitivity for conforming ballots.

Category of voting behavior	Predicted means by procedure						
				3 Candidates		4 Candidates	
	Both	IRV	Borda	IRV	Borda	IRV	Borda
Optimal	0.463	0.406	0.515	0.462	0.618	0.336	0.384
Sincere	0.085	0.099	0.079	0.113	0.055	0.081	0.094
Lifting	0.143	0.174	0.115	0.139	0.112	0.203	0.095
Conforming	0.138	0.123	0.154	0.160	0.159	0.088	0.136
Unclassified	0.171	0.198	0.136	0.126	0.056	0.292	0.291

Table 16: Predicted shares for each category of voting behavior in general and by different dimensions of complexity: procedural complexity (IRV vs. Borda) and environmental complexity (number of candidates).

Category of voting behavior	Complexity		Procedural complexity by environment		Environmental complexity by procedure	
	Procedural (IRV vs. Borda)	Environmental (4 vs. 3 candidates)	(IRV 3 vs. Borda 3)	(IRV 4 vs. Borda 4)	(IRV 4 vs. IRV 3)	(Borda 4 vs. Borda 3)
Optimal	-0.110***	-0.177***	-0.156***	-0.048	-0.126***	-0.233***
Sincere	-0.021	0.013	0.058**	-0.013	-0.032	0.039***
Lifting	0.059***	0.019	0.027	0.108***	0.064*	-0.018
Conforming	-0.032***	-0.052***	0.001	-0.048***	-0.071***	-0.022
Unclassified	0.062***	0.197***	0.070***	0.001	0.165***	0.235***

Table 17: Marginal effects of an increase in complexity caused by the voting procedure and the number of candidates on voting behavior.

Complexity

Tables 16 and 17 summarize mean shares of the marginal effects of the two dimensions of complexity that we study here: the procedural complexity of the voting procedure and the environmental complexity induced by the number of candidates. These statistics provide a good fit for the observed shares. Overall, optimal ballots are chosen most often (46%), followed by unclassified ballots (17%) as the second most frequently chosen ballot category, followed immediately by the lifting and conforming ballots (14% each) and, finally, with the sincere ballot as the least often chosen category of voting behavior (9%). This ranking is similar under each voting procedure. However, we observe that voters chose optimal votes more often under Borda than IRV. The more complex IRV procedure leads to a decrease of optimal votes by 11% compared to Borda. Voters react with a 17.7% decrease in optimal votes more intensively to an increase in the complexity of the environment. This reduction in optimal votes is mainly driven by a decrease in optimal votes under Borda (23.3%) compared to IRV (12.6%).

We had found limited support for the argument that voters tend to submit sincere votes when complexity increases. The limited nature of this support is further strengthened here. We find an increase of 3.9% in sincere votes when the environmental complexity increases but only within the simpler Borda procedure. We find an increase of 5.8% in sincere votes when the procedural complexity increases, but again only for the simple environment with three candidates. However, there are no changes in the use of (non-optimal) sincere votes when the procedural complexity is high, meaning within IRV, or between voting procedures when the environmental complexity is already high, meaning within elections with four candidates.

Voters respond to an increase in the complexity of the voting procedure with an almost equal increase in the use of (non-optimal) lifting and unclassified ballots. Thereby, the 5.9% increase in lifting ballots is mainly due to an increase in procedural complexity when environmental complexity is already high by 10.8%. The increase in unclassified ballots with procedural complexity comes from the less complex environment. What is, however, striking is the 19.7% increase of unclassified ballots when the complexity of the environment increases. This is the case for 16.5% under IRV and even 23.5% under Borda. Finally, we note a smaller decrease of the conforming vote by 3% under the more complex procedure and by 5% when the number of candidates increases. These effects are exclusively driven by a decrease in the use of the conforming vote under IRV with four candidates. The ratio of conforming votes remains essentially constant under Borda, regardless of the number of candidates.

Finally, the change of the share in unclassified ballots depends on the type of complexity. There is a small increase of 6% of those votes when complexity is procedural, compared to an almost 20% increase in the case of environmental complexity.

Summing up:

RESULT 4. We have identified several factors that influence voting behavior in elections with varying levels of complexity.

1. Voters tend to cast optimal ballots less frequently in more complex environments, characterized by IRV or involving more candidates. In contrast to the simpler Borda procedure, where factors like learning or taking more time to choose a ballot improve the likelihood of submitting an optimal ballot, neither of those factors has an effect under IRV.

2. Voters who follow the Sincere or Conforming heuristic submit their ballots swiftly.

3. While highly educated voters are more likely to cast an optimal ballot under IRV, experiential learning does not improve the understanding of how to vote optimally. This contrasts with the less complex Borda, where voters learn how to vote in their best interest.

4. Even when controlling for numeracy skills and education, older voters are less prone to strategic voting, opting more frequently for the sincere, lifting, conforming, and unclassified ballots.

5. We also observe some evidence suggesting that personal values play a role in selecting voting strategies. For example, less lie-averse voters are more inclined to cast ballots consistent with strategic voting behavior. And those with stronger social preferences are more likely to submit ballots that conform to others' votes.

6. We observe heterogeneous responses to an increase in one or both of the two dimensions of complexity that we study here: the procedural complexity and the environmental complexity induced by the number of candidates. Firstly, optimal votes decrease substantially. Secondly, this decrease can be explained by voters submitting more often lifting votes when procedural complexity increases and ballots in the *Unclassified* category when the environmental complexity increases.

6. Conclusion

Advocacy groups in several democracies, like the USA, Canada, and Great Britain, are campaigning actively in favor of electoral reforms that would lead to the adoption of IRV. During the last two decades, they have managed to convince citizens and lawmakers in several places, especially in the USA, to hold political elections under IRV. They argue that one advantage of this voting procedure is that it discourages people from voting strategically and induces them to vote sincerely, meaning to reveal their true preference ranking of the candidates. This assertion can be justified on the basis that IRV is such a complex voting procedure that it is hard for voters to figure out how to vote strategically, and that voters react to the complexity of voting strategically by resorting to sincere voting. Our paper examines this argument. More specifically, we investigate whether complexity does indeed prevent voters from voting strategically. If so, does complexity induce people to vote sincerely? Or do voters react to complexity by adopting other voting heuristics and, if so, which ones?

To address these questions, we designed and ran a laboratory experiment. We capture the effect of complexity on voters' behavior in two ways: firstly by holding elections under IRV and Borda (*procedural complexity*), and secondly by varying the number of candidates (*environmental complexity*). Although both IRV and Borda call for voters to rank candidates, it is more complex to figure out how to vote strategically under IRV than under Borda. Hence, we can analyze the impact of complexity on individuals' voting behavior by contrasting how voters behave under these two voting procedures. Likewise, strategic voting is arguably more complex in elections involving more candidates.

We find some support for the argument that complexity may prevent people from voting strategically. However, we also find that the prevention of strategic voting has drawbacks, as it results in worse outcomes for voters and impairs their capacity to learn from experience. Indeed, in our experiment, participants obtained on average higher payoffs in elections conducted using Borda as opposed to IRV. Even more concerning, participants' payoffs would have been higher under IRV had participants cast under this voting procedure the ballot they submitted under Borda.

Our findings also lend support to the claim that voters react to the difficulty of voting strategically by resorting to voting sincerely. However, this support is limited and needs qualifications, as only a small fraction of voters choose to do so. Additionally, we find that voters react to complexity less often by voting sincerely than by adopting a heuristic of *Lifting*, which consists of inverting the ordering of their two top preferred candidates. We find furthermore that some voters adopt a heuristic of *Conforming* – which consists of submitting a ballot that aligns with those cast by the other voters – as frequently as they vote sincerely. However, contrary to *Sincere* and *Lifting*, we find that voters do not adopt the *Conforming* heuristic as a reaction to the complexity of the voting procedure.

We have left for future research several robustness checks and extensions of this work. Firstly, our paper concentrates on examining the impact of *computational* complexity on voters' behavior, meaning the complexity for a voter to identify an optimal ballot. In future research, we aim to explore the impact of adding *strategic* complexity that concerns *strategic uncertainty*, meaning the complexity of accurately anticipating candidates' winning prospects, on voters' behavior. If introducing strategic complexity leads to changes in voting behavior, this could have significant policy implications concerning the information given to voters before the election, including pre-election polls. Secondly, in our study, participants were compelled to provide a complete ranking of the candidates, like Australian voters are required to do. However, other places that utilize IRV offer voters the option to submit a partial ranking of the candidates. This takes the form of allowing voters to choose the number of candidates they rank (referred to as *ballot truncation*) or permitting them to rank only a limited number of candidates, e.g., as in the 2021 New York City Democratic primary election, in which voters were allowed to rank up to five out of the thirteen candidates. As Dellis et al. (2011) show, allowing ballot truncation can affect significantly voters' behavior and election outcomes. In future research, we aim to explore how ballot truncation and restrictions on the number of candidates that voters can rank on their ballot influence voters' behavior.

References

Bartholdi, J. M. and Orlin, J. B. (1991). Single transferable vote resists strategic voting. *Social Choice Welfare*, 8:341–354.

Bartholdi III, J. J., Tovey, C. A., and Trick, M. A. (1989). The computational difficulty of manipulating an election. *Social Choice and Welfare*, 6:227–241.

Bassi, A. (2015). Voting systems and strategic manipulation: An experimental study. *Journal of Theoretical Politics*, 27(1):58–85.

Bhat, C. R. (2001). Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. *Transportation Research Part B: Methodological*, 35(7):677–693.

Black, D. (1976). Partial justification of the borda count. *Public Choice*, 28:1–15.

Bol, D., Dellis, A., and Oak, M. (2016). Comparison of voting procedures using models of electoral competition with endogenous candidacy. In Gallego, M. and Schofield, N., editors, *The Political Economy of Social Choices*, pages 21–54. Springer.

Bol, D. and Verthé, T. (2021). Strategic voting versus sincere voting. In Redlawsk, D., editor, *Oxford Research Encyclopedia of Political Decision Making*. Oxford University Press.

Bossaerts, P. and Murawski, C. (2017). Computational complexity and human decision-making. *Trends in Cognitive Sciences*, 21(12):917–929.

Bouton, L. and Castanheira, M. (2012). One person, many votes: Divided majority and information aggregation. *Econometrica*, 80(1):43–87.

Buisseret, P. and Prato, C. (2023). Politics transformed? electoral competition under ranked choice voting.

Callander, S. (2005). Duverger's hypothesis, the run-off rule, and electoral competition. *Political Analysis*, 13(3):209–232.

Chamberlin, J. (1985). An investigation into the relative manipulability of four voting systems. *Behavioral Science*, 30:195–203.

Chamberlin, J., Cohen, J., and Coombs, C. (1984). Social choice observed: Five presidential elections of the american psychological association. *Journal of Politics*, 46:479–502.

Cokely, E. T., Galesic, M., Schulz, E., Ghazal, S., and Garcia-Retamero, R. (2012). Measuring risk literacy: The berlin numeracy test. *Judgment and Decision Making*, 7(1):25–47.

Conitzer, V. and Walsh, T. (2016). Barriers to manipulation in voting. In Brandt, F., Conitzer, V., Endriss, U., Lang, J., and Procaccia, A., editors, *Handbook of Computational Social Choice*, pages 127–145. Cambridge University Press.

Cox, G. (1997). *Making Votes Count*. Cambridge University Press, Cambridge.

Dellis, A., D'Evelyn, S., and Sherstyuk, K. (2011). Multiple votes, ballot truncation and the two-party system: An experiment. *Social Choice and Welfare*, 37:171–200.

Dellis, A., Gauthier-Belzile, A., and Oak, M. (2017). Policy polarization and strategic candidacy in elections under the alternative vote rule. *Journal of Institutional and Theoretical Economics*, 173(4):565–590.

Donovan, T., Tolbert, C., and Harper, S. (2022). Demographic differences in understanding and utilization of ranked choice voting. *Social Science Quarterly*, 103:1539–1550.

Doron, G. and Kronick, R. (1977). Single transferable vote: An example of a perverse social choice function. *American Journal of Political Science*, 21(2):303–311.

Durand, F. (2023). Coalitional manipulation of voting rules: Simulations on empirical data. *Constitutional Political Economy*, 34:390–409.

Eggers, A. C. and Nowacki, T. (2024). Susceptibility to strategic voting: A comparison of plurality and instant-runoff elections. *Journal of Politics*, 86(2).

Eggers, A. C. and Vivyan, N. (2020). Who votes more strategically? *American Political Science Review*, 114(2):470–485.

Elkind, E., Faliszewski, P., and Slinko, A. (2012). Rationalizations of condorcet-consistent rules via distances of hamming type. *Social Choice and Welfare*, 39(4):891–905.

Elkind, E., Grandi, U., Rossi, F., and Slinko, A. (2020). Cognitive hierarchy and voting manipulation in k-approval voting. *Mathematical Social Sciences*, 108:193–205.

Esponda, I. and Vespa, E. (2014). Hypothetical thinking and information extraction in the laboratory. *American Economic Journal: Microeconomics*, 6(4):180–202.

Farrell, D. and McAllister, I. (2006). *The Australian Electoral System: Origins, Variations and Consequences*. University of New South Wales Press, Sydney, Australia.

Forsythe, R., Myerson, R., Rietz, T., and Weber, R. (1993). An experiment on coordination in multicandidate elections: The importance of polls and election histories. *Social Choice and Welfare*, 10:223–247.

Forsythe, R., Myerson, R., Rietz, T., and Weber, R. (1996). An experimental study of voting rules and polls in three-candidate elections. *International Journal of Game Theory*, 25:355–383.

Gibbard, A. (1973). Manipulation of voting schemes: A general result. *Econometrica*, 41(4):587–601.

Granic, D.-G. (2017). The problem of the divided majority: Preference aggregation under uncertainty. *Journal of Economic Behavior & Organization*, 133:21–38.

Greiner, B. (2015). Subject pool recruitment procedures: organizing experiments with orsee. *Journal of the Economic Science Association*, 1:114–125.

Grether, D. M. and Plott, C. R. (1979). Economic theory of choice and the preference reversal phenomenon. *The American Economic Review*, 69(4):623–638.

Grimm, V. and Mengel, F. (2012). An experiment on learning in a multiple games environment. *Journal of Economic Theory*, 147(6):2220–2259.

Hamming, R. W. (1950). Error detecting and error correcting codes. *The Bell System Technical Journal*, 29(2):147–160.

Harrison, G. W. and McDaniel, T. (2008). Voting games and computational complexity. *Oxford Economic Papers*, 60(4):546–565.

Herzberg, R. Q. and Wilson, R. K. (1988). Results on sophisticated voting in an experimental setting. *The Journal of Politics*, 50(2):471–486.

Inglehart, R., Haerpfer, C., Moreno, A., C. Welzel, K. K., Diez-Medrano, J., Lagos, M., Norris, P., Ponarin, E., and (eds.), B. P. (2014). World values survey: Round Two - Country-Pooled Datafile, Version: www.worldvaluessurvey.org/WVSDocumentationWV2.jsp. Madrid: JD Systems Institute.

Kawai, K. and Watanabe, Y. (2013). Inferring strategic voting. *American Economic Review*, 103(2):624–662.

Kube, S. and Puppe, C. (2009). (When and how) do voters try to manipulate? Experimental evidence from borda elections. *Public Choice*, 139:39–52.

Laslier, J.-F. (2016). Heuristic voting under the alternative vote: The efficiency of "sour grapes" behavior. *Homo Oeconomicus*, 33:57–76.

Lichtenstein, S. and Slovic, P. (1971). Reversal of preferences between bids and choices in gambling decisions. *Journal of Experimental Psychology*, 89:46–55.

Loewen, P. J., Hinton, K., and Sheffer, L. (2015). Beauty contests and strategic voting. *Electoral Studies*, 38:38–45.

Messer, K. D., Poe, G. L., Rondeau, D., Schulze, W. D., and Vossler, C. A. (2010). Social preferences and voting: An exploration using a novel preference revealing mechanism. *Journal of Public Economics*, 94(3):308–317.

Murawski, C. and Bossaerts, P. (2016). How humans solve complex problems: The case of the knapsack problem. *Scientific Reports*, 6:34851.

Myatt, D. P. (2007). On the theory of strategic voting. *The Review of Economic Studies*, 74(1):255–281.

Myerson, R. (1993a). Effectiveness of electoral systems for reducing government corruption: A game theoretic analysis. *Games and Economic Behavior*, 5:118–132.

Myerson, R. (1993b). Incentives to cultivate favored minorities under alternative electoral systems. *American Political Science Review*, 87(3):856–869.

Nielsen, K. and Rehbeck, J. (2022). When choices are mistakes. *American Economic Review*, 112(7):2237–2268.

Oprea, R. (2020). What makes a rule complex? *American Economic Review*, 110(12):3913–3951.

Rae, D. W. (1967). *The political consequences of electoral laws*. Yale University Press, New Haven.

Riker, W. H. (1981). A confrontation between the theory of social choice and the theory of democracy. In Braham, R. L., editor, *Social Justice*, pages 95–119. Martinus Nijhoff Publishing, Boston/The Hague/London.

Rubinstein, A. (2016). A typology of players: Between instinctive and contemplative. *The Quarterly Journal of Economics*, 131(2):859–890.

Saltsman, M. and Paxton, R. (2021). Start spreading the news—ranked-choice voting is a mess; new yorkers won't know for weeks who won the june 22 democratic primary for mayor. *Wall Street Journal*.

Santucci, J. (2021). Variants of ranked-choice voting from a strategic perspective. *Politics and Governance*, 9(2):344–353.

Satterthwaite, M. A. (1975). Strategy-proofness and arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions. *Journal of Economic Theory*, 10(2):187–217.

Train, K. (2000). Halton sequences for mixed logit. *UC Berkeley: Department of Economics Working Paper series, Institute for Business and Economic Research.*

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*, 35(3):435–472.

Wilcox, N. T. (1993). Lottery choice: Incentives, complexity and decision time. *The Economic Journal*, 103(421):1397–1417.

Appendix

A. Design

Vote profile	Optimal		Sincere	Lifting	Conforming
	IRV	**Borda**			
Lift-3	B–	BAC	ABC	BAC	CBA
Lift-4	B–	B–D	ABCD	BACD	DBAC
Lift-Overstate-4	BACD, BC–	B–C	ABCD	BACD	CBAD
Sincere-Bury-3	ABC	ACB	ABC	BAC	BAC
Sincere-Bury4	AB–	ACBD	ABCD	BACD	DBAC
Sincere-3	ABC	ABC	ABC	BAC	CAB

Table A1: Ballots consistent with the different categories of voting behavior

	Winning probabilities for candidates								Voting strategies/heuristics					observed frequency	
	Borda				*IRV*				Optimal		Sincere	Lifting	Conforming		
Ballot	*A*	*B*	*C*	*D*	*A*	*B*	*C*	*D*	*Borda*	*IRV*				*Borda*	*IRV*
	vote profile: Lift-3 (IRV: B- ; Borda: BAC)														
ABC	0	0	1	-	0	1/4	3/4	-	0	0	1	0	0	11	40
ACB	0	0	1	-	0	0	1	-	0	0	0	0	0	4	1
BAC	0	1/2	1/2	-	0	1/2	1/2	-	1	1	0	1	0	115	85
BCA	0	0	1	-	0	1/2	1/2	-	0	1	0	0	0	14	20
CAB	0	0	1	-	0	0	1	-	0	0	0	0	0	10	25
CBA	0	0	1	-	0	0	1	-	0	0	0	0	1	54	37
	vote profile: Lift-4 (IRV: B- ; Borda: B-D)														
ABCD	0	1/2	0	1/2	0	1/4	0	3/4	0	0	1	0	0	17	26
ACBD	0	0	0	1	0	1/4	0	3/4	0	0	0	0	0	0	2
ACDB	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1
ABDC	0	0	0	1	0	1/4	0	3/4	0	0	0	0	0	0	0
ADBC	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1
ADCB	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
BACD	0	1	0	0	0	1/2	0	1/2	1	1	0	1	0	115	88
BCAD	0	1	0	0	0	1/2	0	1/2	1	1	0	0	0	18	17
BCDA	0	1/2	0	1/2	0	1/2	0	1/2	0	1	0	0	0	1	2
BADC	0	1/2	0	1/2	0	1/2	0	1/2	0	1	0	0	0	8	8
BDAC	0	0	0	1	0	1/2	0	1/2	0	1	0	0	0	7	9
BDCA	0	0	0	1	0	1/2	0	1/2	0	1	0	0	0	1	5
CABD	0	0	0	1	0	1/4	1/4	1/2	0	0	0	0	0	0	2
CBAD	0	1/2	0	1/2	0	1/4	1/4	1/2	0	0	0	0	0	1	5
CBDA	0	0	0	1	0	1/4	1/4	1/2	0	0	0	0	0	0	0
CADB	0	0	0	1	0	0	1/4	3/4	0	0	0	0	0	0	0
CDAB	0	0	0	1	0	0	1/4	3/4	0	0	0	0	0	1	0
CDBA	0	0	0	1	0	0	1/4	3/4	0	0	0	0	0	0	0
DABC	0	0	0	1	0	0	0	1	0	0	0	0	0	5	8
DBAC	0	0	0	1	0	0	0	1	0	0	0	0	1	19	16
DBCA	0	0	0	1	0	0	0	1	0	0	0	0	0	12	11
DACB	0	0	0	1	0	0	0	1	0	0	0	0	0	1	5
DCAB	0	0	0	1	0	0	0	1	0	0	0	0	0	2	1
DCBA	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1

Details of the vote profile characteristics

	Winning probabilities for candidates								Voting strategies/heuristics					observed frequency	
	Borda				*IRV*				Optimal		Sincere	Lifting	Conforming		
Ballot	*A*	*B*	*C*	*D*	*A*	*B*	*C*	*D*	*Borda*	*IRV*				*Borda*	*IRV*
	vote profile: Lift-Overstate-4 (IRV: BACD, BC - ; Borda: B -C)														
ABCD	0	0	1	0	0	0	1	0	0	0	1	0	0	14	29
ACBD	0	0	1	0	0	0	1	0	0	0	0	0	0	0	2
ACDB	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1
ABDC	0	1/2	1/2	0	0	0	3/4	1/4	0	0	0	0	0	4	4
ADBC	0	0	1	0	0	0	3/4	1/4	0	0	0	0	0	0	1
ADCB	0	0	1	0	0	0	3/4	1/4	0	0	0	0	0	0	1
BACD	0	1/2	1/2	0	0	1/4	3/4	0	0	1	0	1	0	21	39
BCAD	0	0	1	0	0	1/4	3/4	0	0	1	0	0	0	13	22
BCDA	0	0	1	0	0	1/4	3/4	0	0	1	0	0	0	2	1
BADC	0	1	0	0	0	1/4	1/2	1/4	1	0	0	0	0	74	2
BDAC	0	1	0	0	0	1/4	1/2	1/4	1	0	0	0	0	16	1
BDCA	0	1/2	1/2	0	0	1/4	1/2	1/4	0	0	0	0	0	0	0
CABD	0	0	1	0	0	0	1	0	0	0	0	0	0	3	17
CBAD	0	0	1	0	0	0	1	0	0	0	0	0	1	56	56
CBDA	0	0	1	0	0	0	1	0	0	0	0	0	0	5	10
CADB	0	0	1	0	0	0	1	0	0	0	0	0	0	0	1
CDAB	0	0	1	0	0	0	1	0	0	0	0	0	0	0	3
CDBA	0	0	1	0	0	0	1	0	0	0	0	0	0	0	8
DABC	0	0	1	0	0	0	1/2	1/2	0	0	0	0	0	0	3
DBAC	0	0	1	0	0	0	1/2	1/2	0	0	0	0	0	0	6
DBCA	0	1/2	1/2	0	0	0	1/2	1/2	0	0	0	0	0	0	1
DACB	0	0	1	0	0	0	1/2	1/2	0	0	0	0	0	0	0
DCAB	0	0	1	0	0	0	1/2	1/2	0	0	0	0	0	0	0
DCBA	0	0	1	0	0	0	1/2	1/2	0	0	0	0	0	0	0

Details of the vote profile characteristics

	Winning probabilities for candidates								Voting strategies/heuristics					observed frequency	
	Borda				*IRV*				Optimal		Sincere	Lifting	Conforming		
Ballot	*A*	*B*	*C*	*D*	*A*	*B*	*C*	*D*	*Borda*	*IRV*				*Borda*	*IRV*
	vote profile: Sincere-Bury-3 (IRV: ABC- ; Borda: ACB)														
ABC	1/2	1/2	0	-	1/4	3/4	0	-	0	1	1	0	0	38	61
ACB	1	0	0	-	1/4	1/2	1/4	-	1	0	0	0	0	103	8
BAC	0	1	0	-	0	1	0	-	0	0	0	1	1	61	108
BCA	0	1	0	-	0	1	0	-	0	0	0	0	0	5	22
CAB	1/3	1/3	1/3	-	0	1/2	1/2	-	0	0	0	0	0	0	8
CBA	0	1	0	-	0	1/2	1/2	-	0	0	0	0	0	1	1
	vote profile: Sincere-Bury-4 (IRV: AB-; Borda: ACBD)														
ABCD	1/2	1/2	0	0	5/16	7/16	0	4/16	0	1	1	0	0	31	51
ACBD	1	0	0	0	5/16	4/16	3/16	4/16	1	0	0	0	0	81	3
ACDB	1/2	0	0	1/2	5/16	3/16	3/16	5/16	0	0	0	0	0	1	3
ABDC	1/3	1/3	0	1/3	5/16	7/16	0	4/16	0	1	0	0	0	4	5
ADBC	0	0	0	1	5/16	3/16	0	8/16	0	0	0	0	0	2	0
ADCB	0	0	0	1	5/16	2/16	1/16	8/16	0	0	0	0	0	0	0
BACD	0	1	0	0	0	3/4	0	1/4	0	0	0	1	0	28	69
BCAD	0	1	0	0	0	3/4	0	1/4	0	0	0	0	0	2	18
BCDA	0	1	0	0	0	3/4	0	1/4	0	0	0	0	0	3	6
BADC	0	1	0	0	0	3/4	0	1/4	0	0	0	0	0	14	10
BDAC	0	1/2	0	1/2	0	3/4	0	1/4	0	0	0	0	0	11	14
BDCA	0	1/2	0	1/2	0	3/4	0	1/4	0	0	0	0	0	2	4
CABD	1/4	1/4	1/4	1/4	0	1/4	1/2	1/4	0	0	0	0	0	0	1
CBAD	0	1	0	0	0	1/4	1/2	1/4	0	0	0	0	0	0	0
CBDA	0	1/2	0	1/2	0	1/4	1/2	1/4	0	0	0	0	0	1	0
CADB	0	0	0	1	0	1/4	1/2	1/4	0	0	0	0	0	0	1
CDAB	0	0	0	1	0	1/4	1/2	1/4	0	0	0	0	0	0	2
CDBA	0	0	0	1	0	1/4	1/2	1/4	0	0	0	0	0	0	0
DABC	0	0	0	1	0	0	0	1	0	0	0	0	0	3	3
DBAC	0	0	0	1	0	0	0	1	0	0	0	0	1	24	11
DBCA	0	0	0	1	0	0	0	1	0	0	0	0	0	0	6
DACB	0	0	0	1	0	0	0	1	0	0	0	0	0	1	0
DCAB	0	0	0	1	0	0	0	1	0	0	0	0	0	0	1
DCBA	0	0	0	1	0	0	0	1	0	0	0	0	0	0	0
	vote profile: Sincere-3 (ABC)														
ABC	1	0	0	-	1/4	1/4	1/2	-	1	1	1	0	0	119	50
ACB	1/2	0	1/2	-	1/4	0	3/4	-	0	0	0	0	0	22	19
BAC	1/3	1/3	1/3	-	0	1/2	1/2	-	0	0	0	1	0	18	69
BCA	0	0	1	-	0	1/2	1/2	-	0	0	0	0	0	3	15
CAB	0	0	1	-	0	0	1	-	0	0	0	0	1	43	44
CBA	0	0	1	-	0	0	1	-	0	0	0	0	0	3	11

Table A2: Details of the vote profile characteristics

B. Decision time

	Decision time			*Proportion of optimal ballots*		
Vote profile	*IRV*	*Borda*	*Mean*	*IRV*	*Borda*	Mean
Lift-3	67	49	58	1	1	1
Sincere-Bury-3	40	48	44	0.80	1	0.90
Sincere-3	107	52	80	1	1	1
Mean - 3	71	50	61	0.93	1	0.97
Lift-4	229	96	162	1	1	1
Lift-Overstate-4	116	52	84	1	0.80	0.90
Sincere-Bury-4	79	84	82	0.80	1	0.90
Mean - 4	141	77	109	0.93	0.93	0.93
Mean - all	106	63	85	0.93	0.97	0.95

Table A3: Average decision time in seconds by voting procedure, vote profile, and number of candidates for arriving at the optimal ballot for voters with 11 or more optimal first-attempt ballots. ($N = 5$)

Vote profile	*Decision time* *IRV*	*Borda*	Mean	*Proportion of optimal ballots* *IRV*	*Borda*	Mean
Lift-3	99	70	84	1	1	1
Sincere-Bury-3	31	56	44	0.67	1	0.83
Sincere-3	95	57	76	1	1	1
Mean - 3	75	61	68	0.89	1	0.94
Lift-4	164	105	135	0.78	1	0.89
Lift-Overstate-4	129	94	111	0.78	0.89	0.83
Sincere-Bury-4	113	117	115	0.67	1	0.83
Mean - 4	135	105	120	0.74	0.96	0.85
Mean - all	105	83	94	0.82	0.98	0.90

Table A4: Average decision time in seconds by voting procedure, vote profile, and number of candidates for arriving at the optimal ballot for voters with 10 or more optimal first-attempt ballots. ($N = 9$)

Vote profile	*Decision time* *IRV*	*Borda*	Total	*Proportion of optimal ballots* *IRV*	*Borda*	Total
Lift-3	80	76	78	0.94	1	0.97
Sincere-Bury-3	31	59	45	0.53	1	0.76
Sincere-3	59	61	60	0.59	1	0.76
Mean - 3	57	65	61	0.69	1	0.84
Lift-4	118	104	111	0.82	1	0.91
Lift-Overstate-4	86	73	80	0.71	0.82	0.76
Sincere-Bury-4	81	108	94	0.59	0.94	0.76
Mean - 4	95	95	95	0.71	0.92	0.81
Mean - all	75	80	78	0.70	0.96	0.82

Table A5: Average decision time in seconds by voting procedure, vote profile, and number of candidates for arriving at the optimal ballot for voters with 9 or more optimal first-attempt ballots. ($N = 17$)

C. Comparison with Random Voting

Procedure		Random	1^{st} attempt		2^{nd} attempt	
			Observed	*p-value*	**Observed**	*p-value*
IRV	**Optimal**	18.8%	35.4%	*0.0008*	38.8%	*< 0.001*
	Sincere	10.4%	20.0%	*0.0203*	21.2%	*0.071*
	Lifting	10.4%	38.5%	*< 0.001*	34.9%	*< 0.001*
	Conforming	10.4%	23.4%	*< 0.001*	20.2%	*< 0.001*
Borda	**Optimal**	11.8%	49.2%	*< 0.001*	53.5%	*< 0.001*
	Sincere	10.4%	18.6%	*< 0.001*	18.3%	*< 0.001*
	Lifting	10.4%	28.0%	*< 0.001*	29.3%	*< 0.001*
	Conforming	10.4%	22.4%	*0.2999*	18.8%	*0.0193*

Table A6: Comparison for each voting procedure of the frequencies of submitted votes consistent with each of the four categories of voting behavior (Optimal, Sincere, Lifting, and Conforming) in the first and the second attempts with the frequencies we would have observed had voters submitted random votes. Statistical tests: Wilcoxon-Mann-Whitney tests (two-sided).

D. Heuristics

Voting procedure	Optimal in	Optimal ballot(s)	Frequencies L3	SB3	S3	*p-value* *L3 vs SB3*	*L3 vs S3*	*SB3 vs S3*
IRV	**L3**	B–	49.0%	68.2%	43.3%	*0.0098*	*0.3304*	
	SB3	ABC	18.3%	25.0%	22.1%	*0.223*		*0.6015*
	S3	ABC	18.3%	25.0%	22.1%		*0.4652*	*0.6015*
Borda	**L3**	BAC	51.9%	30.8%	9.6%	*0.0116*	*<0.001*	
	SB3	ACB	1.9%	48.1%	9.6%	*<0.001*		*<0.001*
	S3	ABC	5.8%	18.3%	54.8%		*<0.001*	*<0.001*
			L4	**LO4**	**SB4**	*L4 vs LO4*	*L4 vs SB4*	*LO4 vs SB4*
IRV	**L4**	B–	62.6%	28.9%	57.7%	*<0.001*	*0.4233*	
	LO4	BACD, BC–	54.9%	26.9%	45.2%	*<0.001*		*0.0038*
	SB4	AB–	9.6%	21.1%	27.0%		*0.0002*	*0.2733*
Borda	**L4**	B–D	62.5%	15.4%	14.4%	*<0.001*	*<0.001*	
	LO4	B–C	6.7%	41.3%	9.6%	*<0.001*		*<0.001*
	SB4	ACBD	0.0%	0.0%	36.5%		*<0.001*	*<0.001*

Table A7: Frequency with which an optimal ballot in a vote profile is cast in the other two vote profiles with the same number of candidates

Note: L stands for *Lift*, SB for *Sincere-Bury*, S for *Sincere*, and LO for *Lift-Overstate*. Underlined frequencies indicate when the ballot(s) is (are) optimal.

E. Winning probabilities

		Candidate			
		A	**B**	**C**	**D**
IRV	**1st attempt**	4.0%	40.8%	38.6%	33.3%
	2nd attempt	4.7%	40.7%	37.8%	33.7%
	p-value	*0.115*	*0.8971*	*0.2012*	*0.8675*
Borda	**1st attempt**	27.6%	36.8%	26.8%	17.5%
	2nd attempt	29.5%	38.4%	24.8%	14.7%
	p-value	*0.0024*	*0.0658*	*0.0031*	*0.0794*

Table A8: Winning probabilities of the different candidates

Note: Winning probabilities for *D* are computed for the three 4-candidate vote profiles. For example, in the first attempts, *D* wins 33.3% of the time in the three 4-candidate vote profiles, while *C* wins 38.6% of the time in the six vote profiles.

F. Realized payoffs

Vote profile	Voting procedure	1st attempt	2nd attempt	*p-value*
All profiles	**IRV**	22.7$	23.3$	*0.0906*
	Borda	28.4$	29.1$	*0.006*
	p-value	*<0.001*	*<0.001*	
Lift-3	**IRV**	22.3$	23.3$	*0.1102*
	Borda	22.3$	23.0$	*0.324*
	p-value	1	0.7428	
Lift-4	**IRV**	16.3$	17.0$	*0.6271*
	Borda	23.8$	24.4$	*0.6072*
	p-value	*<0.001*	*<0.001*	
Lift-Overstate-4	**IRV**	20.2$	20.2$	*1*
	Borda	24.9$	25.0$	*1*
	p-value	*<0.001*	*<0.001*	
Sincere-Bury-3	**IRV**	30.7$	30.5$	*0.8036*
	Borda	35.8$	36.3$	*0.2379*
	p-value	*<0.001*	*<0.001*	
Sincere-Bury-4	**IRV**	23.4$	24.5$	*0.3489*
	Borda	30.6$	32.1$	*0.0449*
	p-value	*<0.001*	*<0.001*	
Sincere-3	**IRV**	23.6$	24.5$	*0.1841*
	Borda	32.9$	33.7$	*0.3018*
	p-value	*<0.001*	*<0.001*	

Table A9: Average realized payoffs

G. Construction data base for multinomial analysis

For both voting procedures, we have chosen 6 profiles, for which we consider in our analysis a total of 5 heuristics. Participants make 2 choices for each profile. Thus, each strategy could have been used at maximum 24 times (2 rules x 6 profiles x 2 attempts) and a maximum strategy space of 120 : 24 x 5 heuristics.

However, heuristics coincide for certain profiles sometimes differently under both rules. Only the *Optimality* and *Unclassified* heuristics can be used at maximum 24 times. *Sincere* coincides with *Optimality* under the Sincere 3 profile and for IRV only under Sincere-Bury 3 and 4. *Lifting* coincides with *Optimality* under the Lift 3 and 4 profiles, for IRV only under Lift-Overstate-4, and coincides with the *Conforming* heuristic under Sincere-Bury-3.

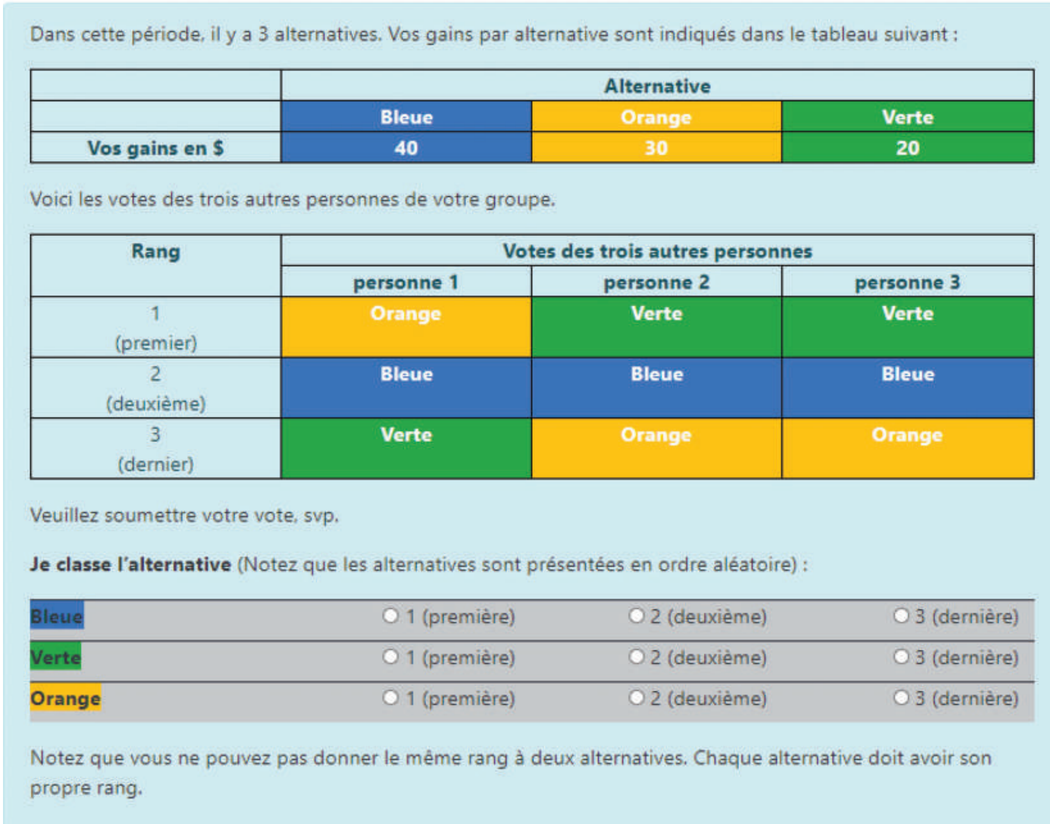
Table A10 presents the possibility to observe strategies by profile and the total number of observations in our sample.

	Heuristics								
	Opt	Sinc		Lift		Conf	Unclass	max possible	
Profile / Rule	IRV/Borda	IRV	Borda	IRV	Borda	IRV/Borda	IRV/Borda	IRV	Borda
Lift-3	X	X		-		X	X	4	
Lift 4	X	X		-		X	X	4	
Lift-Overstate-4	X	X		-	X	X	X	4	5
Sincere-Bury-3	X	-	X	(x)		(x)	X	3	4
Sincere-Bury-4	X	-	X	X		X	X	4	5
Sincere-3	X	-		X		X	X	4	
all Lift count	6	3	5	3	4	5	6		
all Conf count	6	3	5	2	3	6	6		
both rules	12	8		5/7		10/12	12	**49**	
2 attempts	24	16		10/14		20/24	24	**98**	
104 participants									
all Lift count	2.496	1.664		1.456		2.080	2.496	**10.192**	
all Conf count	2.496	1.664		1.040		2.496	2.496	**10.192**	

Table A10: Constructing data set with unbalanced vote categories.

Supplementary Online Appendix

Screenshots



Dans cette période, il y a 3 alternatives. Vos gains par alternative sont indiqués dans le tableau suivant :

	Alternative		
	Bleue	Orange	Verte
Vos gains en $	40	30	20

Voici les votes des trois autres personnes de votre groupe.

Rang	Votes des trois autres personnes		
	personne 1	personne 2	personne 3
1 (premier)	Orange	Verte	Verte
2 (deuxième)	Bleue	Bleue	Bleue
3 (dernier)	Verte	Orange	Orange

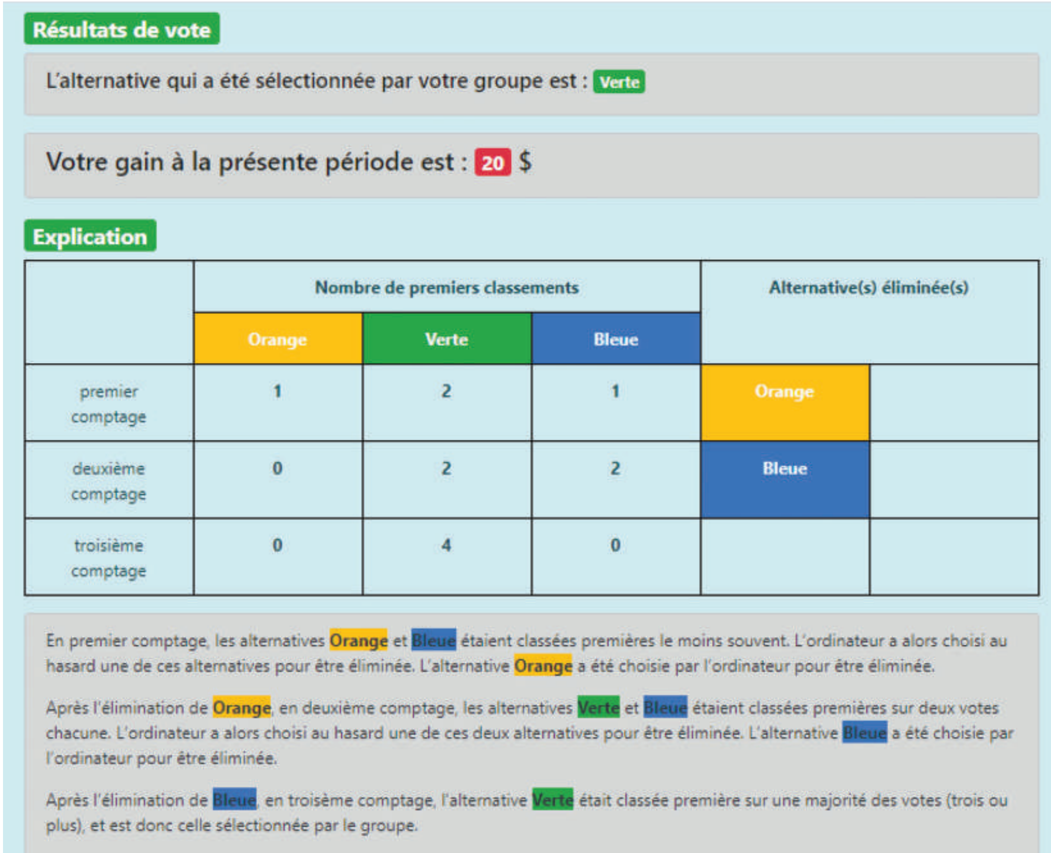
Veuillez soumettre votre vote, svp.

Je classe l'alternative (Notez que les alternatives sont présentées en ordre aléatoire) :

Bleue	1 (première)	2 (deuxième)	3 (dernière)
Verte	1 (première)	2 (deuxième)	3 (dernière)
Orange	1 (première)	2 (deuxième)	3 (dernière)

Notez que vous ne pouvez pas donner le même rang à deux alternatives. Chaque alternative doit avoir son propre rang.

Fig. SupA1: Ballot screen


Résultats de vote

L'alternative qui a été sélectionnée par votre groupe est : Verte

Votre gain à la présente période est : 20 $

Explication

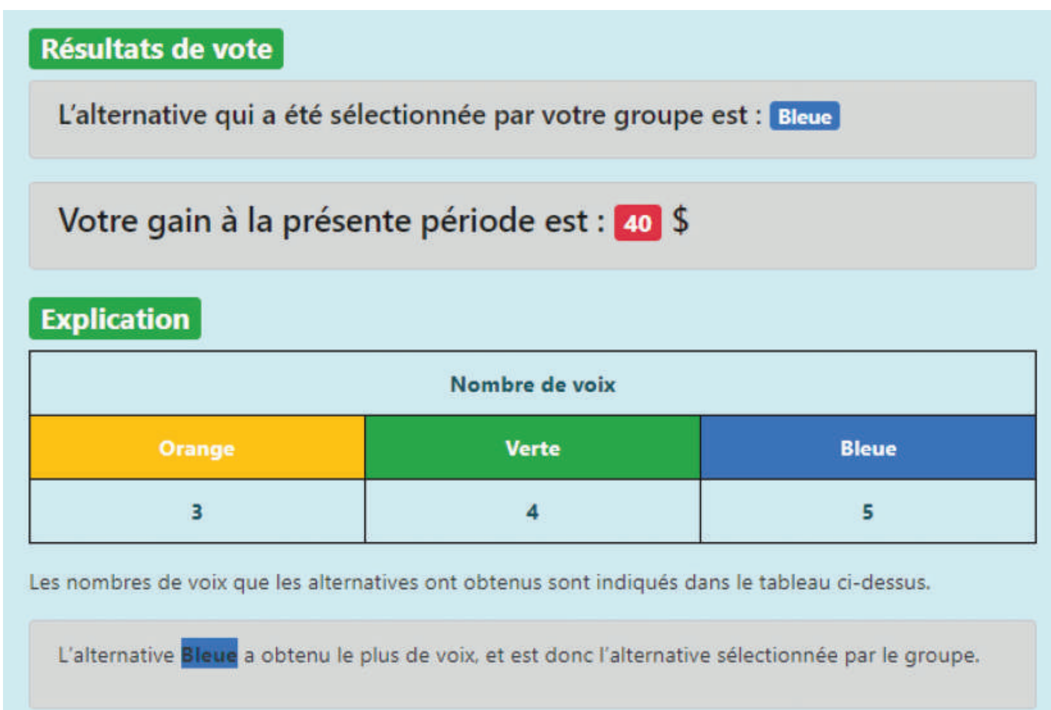
	Nombre de premiers classements			Alternative(s) éliminée(s)	
	Orange	Verte	Bleue		
premier comptage	1	2	1	Orange	
deuxième comptage	0	2	2	Bleue	
troisième comptage	0	4	0		

En premier comptage, les alternatives Orange et Bleue étaient classées premières le moins souvent. L'ordinateur a alors choisi au hasard une de ces alternatives pour être éliminée. L'alternative Orange a été choisie par l'ordinateur pour être éliminée.

Après l'élimination de Orange, en deuxième comptage, les alternatives Verte et Bleue étaient classées premières sur deux votes chacune. L'ordinateur a alors choisi au hasard une de ces deux alternatives pour être éliminée. L'alternative Bleue a été choisie par l'ordinateur pour être éliminée.

Après l'élimination de Bleue, en troisème comptage, l'alternative Verte était classée première sur une majorité des votes (trois ou plus), et est donc celle sélectionnée par le groupe.

Fig. SupA2: Result screen under RCV


Résultats de vote

L'alternative qui a été sélectionnée par votre groupe est : Bleue

Votre gain à la présente période est : 40 $

Explication

Nombre de voix		
Orange	Verte	Bleue
3	4	5

Les nombres de voix que les alternatives ont obtenus sont indiqués dans le tableau ci-dessus.

L'alternative Bleue a obtenu le plus de voix, et est donc l'alternative sélectionnée par le groupe.

Fig. SupA3: Result screen under Borda

Instructions and Questionnaire

English translation from the French Instructions and Questionnaire for the sequence RCV first – Borda second.

SCREEN 0

Welcome to the experiment!

The experiment contains two parts. In the first part, we ask you to make decisions, following the instructions we are about to give you. In the second part, we ask you to answer a series of questions.

If you have questions, please raise your hand and we will come answering them in private. We ask you to not communicate with the other participants.

This session will last approximately two hours.

TRANSITION SCREEN

Instructions

SCREEN 1 -- INSTRUCTIONS

The session will consist of a series of independent periods. During each period, you will be part of a group of four persons who must choose by vote one among several alternatives.

Your gains will depend on the alternative that will be selected by the group. Specifically, each alternative will be referred to as a color and associated with a number corresponding to your gain if this alternative is selected by the group.

For example, consider the case with three alternatives described in the table below. You would receive 20$ if the group were to choose « Blue », 40$ if the chosen alternative is « Orange » and 5$ if the chosen alternative is « Green ».

	Alternative		
	Blue	Orange	Green
Your gain in $	20	40	5

The other three persons are not present during the session, but the computer will indicate their votes before you choose your own vote. All four votes (yours and the votes of the other three persons presented by the computer) will be used to select the alternative for your group.

Next, the computer will report on the screen the alternative that has been selected by your group, as well as your gain during this period. Once you will have read the results, the next period will start.

We explain the voting procedure on the next screen. The voting procedure will change during the session. You will be informed when this happens.

There will be a total of 24 voting periods. A voting situation is repeated twice, during two consecutive periods, for example, periods 1 and 2, periods 3 and 4, and so on. Your group will **choose between three or four alternatives**. At the end of the session, the computer will report on the screen your gain in each period and one period will be randomly chosen. Your compensation for your participation in the experiment will be equal to your gain during the randomly selected period.

SCREEN 2 -- INSTRUCTIONS

Voting procedure

We now describe a voting period.

After being informed about your gain from each alternative, the computer will report the ballots of the other three persons in your group and will ask you to submit your own ballot. A ballot consists of **a ranking of all three or four alternatives, from first to last.**

Here is an example with three alternatives.

	Your ballot	**Ballots of the other three persons**		
Rank	**You**	**person 1**	**person 2**	**person 3**
1 (first)		Orange	Orange	Green
2 (second)		Green	Blue	Orange
3 (last)		Blue	Green	Blue

In this example, person 1 has ranked the alternative Orange first, Green second, and Blue last. Person 2 ranked the alternative Orange before Blue and ranked the alternative Green last. Finally, person 3 has submitted the following ranking: Green, Orange, and, finally, Blue.

You will then be invited to submit your own ballot, that is, rank the alternatives from first to last.

Once you will have submitted your ballot, the computer will determine the selected alternative as follows.

SCREEN 3 -- INSTRUCTIONS

- If one alternative is ranked first on a majority of the four ballots, that is, on **three or four ballots**, then this alternative will be the one selected by your group during this period. The process stops there.

  For instance, if you rank Orange first, Blue second, and Green last,

	Your ballot	**Ballots of the other three persons**		
Rank	**You**	**person 1**	**person 2**	**person 3**
1 (first)	Orange	Orange	Orange	Green
2 (second)	Blue	Green	Blue	Orange
3 (last)	Green	Blue	Green	Blue

then Orange is ranked first on a majority of ballots (3 out of 4 votes). Orange is then the selected alternative, and the process stops there.

SCREEN 4 -- INSTRUCTIONS

- If neither alternative is ranked first on a majority of the four ballots, that is, neither alternative is ranked first at least three times, then:

1) The computer will eliminate the alternative which is ranked first the least often. If several alternatives are in that case, that is, if there is a tie, then the computer will eliminate randomly one of these alternatives, that is, one of the alternatives ranked first the least often.

2) The eliminated alternative is then removed from the ballots.

3) Steps 1 and 2 will be repeated with the remaining alternatives until one alternative will be ranked first on a majority of ballots. This alternative will then be the selected alternative during this period. The process stops there.

For example, if you rank Blue first, Orange second, and Green last,

	Your ballot	**Ballots of the other three persons**		
Rank	**You**	**person 1**	**person 2**	**person 3**
1 (first)	Blue	Orange	Orange	Green
2 (second)	Orange	Green	Blue	Orange
3 (last)	Green	Blue	Green	Blue

then alternative Orange is ranked first twice, while alternatives Blue and Green are each ranked first once. Hence, neither alternative is ranked first on a majority of ballots. The alternative ranked first the least often is then eliminated. In this example, there are two alternatives that are ranked first the least often: Blue and Green.

The computer eliminates randomly alternative Blue or alternative Green, which are equally likely to be eliminated.

1) If Blue is the one eliminated,

	Your ballot	Ballots of the other three persons		
Rank	**You**	**person 1**	**person 2**	**person 3**
1 (first)	~~Blue~~ *(eliminated)*	Orange	Orange	Green
2 (second)	Orange	Green	~~Blue~~ *(eliminated)*	Orange
3 (last)	Green	~~Blue~~ *(eliminated)*	Green	~~Blue~~ *(eliminated)*

then Orange and Green are the two remaining alternatives.

	Your ballot	Ballots of the other three persons		
Rank	**You**	**person 1**	**person 2**	**person 3**
1 (first)	Orange	Orange	Orange	Green
2 (second)	Green	Green	Green	Orange

Orange is now ranked first three times out of four; Orange is then ranked first on a majority of ballots. Orange is the selected alternative, and the process stops there.

2) If Green is the one eliminated,

	Your ballot	Ballots of the other three persons		
Rank	**You**	**person 1**	**person 2**	**person 3**
1 (first)	Blue	Orange	Orange	~~Green~~ *(eliminated)*
2 (second)	Orange	~~Green~~ *(eliminated)*	Blue	Orange
3 (last)	~~Green~~ *(eliminated)*	Blue	~~Green~~ *(eliminated)*	Blue

then Orange and Blue are the two remaining alternatives.

	Your ballot	Ballots of the other three persons		
Rank	**You**	**person 1**	**person 2**	**person 3**
1 (first)	Blue	Orange	Orange	Orange
2 (second)	Orange	Blue	Blue	Blue

Orange is now ranked first three times out of four; Orange is then ranked first on a majority of ballots. Orange is the selected alternative, and the process stops there.

End of instructions

SCREEN 5 -- QUESTIONNAIRE

This ends the description of the first voting procedure. Before starting with the first voting period, please answer the following comprehension questions. You can read again the instructions at any time by clicking on the 'instructions' button.

Instructions and Questionnaire

Question 1. Suppose there are three alternatives: Blue, Orange, and Green. Suppose that the group consists of five persons who rank the alternatives as indicated in the table:

	Ballots				
Rank	**Person 1**	**Person 2**	**Person 3**	**Person 4**	**Person 5**
1	Orange	Blue	Green	Orange	Blue
2	Green	Orange	Orange	Blue	Green
3	Blue	Green	Blue	Green	Orange

1. How many times is « Blue » ranked first?
   How many times is « Orange » ranked first?
   How many times is « Green » ranked first?

Answers: 2, 2, and 1.

2. Is one alternative ranked first on a majority of ballots?

Answer: No.

3. Which alternative is eliminated first?

Answer: Green.

4. After the elimination of Green, how many times is Blue ranked first?
   After the elimination of Green, how many times is Orange ranked first?

Answer: 2 and 3.

If the answer is wrong:

Here is a table reporting the votes after the elimination of Green.

	Ballots				
Rank	**Person 1**	**Person 2**	**Person 3**	**Person 4**	**Person 5**
1	Orange	Blue	Orange	Orange	Blue
2	Blue	Orange	Blue	Blue	Orange

Please reconsider your answer.

If the answer is still wrong:

Orange is now ranked first by persons 1, 3, and 4, and Blue by persons 2 and 5.

5. After the elimination of Green, is one alternative ranked first on a majority of ballots?

Answer: Yes.

6. Which one of the three alternatives is selected?

Answer: Orange.

If the answer is wrong, depending on the reported alternative:

[Green:] The selected alternative is the one ranked first on a majority of ballots. Green has been eliminated.

[Blue:] The selected alternative is the one ranked first on a majority of ballots. After Green has been eliminated, Orange is now ranked first three times, by persons 1, 3, and 4, and Blue is ranked first twice by persons 2 and 5.

Please reconsider your answer. If you still have questions, please do not hesitate to call the lab assistant.

SCREEN 6 -- QUESTIONNAIRE

Question 2. Suppose there are three alternatives: Blue, Orange, and Green. Suppose that your gains are as indicated in the following table:

	Alternative		
	Blue	Orange	Green
Your gain in $	10	8	3

1. If alternative Blue is selected, what is your gain?

Answer: 10.

2. If alternative Orange is selected, what is your gain?

Answer: 8.

3. If alternative Green is selected, what is your gain?

Answer: 3.

TRANSITION SCREEN

This concludes the instructions.

Beginning of the experiment

SCREEN 7 -- DECISION

We now begin the first voting period. This vote will be repeated in the next period. You can look again at the instructions at any time by clicking on the button « Instructions ».

SCREEN 8 -- DECISION

Please confirm your choice.

SCREEN 9 -- RESULT

Results of the vote:

The alternative selected by your group is …

Your gain in this period is …

Explanation:

	Number of first rankings			Eliminated alternative
	Orange	Blue	Green	
First count				
Second count				
Third count				

Alternative X was ranked first on a majority of ballots (three or more) and thus is the one selected by the group.

In the first count, alternative X was ranked first less often than the other alternatives and thus is eliminated.

In the first count, alternatives X and Y were ranked first the least often. The computer has then chosen randomly one of these alternatives to be eliminated. Alternative X has been chosen by the computer to be eliminated.

After the elimination of X, in the second count, alternative Y was ranked first on a majority of ballots and is thus the alternative selected by the group.

After the elimination of X, in the second count, alternatives Y and Z were ranked first on two ballots each. The computer has then chosen randomly one of these two alternatives to be eliminated. Alternative Y has been chosen by the computer to be eliminated.

After the elimination of X, in the second count, alternative Y was ranked first the least often and is thus eliminated.

After the elimination of X, in the second count, alternatives Y and Z were ranked first the least often. The computer has then chosen randomly one of these alternatives to be eliminated. Alternative Y has been chosen by the computer to be eliminated.

After the elimination of Y, in the third count, alternative Z was ranked first on a majority of ballots (three or more) and is thus the alternative selected by the group.

After the elimination of Y, in the third count, the remaining two alternatives were ranked first on two ballots each. The computer has chosen randomly one of these two alternatives as the alternative selected by the group. Alternative Z has been chosen by the computer as the selected alternative.

You will find below a reminder of your ballot and those of the other persons, as well as your gain from each alternative.

TRANSITION SCREEN

We are now going to change the voting procedure. During the next periods, you will choose alternatives using another voting procedure. We explain this voting procedure on the next screen.

SCREEN 10 -- INSTRUCTIONS
Voting procedure

We now describe a voting period.

After being informed about your gain from each alternative, the computer will present the ballots of the other three persons in your group and will ask you to submit your own ballot. A ballot consists of **a ranking of all alternatives, from first to last.**

Each position on a ranking corresponds to the number of votes that the alternative will receive. More precisely, if there are three alternatives, the alternative ranked first receives 2 votes, the alternative ranked second receives 1 vote and the alternative ranked last receives 0 vote. If there are four alternatives, the alternative ranked first receives 3 votes, the alternative ranked second receives 2 votes, the alternative ranked third receives 1 vote and the alternative ranked last receives 0 vote.

Here is an example with three alternatives.

	Your ballot	**Ballots of the other three persons**		
Rank	**You**	**person 1**	**person 2**	**person 3**
1 (first = 2 votes)		Orange	Orange	Blue
2 (second = 1 vote)		Green	Blue	Green
3 (last = 0 vote)		Blue	Green	Orange

In this example, person 1 has ranked alternative Orange first, Green second, and Blue last. Person 2 has ranked alternative Orange before Blue and ranked alternative Green last. Finally, person 3 has posted the following ranking: Blue, Green, and, finally, Orange. Orange has then 4 votes (2 from person 1, 2 from person 2, and 0 from person 3). Blue has 3 votes (0 from person 1, 1 from person 2, and 2 from person 3). Finally, Green has 2 votes (1 from person 1, 0 from person 2, and 1 from person 3).

You will then be invited to submit your own ballot, that is, rank the three alternatives.

Once you will have submitted your ballot, the computer will determine the selected alternative as follows.

SCREEN 11 -- INSTRUCTIONS

- If one alternative has received the most votes from the four ballots, then this alternative will be the one selected by your group during this period. The process stops there.

  For example, if you rank Orange first, Blue second, and Green last,

	Your ballot (Number of votes)	Ballots of the other three persons (Number of votes)			Result (Total number of votes)
Rank	**You**	**person 1**	**person 2**	**person 3**	
1 (first = 2 votes)	Orange (2)	Orange (2)	Orange (2)	Blue (2)	Orange (2+2+2+0=6)
2 (second = 1 vote)	Blue (1)	Green (1)	Blue (1)	Green (1)	Blue (1+0+1+2=4)
3 (last = 0 vote)	Green (0)	Blue (0)	Green (0)	Orange (0)	Green (0+1+0+1=2)

then Orange receives 6 votes, Blue receives 4 and, finally, Green receives 2. Orange is thus the selected alternative, and the process stops there.

SCREEN 12 -- INSTRUCTIONS

- If several alternatives have received the most votes, that is, if there is a tie, then the computer will choose randomly one alternative among those that have received the most votes. This alternative will be the selected alternative during this period.

For example, if you rank Blue first, Orange second, and Green last,

	Your ballot (Number of votes)	Ballots of the other three persons (Number of votes)				Result (Total number of votes)
Rank	You	person 1	person 2	person 3		
1 (first = 2 votes)	Blue (2)	Orange (2)	Orange (2)	Blue (2)		Orange (1+2+2+0=5)
2 (second = 1 vote)	Orange (1)	Green (1)	Blue (1)	Green (1)		Blue (2+0+1+2=5)
3 (last = 0 vote)	Green (0)	Blue (0)	Green (0)	Orange (0)		Green (0+1+0+1=2)

then alternatives Orange and Blue obtain the same total number of votes (5 votes each) while Green obtains fewer votes (2 votes).

Thus the computer chooses randomly between alternatives Orange and Blue.

1) If Orange is chosen by the computer, then Orange is the selected alternative, and the process stops there.
2) If Blue is chosen by the computer, then Blue is the selected alternative, and the process stops there.

End of instructions.

SCREEN 13 -- QUESTIONNAIRE

This ends the description of the voting procedure. Before starting with the first voting period, please answer the following comprehension questions. You can read again the instructions at any time by clicking on the 'instructions' button.

Question. Suppose there are three alternatives: Blue, Orange, and Green. Suppose that the group consists of five persons who rank the alternatives as indicated in the table:

	Ballots				
Rank	**Person 1**	**Person 2**	**Person 3**	**Person 4**	**Person 5**
1 (first = 2 votes)	Orange	Blue	Green	Orange	Blue
2 (second = 1 vote)	Green	Orange	Orange	Blue	Green
3 (last = 0 vote)	Blue	Green	Blue	Green	Orange

1. How many votes does alternative Blue receive?
   How many votes does alternative Orange receive?
   How many votes does alternative Green receive?

Answers: 5, 6 and 4.

2. Does one alternative receive more votes than any of the other alternatives?
   - O No
   - O Yes, Blue
   - O Yes, Orange
   - O Yes, Green

Answer: Yes, Orange.

3. Which one of the three alternatives is selected?

Answer: Orange.

SCREEN 14 -- DECISION

We now begin the first voting period. This vote will be repeated in the next period. You can look again at the instructions at any time by clicking on the button « Instructions ».

SCREEN 15 -- DECISION

Please confirm your choice.

SCREEN 16 -- RESULT

Results of the vote:

The alternative selected by your group is …

Your gain in this period is …

Explanation:

Number of votes		
Orange	Blue	Green

The number of votes that each alternative has received is reported in the table above.

Alternative X has received the most votes and thus is the alternative selected by the group.

Alternatives X and Y / X, Y, and Z have each obtained the most votes. The computer has then chosen randomly one of these alternatives as the alternative selected by the group. Alternative X has been chosen by the computer.

You will find below a reminder of your ballot and those of the other persons, as well as your gain from each alternative.

SCREEN 17 QUESTIONNAIRE

You have successfully participated in all the votes of the experiment. We now ask you to answer a couple of questions. Thank you!

Your gender: O woman

O man

O I prefer not to answer

Your year of birth:

Your degree: O Bachelor

O Master

O Ph.D.

O other

Your (last) field of study:

Please, explain how you have made your decisions during the experiment:

SCREEN 18 – BERLIN NUMERACY TEST

Please, answer the following questions. Do not use a calculator.

You can take notes or make your computations on paper.

1) Among 1,000 persons in a small town, 500 are members of a choir.
Among these 500 persons, 100 are men.
Among the 500 persons who are not members of the choir, 300 are men.
What is the probability that a randomly selected man is a member of the choir (in percent)?

2) [a] Imagine that we roll 50 times a five-sided dice *(« 1 », « 2 », « 3 », « 4 », or « 5 »)*. Out of these 50 dice rolls, how many times do you expect the dice will indicate an uneven number (« 1 », « 3 » or « 5 »)?

2) [b] Imagine that we roll 70 times a six-sided dice *(« 1 », « 2 », « 3 », « 4 », « 5 » or « 6 »)*. For each number other than 6, the probability the dice indicates that number is half the probability the dice indicates a « 6 ». Out of these 70 dice rolls, how many times do you expect the roll will land on « 6 »?

3) In a forest, 20% of mushrooms are red, 50% brown and 30% white.
A red mushroom is poisonous with a 20% probability.
A mushroom that is not red is poisonous with a 5% probability.
What is the probability that a poisonous mushroom in the forest is red (in percent between 0 and 100)?

Correct Answers: Question 1: 25; 2a: 30; 2b: 20; 3: 50.
Q1 correct -> Q2b correct -> end score: 4.
Q1 correct -> Q2b wrong -> Q3 correct -> end score: 4.
Q1 correct -> Q2b wrong -> Q3 wrong -> end score: 3.
Q1 wrong -> Q2a correct -> end score: 2.

Q1 wrong -> Q2a wrong -> end score: 1.

SCREEN 19 – MEASURING LYING AVERSION

For each of the following statements, please indicate whether you think it can always be justified, can never be justified, or something in between by choosing the right degree.

	never justifiable 1	2	3	4	5	6	7	8	9	always Justifiable 10
Refusing to pay for a ride on public transportation	O	O	O	O	O	O	O	O	O	O
Cheat on tax payments if there is an opportunity to do so	O	O	O	O	O	O	O	O	O	O
Lying in general	O	O	O	O	O	O	O	O	O	O
Lying for one's own personal gain	O	O	O	O	O	O	O	O	O	O
Lying when there are few consequences for others	O	O	O	O	O	O	O	O	O	O
Lying when there are no consequences for others	O	O	O	O	O	O	O	O	O	O

SCREEN 20 – MEASURING SOCIAL PREFERENCES

Consider the following hypothetical situation. You receive 100$ that you can share with another person who participates in this experiment.

[Yellow hightlight: first screen]

Specifically, if you offer Y$ to the other person, then you will keep (100 – Y)$ for yourself, and the other participant will receive Y$.

How would you share the 100$ with this person?

I give the other person: …. $ and I keep for myself: …. $

[Blue hightlight: second screen]

If the experimenter were to double each dollar you give the other person, how would you share the 100$ with this person? Otherwise stated, if you offer Y$ to the other person, then you will keep (100 – Y)$ for yourself, and the other person will receive Y$ from you and another Y$ from the experimenter.

I give the other person: …. $ and I keep for myself: …. $

The other person receives: ________________$ [=2*Y]

and I receive: ______________$ [100 – Y]

[Green hightlight: third screen]

If the experimenter were to double each dollar you keep for yourself, how would you share the 100$ with the other person? Otherwise stated, if you offer Y$ to the other person, then you will keep (100 – Y)$ for yourself and will receive another (100 – Y)$ from the experimenter. The other person will receive Y$ from you.

I give the other person: …. $ and I keep for myself: …. $

The other person receives: ____________________ $ [Y]

and I receive: ___________________ $ [2* (100 – Y)]

SCREEN 21 – MEASURING RISK AVERSION

Consider the following hypothetical situation. You draw a card from a pack containing 100 cards numbered from 1 to 100. You win 100$ if the number on the card you draw lies above 75 (i.e., 76, 77, … 99, or 100). Otherwise, you receive nothing.

What is the smallest amount of money at which you are indifferent between receiving this amount of money for sure and playing the lottery that was just described?

[sliderbar] 0-------------------|-----------------------------------100$

SCREEN 22

Summary and drawing of the period determining your gain:

Here is a summary of your gain in each period:

Period	**The alternative selected by the group**	**Your gain**
1		
2		
…		
24		

As mentioned in the instructions at the beginning of the experiment, the computer is going to choose with your help one of the 24 periods to determine your gain in the first part of the experiment. Each period has the same chance of being chosen. Click on the button below to determine the period.

[Start the draw]

[Stop the wheel to determine the period]

The chosen period is …..

Your gain for the first part of the experiment is … $.

SCREEN 23 -- END

Thank you for your participation!

Here is a reminder of your total compensation:
Gain at period ….: …. $
Show-up fee: 15 $
Total: …$

Please raise your hand when you are done.
Please stay sit.
We are about to bring you your compensation.

Your participant number is: …