The Effects of District Magnitude on Voting Behaviour*

Simon Hix
London School of Economics and Political Science

Rafael Hortala-Vallve
London School of Economics and Political Science

Guillem Riambau
Yale - NUS College

June 17, 2013

Abstract. Is there less strategic voting in multi-member districts than in single-member districts? Existing research on this question is inconclusive, at least in part because it is difficult in observational data to isolate the effect of district magnitude on voting behavior independently from voters' preferences or parties' positions. Hence, we investigate this issue in a laboratory experiment, where we vary district magnitude while keeping voters' preferences and parties' positions constant. We find that voting for the preferred party (sincere voting) increases with district magnitude and we are able to explain this in terms of a mechanical effect and a psychological one. We also find a high incidence of voting for the frontrunner in all our elections, even when there are no incentives for doing so.

*For helpful comments we thank Manuel Arellano, André Blais, Raymond Duch, Iván Fernández-Val, Alex Fouirnaies, Simon Hug, Becky Morton, Matt Shugart and audience members at the LSE/NYU conference in Political Economy, APSA 2011, EPSA 2012, and MPSA 2013.
INCOMPLETE!

1. Introduction

The reform of electoral systems is a salient policy concern for new democracies as well as many advanced democracies. One key issue in the design of electoral systems is the ideal *district magnitude*: the number of candidates to be elected in each district. For example, in 2012 Romania was considering switching from proportional representation to single-member districts, Israel was considering switching from a single national multi-member district to smaller multi-member districts, and Tunisia introduced small multi-member districts for its first democratic elections. What are the consequences of district magnitude in terms of the behavior of voters, the effective number of political parties, and the overall quality of representation and democracy? We still have only partial answers to these questions. In this paper we analyze the behavior of voters in a controlled laboratory setting, in which we change district magnitude but keep all other relevant political variables constant.

? suggests that as district magnitude increases, the proportion of voters that behave *strategically* decreases, while the proportion who votes *sincerely* for their most preferred party increases. Indeed, **?**, p. 100 claims that: "strategic voting ought to fade out in multi-member districts when the district magnitude gets much above five". This argument is similar to a claim made by **?**, p. 279 much earlier: "The general rule is that the progression from maximal manipulative impact [via strategic voting] to sheer ineffectiveness follows, more than anything else, the size of the constituency".

These intuitions might explain the patterns in the aggregate election outcome data, in that voters in small multi-member districts strategically coordinate around larger parties, which leads to a low number of wasted votes (and a closer relationship between vote-shares and seat-shares of parties) as well as a low number of parties elected to parliaments and fewer parties in government. In contrast in large multi-member districts, where most voters simply vote sincerely for their most preferred party, voting behavior fragments and government formation then becomes more difficult. In other words, voters behave similarly, in terms of their ability to strategically coordinate around viable candidates, in small multi-member districts and single-member districts, but behave quite differently in large multi-member districts, overwhelmingly voting for their most preferred parties.

We still do not know what exactly might be driving these expected empirical patterns. Following **?**, but with a contemporary twist, we should ask ourselves whether the effect of district magnitude on voting behavior is purely *mechanical* or *psychological*. The former is due to the fact that as district magnitude increases the proportion of voters whose strategic and sincere motivations coincide increases as the number of viable parties/candidates increases. Te latter instead would capture changes in voters' strategies possibly due to the increased complexity of the electoral system.

It is impossible to isolate such micro-level effects using actual voting data which may explain why the presence of strategic (non-expressive) voting does not seem to vary with district magnitude –see **?**. In cross country research, variations in district magnitude covary with a number of other factors which influence how voters behave, such as the number of parties, societal cleavage structures, institutional effects such as regime type, the level of political and economic development of a country, and so on. In within-country research, district magnitude variations also correlate with other political variations. In Spain, Brazil or Switzerland, where district magnitude varies enormously, elections are held at the same time under the same political institutions and political contexts, but the number and type of candidates and parties competing in each district vary considerably and may be endogenous to expected voting behavior. Similar, in presidential primary elections in the United States, where there is variation in whether delegates are rewarded in proportion to vote-shares or on a winner-takes-all basis, the effects of the electoral rule are confounded with the timing of the elections.

It is also difficult to investigate these particular micro-level voting processes with formal models mainly because of the problem of multiple equilibria in multi-candidate and multi-seat elections.[1] Intuitively, it seems that should voters focus their attention in the close races for the last seat, strategic voting should be invariant to district magnitude (see **?**).

We define voters being strategic or sophisticated when they act "in accordance with both their preferences for the candidates and their perceptions of the relative chances of various pairs of candidates being in contention for victory".[2] Just as pre-election polls serve to inform the electorate about the relative chances of the candidates (**?**), in our multi-election setting past voting behavior helps voters form expectations on their chances of influencing the outcome. Our goal is to analyze whether voters' behavior varies as we modify the number of candidates that need to be elected (i.e. the district magnitude).

---

[1]Cox (1994) analyzes voting equilibria under single non transferable vote and shows support for his now classic $M+1$ rule. We instead analyze voters' behavior in non majoritarian multi member elections.
[2]**?**, p. 135

We consequently investigate the effect of district magnitude on voting behavior via a laboratory experiment designed to isolate the motivations behind voter choices. Throughout we keep constant the distribution of voters' preferences and the number and policy location of parties. This allows us to clearly observe the behavior of voters under different district magnitude treatments. We build our analysis on two stylized types of behavior: (1) *sincere behavior*, where a person votes for the party that yields the highest utility regardless of information about the electoral chances of the party; and (2) *sophisticated behavior*, where a person takes into account the votes parties have received in previous elections when deciding whom to vote for. After reporting aggregate results in Section 3, we observe a large proportion of subjects who do not vote sincerely even when their sincere and sophisticated action coincides –i.e. when there is no tension between the honest expression of their preferences and the consideration of casting a 'useful' vote. To our surprise we find that among the subjects who do not vote for their preferred party many vote for the party that obtained most votes in the previous election. This further motivates the characterization of a third type of behavior: *frontrunner behavior*, where a person votes for the party that obtained most votes in the previous election. Recently, **?** also observe frontrunner voting in French presidential elections. Within the formal literature, we are only aware of **?** where it is shown that "when voters care about the winning candidate a unique responsive equilibrium exists", however "the addition of a desire to win creates multiple equilibria".

The paper is organized as follows: we first describe in detail our experimental setup in Section 2. We then summarize our main findings in Section 3 by showing the proportion of sincere, sophisticated and frontrunner votes in our different experimental treatments: we find that sincere voting increases with district magnitude, sophisticated decreases, and frontrunner voting can never be discarded in any of our treatments. Later in Section 4 we analyze individual voting behavior and classify each subject as one of our three types. We show that sincere behavior increases with district magnitude and this is due to both mechanical and psychological factors.

2. The experiment

Our experiment consists of four treatments, each corresponding to a different district magnitude: a single-member district (M=1); a two-member district (M=2); a three-member district (M=3); and pure proportional representation (M=PR). Subjects participate in 60 elections by casting a single vote for one of five parties.[3] In the M=1 treatment, a candidate from the party that receives the most votes is elected, and each subject receives a payoff from the election equivalent to his or her utility for that party. In the M=2 and M=3 treatments, we apply a form of closed-list proportional representation, where seats are allocated to the parties in proportion to their vote-shares (using the Sainte-Laguë divisor method), and each subject receives a pay-off from the election equivalent to his or her utility for the party of each candidate that is elected. Finally, in the PR treatment, each subject receives a pay-off in direct proportion to the share of votes each party receives.

Table 1 shows how we allocated 212 subjects to our four treatments.[4] Given that each subject participated in 60 elections we have 12,720 observations.

treatment	*M=1*	*M=2*	*M=3*	*PR*	**total**
number of groups	*2*	*2*	*3*	*2*	*9*
participants per group	*24,24*	*25,25*	*22,24,24*	*20,24*	*212*

Table 1. Participants and Treatments

In all treatments, the utility that subjects derived for each of the parties waas privately announced. Every 5 periods, subjects were randomly assigned a preferred policy so that the overall distribution in a bounded two dimensional policy space was uniform; political parties had fixed positions and a subject's utility for each of the five parties was assigned by assuming quadratic (single-peaked and symmetric) preferences. Voters were only told the utility they derived for each party and did not know their policy location relative to the locations of the other voters, nor the relative positions of the parties in the policy space.[5] In our experiment, voters never observe other voters' preferences, they can only infer them by observing past voting behavior.

The same procedure was used in all sessions. Instructions[6] were read aloud and questions answered in private. Students were asked to answer a questionnaire to check that they fully understood the experimental design, the seat-allocation method, and the pay-off structure for their particular

[3]Casting a vote for a single party is the most common ballot-structure in single-member as well as multi-member districts in national parliamentary elections in democracies (**?**).

[4]No subject participated in more than one session. Students were recruited through the online recruitment system ORSEE (**?**) and the experiment took place on networked personal computers in Centre for Experimental Social Sciences at Nuffield College, Oxford in November 2011. The experiment was programmed and conducted with the software z-Tree (**?**). The data and program code for the experiment are available upon request.

[5]We imposed little knowledge on the underlying structure of the experiment, to avoid favoring subjects that are comfortable with spatial models of electoral competition and can deduce the viability of each party from its spatial location.

[6]See the Appendix for the instructions for the M=2 groups and various screenshots of our program,.

treatment group. If any of their answers were wrong, we referred the participant to the section of the instructions where the correct answer was provided. Students were isolated and could only communicate through the computer terminals.

In the first election each participant was shown a screen with their utility from each of the five parties and was asked to cast a single vote for one of the parties. Abstention was not allowed. The participants were then informed of the outcome of the first election: the number of votes each party received; which candidate(s) was (were) elected; and the payoff they received from the election. The participants were then asked to vote again for one of the parties. This procedure (in which we counted the votes for each party, we assigned seats, and we informed participants about the outcome of the election and their payoff) was repeated for five elections. Then, after five elections, the participants' preferences were redrawn and the participants interacted for a further five elections, after which the preferences were redrawn again. In other words the experiment was organized as 12 sets of five rounds (60 elections in total) and for each set of elections, participants' preferences and party labels were redrawn.

At the end of the last election, the computer randomly selected four elections and subjects were paid the profits they obtained in those four elections (in Pence Sterling). In addition, subjects received a show-up fee of £3 for taking part in the experiment. At the end of each session, participants were asked to fill in a questionnaire on the computer and were given their final payment in private. Session length, including waiting time and payment, was around 90 minutes. The average payment was 15.71 GBP (approximately 25 USD).

3. Aggregate Results

As an illustration of our results, Table 2 shows the outcomes of elections 11 to 15 for one of the groups in each treatment (we report the votes received by each of the five parties and the candidates assigned to each party). The results with M=1 shows voters *coordinating* around the first two parties A and B, with support for the other three parties declining over time. This suggests a high proportion of sophisticated behavior, with voters whose preferred party was C, D or E realizing that their most preferred party had no chance of winning. When M=2, voters appear to *coordinate* around three parties (A, B and C) as **?** would have predicted, with a non-negligible proportion of voters still supporting the two uncompetitive parties (D and E); in contrast, when M=3 party C ran away with the election after a few rounds. Finally, in the fully proportional (control) treatment group, there were considerable shifts in voting patterns, despite the fact that the optimal behavior for each participant in this treatment was to vote sincerely. Most strikingly we observe a tendency to vote for the frontrunner candidate. We will leave this aspect aside for now and will revisit it at the end of this Section.

election	**M=1**	**M=2**	**M=3**	**PR**
11	(6,7*,4,4,3)	(7*,6,6*,3,3)	(5*,5*,6*,4,4)	(5,6,3,6,4)
12	(7,12*,4,1,0)	(8*,7,8*,1,1)	(4*,7*,10*,1,2)	(6,7,3,5,3)
13	(8,12,*3,1,0)	(7,8*,7*,1,2)	(4*,3,16**,0,1)	(9,7,2,4,2)
14	(8,14,*2,0,0)	(7,8*,8*,0,2)	(6*,3,13**,1,1)	(11,6,1,5,1)
15	(9,14*,1,0,0)	(7,8*,8*,1,1)	(3,4*,14**,2,1)	(13,5,2,3,1)

Table 2. Sample of election results for each treatment.
In each cell we indicate the votes received by parties A, B, C, D, and E (resp.) and we identify with one or two stars (* or **) the parties that obtained 1 or 2 candidates, respectively.

In what follows we classify a vote as *sincere* when the subject votes for his/her most preferred party, the one that yields maximum payment.[7] A *sophisticated* vote is instead a vote in which the subject not only considers his/her preferences for all parties but also the likelihood that his or her vote will be pivotal. In the Appendix we offer a detailed explanation of the computation of expected utilities when voting for each party. For this purpose we build on **?**, and assume that our subjects best respond to the probability that each voter will vote for each of the parties when these probabilities coincide with the previous period frequency of votes.[8] Note that our definition of a sophisticated vote differs from the one in the electoral studies literature. Following game theoretic conventions, we define a sophisticated vote simply as the vote which maximizes expected utility (takes into account the utility for each candidate as well as the probability tat the vote is pivotal). Instead, the empirical electoral studies literature often treats a sophisticated (or strategic) vote as the vote which maximizes expected utility and does not coincide with an expressive vote. In our definition, sincere and sophisticated can either occur uniquely or can be present at the same time.

[7]This kind of behavior is often also referred as expressive, honest or straightforward vote. See **?** or **?**.
[8]**?** is the only experimental work we are aware of that uses past election information as the cue from which voters form expectations about their probability of being pivotal. From a game theoretical perspective, we are assuming that subjects engage in a type of fictitious play, where subjects myopically best respond to previous period play.

The frequency of sincere and/or sophisticated voting behavior in our four treatments is shown in Table 3. The long standing hypothesis in the electoral studies literature that sincere voting should increase with district magnitude seems to find little support in our aggregated data.

	M=1	M=2	M=3	PR
% sincere	*70.5*	*72.4*	*72.7*	*89.8*
% sophisticated	*84.2*	*74.1*	*70.5*	*89.8*
% observed both	*64.5*	*67.0*	*66.6*	*89.8*
% predicted both	*72.5*	*85.5*	*91.2*	*100*

Table 3. Frequency of Types of Behavior by Treatment.
"Observed Both": a subject is *both* sincere and sophisticated.
"Predicted Both": the sincere vote and the sophisticated vote coincide

Most surprisingly, we see a decrease in sophisticated voting as we increase district magnitude from 1 to 3. We expected the opposite because as we increase district magnitude, the likelihood that a sophisticated vote will coincide with a sincere vote increases. This increased coincidence in both types of voting is captured in the table last row: the 'predicted' values are the percentage of observations where voting sincerely for the most preferred party in an election can also be classified as a sophisticated vote for the the party that maximizes the expected payoff.

Based on the classical distinction between mechanical and psychological effects of electoral rules, we identify the effects of district magnitude on voting behavior using this typology. We define *the mechanical effect of district magnitude on voting behavior* as the change in sincere behavior not caused by a voters changing their *strategy* but due to the increased likelihood a sophisticated vote is sincere. Instead, the changes in *strategy* are coined *the psychological effect of district magnitude on voting behavior*. Consider for instance a sophisticated voter whose preferred party is the third ranked in number of votes: when district magnitude is 1 his/her vote is less likely to coincide with his/her sincere one than when district magnitude is 2. Note that the *strategy* of the voter is not changing with district magnitude yet the way we classify his actions is.

Following the above distinction, the results in Table 3 seem to suggest that the increase in sincere voting (as we increase district magnitude) is not driven by a mechanical effect but by a psychological one: possibly due to the higher complexity of computing the correct sophisticated action when district magnitude is large, voters seem less likely to vote sophisticatedly and instead vote sincerely.

Something that seems puzzling in table 3 is the large difference between the percentage of observations that are both sincere and sophisticated and the situations that are predicted to be so. It seems that in situations in which both sincere and sophisticated actions coincide, the voter should have no conflict about supporting his preferred party. However we observe that around a 20% of subjects fail to choose this action when it is optimal to do so! Whom are they voting for? To our surprise we see that 50% of the subjects who do not vote for their most preferred party (when sincere and sophisticated actions coincide) are voting for the party that obtained most votes in the previous round of elections. We consequently define a third type of voting behavior: we classify a vote as *frontrunner* when the subject is voting for the party that obtained the most votes in the previous period of play.[9] Together with this third classification, our three types of behavior describe more than 90% of all vote choices.

Further evidence towards frontrunner behavior is found in our control treatment with a fully proportional electoral system where voting sincerely is the dominant strategy. Below we depict the data of our control treatment aggregated by each set of five elections (recall that preferences are redrawn every five periods) whenever the frontrunner action does not coincide with the sincere one. That is, we drop observations in which the subject's preferred party is the most voted party in the previous election. We see that most votes are sincere yet a large proportion (around 11%) are voting for the frontrunner of the previous election when this is clearly defined in rounds 2, 3, 4 and 5.

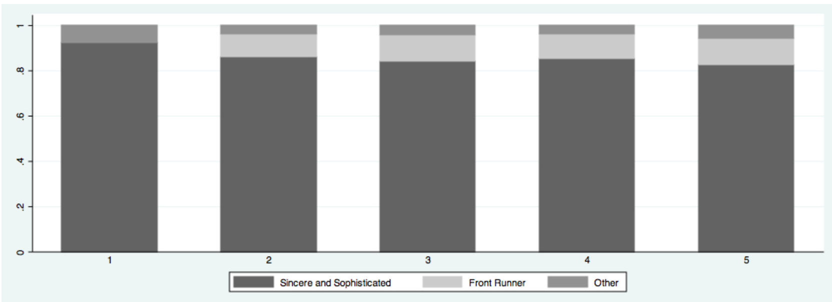


FIGURE 1. Frequencies of Votes by Election Round when PR and Voting Sincerely is not Voting for the Frontrunner

[9]Frontrunner voting is only defined for election rounds 2,3,4 and 5 given that in the first election (round 1) preferences have been redrawn and there is no previous period of play with the same preferences. In common value situations, voting for the winner can be understood in terms of herding **?**, information aggregation ((**?**), or favoring a stable governing party **?**. There is no room for such rationalizations in our setup.

A key problem when analyzing our data (as with actual voting data) is that many observations can be simultaneously classified as more than one type. Consider for instance a subject whose preferred party is the one that obtained most votes in the previous period when district magnitude is 1: when the subject votes expressively, he is also voting for the frontrunner and most likely his vote also coincides with his sophisticated action. Below we look at the subsample of observations in which the three types of actions do not coincide.[10].

	M=1	**M=2**	**M=3**
% sincere	*14.4*	*12.4*	*22.5*
% sophisticated	*70.5*	*53.7*	*42.8*
% frontrunner	*7.4*	*24.8*	*22.5*
% other	*7.7*	*9.2*	*12.3*
observations	*312*	*218*	*138*

Table 4. Frequency of Types of Behavior by Treatment when Sincere, Sophisticated and Frontrunner do not coincide.

The last row in Table 4 above shows yet another manifestation of the mechanical effect of district magnitude: as district magnitude increases it is more likely that the sincere and sophisticated actions coincide, thus our sample becomes thinner. Even when we now only consider less than 7% of our observations, the patterns we observed earlier are preserved in this subsample: sincere voting is greater when district magnitude is 3 rather than 1 and sophisticated voting decreases. Possibly due to the increased complexity of the voting rule we see frontrunner voting and other *behavior* increasing with district magnitude.

By construction, the disjoint set never includes the first round of elections when subjects vote just after preferences have been redrawn. In that case there is no previous information so being sincere and sophisticated coincide in all treatments. As we discussed above when analyzing our control treatment, in such elections it is dominant to vote sincerely for the preferred party. Indeed we observe a high incidence of such behavior in our data from the very first election –82% of our first election observations are sincere– and such behavior increases as the experiment unfolds reaching 90% of sincere observations in the first election of the last round of five elections (election 56).

The patterns we observe are clearly suggestive of both mechanical and psychological effects on voting behavior in our experiment. The increase in sincere voting is (at least partly) driven by the

[10]This is indeed the approach followed by many studies in the literature, see for example (**?**)

mechanical effect of district magnitude. However, if only mechanical effects are present in our data we should also be observing an increase in the number of votes that are qualified as sophisticated. Moreover, frontrunner voting should not increase with district magnitude. Overall there seems to be a tendency to move away from sophisticated actions, towards non-rational behavior (as expressed by sincere and frontrunner voting) as we increase district magnitude. A possible explanation is that as district magnitude increases it becomes exponentially difficult to compute the *right* sophisticated action.

Both tables 3 and 4 are indicative of the heterogeneous effects of district magnitude in our population: if all subjects were sincere, we should observe 100% of observations as sincere, sophisticated voting should increase with district magnitude due to the mechanical effect and frontrunner voting should remain unchanged. Instead, if all subjects were sophisticated, sincere voting should increase with district magnitude, sophisticated voting should always be at 100%, and frontrunner voting should decrease with district magnitude because more parties become viable so less voters need to favor the frontrunner candidate. In the next section we analyze in detail individual voting decisions so that we can understand whether district magnitude has a systematic effect on the heterogeneous behavior of subjects. Our goal is to measure the relative power of the three types of behavior for a representative voter and see how district magnitude influences the relative weight of the different motivations.

4. Individual behavior

Our initial specification assumes that subjects are either sincere or sophisticated: the (unobserved) indicator function z_{it}^{soph} (z_{it}^{sin}) takes value 1 if in round t subject i behaves in a sophisticated (sincere) way, and 0 otherwise. Our goal is to estimate $z_{it} = (z_{it}^{soph}, z_{it}^{sin})$ for all subjects and rounds, i.e. for each of our treatments we want to identify what is the unconditional probability that a subject is of each type.[11] This unconditional probability can also be interpreted as the proportion of subjects who are basing their choice solely on their preferences (sincere types), and the proportion who are also taking into into account their probability of being pivotal (sophisticated types). We will use both interpretations indistinctly throughout the text.

[11]Note that in this first specification we have that $z_{it}^{soph} = 1 - z_{it}^{sin}$.

Let u^*_{ijt} be the utility subject i derives from party j in election t; and v^*_{ijt} be the expected utility i derives from voting for party j in election t. y_{ijt} is a dummy variable that takes value 1 when subject i votes for party j in round t:

$$y_{ijt} = 1 \text{ if } u^*_{ijt} \geq \max(u^*_{i1t}, u^*_{i2t}, ..., u^*_{i5t}) \text{ and } i \text{ is sincere)}$$

$$y_{ijt} = 1 \text{ if } v^*_{ijt} \geq \max(v^*_{i1t}, v^*_{i2t}, ..., v^*_{i5t}) \text{ and } i \text{ is sophisticated)}$$

$$y_{ijt} = 0 \text{ otherwise}$$

where $u^*_{ijt} = \bar{u}_j + \alpha u_{ijt} + \varepsilon^u_{ijt}$ when $z^{ssin}_{it} = 1$ and $v^*_{ijt} = \bar{v}_j + \beta v_{ijt} + \varepsilon^v_{ijt}$ when $z^{soph}_{it} = 1$. We assume $\varepsilon^u_{ijt}, \varepsilon^v_{ijt} \sim$ type I extreme value. We model the probability that subject i votes for j in period t as a multinomial logit (MNL):[12]

$$p_{ijt}(u_{ijt}, v_{ijt}; z_{it}) = \frac{e^{\bar{u}_j + \alpha u_{ijt}}}{\sum_{k=1}^J e^{\bar{u}_k + \alpha u_{ikt}}} \qquad \text{if } z^{soph}_{it} = 0$$

$$p_{ijt}(u_{ijt}, v_{ijt}; z_{it}) = \frac{e^{\bar{v}_j + \beta v_{ijt}}}{\sum_{k=1}^J e^{\bar{v}_k + \beta v_{ikt}}} \quad \text{if } z^{soph}_{it} = 1$$

Note that modeling utilities with a MNL allows us to take into account cardinality of preferences: the likelihood of voting for one party depends not only on the ranking of the party in an ordinal setting, but on the cardinal utility this party yields relative to all other parties.

Given that the type $z = (z^{sin}, z^{soph})$ is unobserved, we can at most infer the probability that each subject is of each type. We estimate z's using the Expectation-Maximization algorithm.[13] **?** elegantly summarizes what we do: given initial values of the parameters, we first calculate the conditional probability that an individual is a particular type; using these conditional probabilities as weights, we treat types as observed and we maximize the (now) additively separable log-likelihood function; given the new parameter estimates, we then update the conditional probabilities of being each of the types and iterate until convergence (**?**).[14] Since the likelihood of being a particular type is always strictly positive, we find that even when a vote seems to unambiguously be of type z, the algorithm can only assign it a probability of being of type z which is arbitrarily close to 1, but

[12] The reported results impose that the constants $\bar{u}_j$ and $\bar{v}_j$ are equal $\forall j$. That is, the mean propensity of voting for each party is the same regardless of the type (we find analogous results when we relax this condition).

[13] See **?** and **?** for detailed descriptions of the algorithm.

[14] Our strategy is similar to that of **?**: in both cases types are unobserved and from the data we can at most infer the probability that each subject is of each type at each round. Nevertheless, whereas **?** use Bayesian methods, we use the computationally simpler EM algorithm, as in **?**.

never 1 (and, in some cases, given the available information, even not necessarily close to one). In the Appendix we specify all computational details of the EM algorithm.

Throughout we do a double exercise: we first pool *all* observations for each district magnitude, and run the EM algorithm just once per district magnitude –i.e. we run it four times (M=0,1,2,3). For each action we find $\hat{z}_{it}$ and we report the average across all individuals and periods (panels A in tables below). In other words, we first assume that all observations are independent. Second, given that for each subject we have many observations, we run the algorithm once per subject.[15] In this case we also report the averages across all subjects (panels B in tables below). This second set of regressions can be interpreted as controlling for individual fixed effects given that we are estimating z's for each subject separately. We can advance that results differ slightly among our two types of regressions but the patterns are the same regardless of which specification we use.

Table 5 reports the results for the two types model. For each district magnitude we report the percentage of individuals that are classified as sincere and those that are classified sophisticated –we also report the % of correct vote predictions using $\hat{z}_{it}$ and the parameters of the model. In both panels A and B, the estimated proportion of sincere types sharply increases with district magnitude. This is a strong result, yet we have to remember that for a great proportion of our observations, sincere and sophisticated actions are observationally equivalent. This means that we are forcing our algorithm to decide between two types when both are potentially correct. Note that there are no results for PR because in that case both sincere and sophisticated behavior are equivalent. Note also that the percentage of observations we correctly classify decreases with district magnitude but always reaches very high scores.

We check statistical significance on the increase in the proportion of sincere voters by running a Kolmogorov-Smirnov test: for each district magnitude, the null hypothesis is that the cumulative distribution function of the distribution of z_{ijt}^{sin}s for $M+1$ is not larger than the same one for M. We do not report the results but we find this hypothesis rejected in both cases. That is, the distribution of z_{it}^{sin}s increase with district magnitude.[16]

[15] Given that we do not include first round observations when sophisticated and sincere coincides we have 48 observations per subject.

[16] Note that panel B has slightly less observations than panel A. This is because in each treatment we had to drop one subject to reach convergence in our algorithm.

	M=1	M=2	M=3
Panel A: Whole sample			
% sincere	*21.98*	*33.49*	*55.99*
% sophisticated	*78.02*	*66.51*	*44.01*
% correct	*87.24*	*76.79*	*68.60*
observations	*2,304*	*2,400*	*3,360*
Panel B: Average across subjects			
% sincere	*30.56*	*36.24*	*51.50*
% sophisticated	*69.44*	*63.76*	*48.50*
% correct	*92.07*	*81.30*	*76.29*
observations	*2,256*	*2,352*	*3,312*

Table 5. Proportion of subjects of each type (2 types)

Following our previous discussion, we next introduce a third type that captures whether subjects vote for the frontrunner party in the previous election. We denote it $z^{FR} = 1 - z^{sin} - z^{soph}$. Formally, for this type y_{ijt} is equal to 1 if in the previous round party j obtained the most votes, and is equal to 0 otherwise. The utility of this type of voter is given by:

$$p_{ijt}(u_{ijt}, v_{ijt}, FR_{ijt}; z_{it}) = \frac{e^{\bar{\gamma}_j + \alpha FR_{ijt}}}{\sum_{k=1}^{J} e^{\bar{\gamma}_k + \alpha FR_{ikt}}} \qquad \text{if } z_{it}^{FR} = 1$$

where FR_j is a dummy that takes value 1 if j was the party with most votes in the previous round and 0 otherwise.

Table 6 shows the results for the three types model. We consistently find the same patterns: in both panels A and B, the proportion of sincere actions increases with district magnitude. This result is statistically significant when using the Kolmogorov-Smirnov test. Interestingly, it is worth noting that around one in every ten votes is classified as frontrunner. This is indeed the proportion we observed in our control treatment with a fully proportional electoral system.

As was noted in Section 3, one of the main issues with our experimental design is that many actions are observationally equivalent. As a robustness check we now look at the subsample of observations in which the three types of action do not coincide. Note that there is no self-selection into this subsample: utilities are randomly assigned every five elections. Besides, expected utilities are computed given the behavior of *all* subjects in the previous round so there is no way a subject can choose a particular ordering of parties according to utilities or expected utilities. The median

	M=1	M=2	M=3	PR
Panel A: Whole sample				
% sincere	*19.99*	*20.41*	*35.86*	*85.91*
% sophisticated	*71.04*	*63.73*	*53.89*	
% frontrunner	*8.97*	*15.86*	*10.25*	*14.09*
% correct	*91.32*	*79.58*	*78.24*	*91.10*
observations	*2,304*	*2,400*	*3,360*	*2,112*
Panel B: Average across subjects				
% sincere	*27.30*	*29.41*	*41.62*	*86.95*
% sophisticated	*63.67*	*55.82*	*38.50*	
% frontrunner	*9.03*	*14.77*	*19.88*	*13.05*
% correct	*95.30*	*87.07*	*85.45*	*94.41*
observations	*2,256*	*2,352*	*3,312*	*2,112*

Table 6. Proportion of subjects of each type (3 types)

number of observations per individual is 16, 9 and 5 in the treatments with district magnitude 1, 2, and 3, respectively. This once again captures the mechanical effect of district magnitude. Table 7 reports results for the subsample of disjoint observations (the observations for PR are considered disjoint when sincere does not coincide with frontrunner): once again we observe that the unconditional probability that an action is sincere increases. Even when the average values do not seem to confirm this last statement (for M=2), the Kolmogov-Smirnov test rejects that the cumulative distribution function for $(\hat{z}^{sin}|M = 3)$ is not larger than the c.d.f for $(\hat{z}|M = 2)$ –this contradicts the reported average levels because the distribution of $\hat{z}^{sin}$'s is more sparse when $M = 2$. A key lesson we take from this last table is the fact that the cases $M = 2$ and $M = 3$ are almost equivalent and distinctly different than $M = 0$ and PR. Finally, it is surprising to see that the propensity to cast a frontrunner vote is now much larger than we observed in the full sample.

4.1. **Robustness Check.** We finally show that our results are not driven by the particular choice of the number of types or the EM algorithm. We do not run a conditional probit as this model estimates the values of the parameters for the average subject and we want to allow for subject heterogeneity –the mixed logit is our model of choice given that we can estimate parameters for each individual and report the distribution of such parameters among our population.

In our approach we have implicitly assumed that the value of each of our parameters is the same within each type and we have then assigned each of our subjects a probability of being each type

	M=1	M=2	M=3	PR
Panel A: Whole sample				
% sincere	*14.47*	*21.16*	*20.24*	*86.34*
% sophisticated	*79.39*	*57.17*	*59.63*	
% frontrunner	*6.14*	*21.67*	*20.13*	*13.66*
% correct	*79.10*	*75.69*	*63.77*	*92.17*
observations	*312*	*218*	*138*	*1405*
Panel B: Average across subjects				
% sincere	*13.95*	*21.96*	*20.67*	*86.67*
% sophisticated	*80.14*	*57.24*	*57.75*	
% frontrunner	*5.91*	*20.80*	*21.57*	*13.33*
% correct	*78.71*	*76.46*	*63.55*	*GUILLEM*
observations	*312*	*218*	*138*	*1405*

TABLE 7. Proportion of subjects of each type (3 types) when M=1,2, or 3 and Sincere, Sophisticated and Frontrunner do not coincide or when PR and Sincere and Frontrunner do not coincide

(z's). Instead, the mixed logit assumes that all subjects simultaneously take into account utilitarian, sophisticated and frontrunner considerations (and allows subjects to place different weight on each component). Formally we have that:

$$p_{ijt}(u_{ijt}, v_{ijt}, FR_{ijt}) = \frac{e^{\alpha_j + \beta^i_{sin} u_{ijt} + \beta^i_{soph} v_{ijt} + \beta^i_{FR} FR_{ijt}}}{\sum_{k=1}^{J} e^{\alpha_k + \beta^i_{sin} u_{ikt} + \beta^i_{soph} v_{ikt} + \beta^i_{FR} FR_{ikt}}}$$

where

$$\begin{pmatrix} \beta^i_{sin} \\ \beta^i_{soph} \\ \beta^i_{FR} \end{pmatrix} \sim \mathbf{N} \begin{pmatrix} \beta_{sin} & & \sigma^2_{sin} & \sigma_{sinsoph} & \sigma_{sinFR} \\ \beta_{soph} & , & \sigma_{sinsoph} & \sigma^2_{soph} & \sigma_{sophFR} \\ \beta_{FR} & & \sigma_{sinFR} & \sigma_{sophFR} & \sigma^2_{FR} \end{pmatrix}$$

We assume that the distribution of the vector of parameters β follows a multivariate normal distribution. Below we report the estimated means of our mixed logit computations (covariance matrices are available upon request).

Note that the scale of the utility and expected utility values are not the same so parameters in our four different specifications are not directly comparable. However, we can observe similar trends as those we observed earlier: sincere considerations (as captured by β_{sin}) increase with district

	M=1	M=2	M=3	PR
Mean β_{sin}	*2.97*	*1.92*	*3.78*	*14.99*
	(0.50)	*(0.26)*	*(0.32)*	*(1.20)*
Mean β_{soph}	*90.84*	*80.24*	*63.61*	
	(9.57)	*(6.14)*	*(6.68)*	
Mean β_{FR}	*1.66*	*0.77*	*1.02*	*1.98*
	(0.15)	*(0.12)*	*(0.11)*	*(0.88)*
% correct	*78.26*	*70.75*	*66.90*	*85.23*
observations	*2,304*	*2,400*	*3,360*	*2,112*

Table 8. Mixed Logit Results
(standard errors in brackets)

magnitude in all cases apart from $M = 1$ to $M = 2$; sophisticated considerations (as captured by β_{soph}) decrease with district magnitude in all cases; and finally, frontrunner considerations are always present. In terms of the percentage of correct observations we see that our EM algorithm has a much greater predictive power than the mixed logit model. **?** run a similar comparison with the mixed logit model and define the *Proportional Reduction in Error* (PRE) as:

$$\mathbf{PRE} = \frac{\#\{\text{correct types model}\} - \#\{\text{correct mixed logit }\}}{N - \#\{\text{correct mixed logit}\}}$$

where N is the total number of observations. We can compute this score for each of our treatments and we find that the relative improvement of the EM algorithm with respect to the more standard mixed logit is always above 30% and reaches its maximum for the M=1 treatment in which PRE is 60.1%.[17]

5. Conclusions

A widely-held assumption in political science is that non-sincere voting should be lower in larger districts. Yet, formal work on the probability of being pivotal as well as actual voting data from elections suggests that voters in multi member districts are just as sophisticated as they are in single member districts. One problem for empirical research, however, is that it is almost impossible to isolate the effect of district magnitude on voting behavior independently from voters' preferences

[17]PRE values are 60.1%, 30.2%, 34.3% and 39.7% for DM=1,2,3 and PR, respectively.

or parties' positions. We hence designed a lab experiment to isolate this effect, by varying district magnitude while keeping voters' preferences and parties' positions constant.

We find a decrease in levels of sophisticated voting as we increase district magnitude. We also find evidence of a mechanical effect of district magnitude on the propensity to vote sincerely for a voter's most preferred party. As district magnitude increased, the proportion of voters who found that their most preferred party would also now wield them the highest expected utility (and hence that sincere and sophisticated motivations coincided) also increased. Nevertheless, not as many participants voted for their most preferred party when their sincere and sophisticated motivations coincided as we expected.

We see that sophisticated voting does not increase with district magnitude and there is a high incidence of voting for the winner of the previous election (the frontrunner) in all our treatments (we even find the latter to be the case in our control with fully proportional elections -above 10% of such observations). Following our classification of mechanical and psychological effects we have indeed seen that as we increase district magnitude we find that there is an increase in sincere voting possibly due to a mechanical effect. However, the decrease in sophisticated voting and the increase in frontrunner voting can only be classified as a psychological effect of the increase in district magnitude.

This last aspect, the presence of frontrunner voting, has not received much attention in the empirical and theoretical studies of voting, however, as **?** points out: "voting for the winner is no less plausible than the assumption that voters believe they can be pivotal".

6. Appendix

6.1. **Instructions.** (treatment M=2)
Thank you for agreeing to participate in our voting experiment. The sum of money you will earn during the session will be given privately to you at the end of the experiment. From now on (and until the end of the experiment) you cannot talk to any other participant. If you have a question, please raise your hand and one of the instructors will answer your questions privately. Please do not ask anything aloud! You belong to a group of 25 participants with whom you will interact for 60 elections. The rules are the same for all participants and for all elections. In each election the group will vote to elect two candidates. The winning candidates will be selected by a form of proportional representation, where each party will win seats in proportion to their share of the vote. After each election you will be announced the outcome and your "profit" in such election. At the end of the experiment you'll be asked to answer a questionnaire.

6.1.1. *Voting procedure.* The party with the most votes wins the first seat, and its vote-total is then divided by 3. The party with the highest remaining votes wins the second seat. In the case of a tie, the winner is determined randomly. As an illustration consider the following example:

Party	A	B	C	D	E
Votes	2	**9**	2	**7**	5
Votes÷		3			

As a result, parties B and D each obtain a candidate because 9 and 7 are the highest numbers. Now consider a different example where parties obtain the following number of votes:

Party	A	B	C	D	E
Votes	**15**	1	4	2	3
Votes÷	5				

In this second example, Party A obtains 2 candidates because 15 and 5 are the highest numbers.

6.1.2. *Profits in each election.* The profits you receive in each election depend on the candidates elected by the group regardless of whether you voted for any of them. Your profit will be equal to the sum of your valuation of the party of each elected candidate. The table below shows five hypothetical valuations for each of the five parties:

Party	A	B	C	D	E
Your valuations	500	1200	100	1800	500

So, if the 2 candidates from party A are elected, you obtain a profit of 1000 (= 500 + 500). Alternatively, if one candidate from party C is elected, and one candidate from party D is elected you obtain a profit of 1900 (= 100 + 1800). It is important to note that (a) your valuations are different from the valuations of all other voters; and (b) that no other voter knows the valuations of any other voter.

6.1.3. *Final Payment.* At the end of the last election, the computer will randomly select 4 elections and you will earn the sum of the profits on those elections in pennies. Additionally you will be paid three pounds for taking part in the experiment.

6.1.4. *Questionnaire.* (prior to the beginning of the session)
1. When the winners of an election are known, do you know your profit in such election? YES/NO
2. When the winners of an election are known, do you know the profit of any other participant? YES/NO
3. Imagine that party A obtains more votes than party B. Could it ever be the case that party B obtains more candidates than party A? YES/NO
4. Imagine a situation where the votes obtained by each party are given by the table below. What would the outcome of the election be?

Party	A	B	C	D	E
Votes	15	2	1	7	0

- Two candidates from party A
- One candidate from party D and one from party E
- One candidate from party A and one from party D
- Two candidates from party B

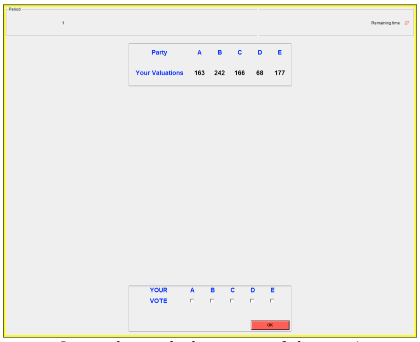
5. Consider a situation where your valuations and the votes obtained by each party are given by the table below. Imagine you voted for party B, what would your profit be?

Party	A	B	C	D	E
Votes	15	2	1	7	0
Your Valuations	500	1200	100	1800	500

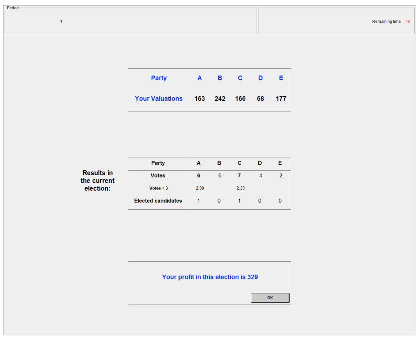
- 1000 or 1200 or 2400 or 0 or 2300

6. Who is going to be paid at the end of the experiment?

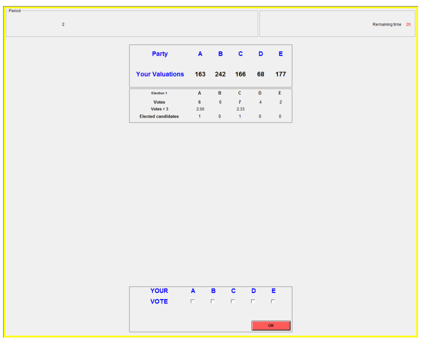
- No one
- 1 person according to his/her profit in various elections
- 2 people according to their profit in various election
- Everyone according to his or her profit in four elections
- Everyone according to his or her average profit throughout the experiment



Screenshot at the beginning of election 1



Screenshot after election 1



Screenshot at the beginning of election 2

FIGURE 2. Screenshots of the Ztree program for the treatment M=2.

6.2. **Location of Political Parties.** In all our sessions, the location of the parties and the voters were re-drawn after each set of five elections. We alternated between two types of party locations, shown in the Figure below. So, elections 1 to 5, 11 to 15, 21 to 25, 31 to 35, 41 to 45, and 51 to 55 where held with the Type A locations of the parties. And, elections 6 to 10, 16 to 20, 26 to 30, 36 to 40, 46 to 50, and 56 to 60 where held with the Type B locations of the parties. Note that in the Type B elections we assumed radial symmetry so that there were many more players that were indifferent between various parties. The labelling of parties (from A to E) was randomly allocated in each set of five elections. Given that the participants were not aware of the two-dimensional location of their preferences and the relative location of the parties, when we switched between the two types of party locations we simply announced that the preferences had been redrawn. We clearly stated that the next five elections were independent from the previous five.

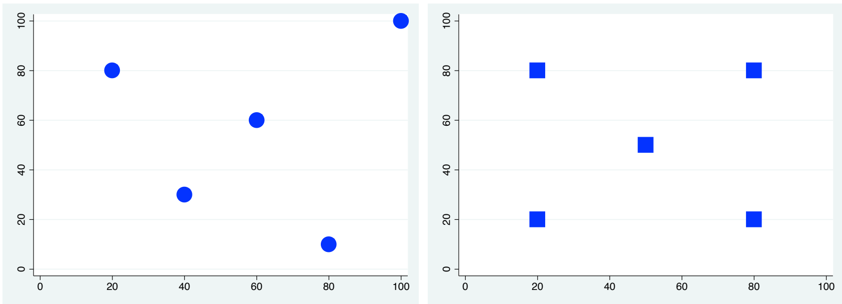


FIGURE 3. Location of political parties. In all treatments, subjects' preferred policies are uniformly distributed in the two dimensional policy space $[0, 100] \times [0, 100]$.

6.3. **Sophisticated voting.** A subject's *sophisticated* vote takes into account not only his or her utility but also the probability his or her vote will be pivotal. The latter component depends on the subject's beliefs about the distribution of votes. Using the *theory of voting equilibria* **?** we will assume that any subject assumes that the probability that any other subject votes for party k is equal to the vote share obtained by each party in the previous round of election. That is, $p_k = \frac{v_k}{v_A+v_B+v_C+v_D+v_E}$, where $(v_A, v_B, v_C, v_D, v_E)$ are the votes received by each of the five parties in the previous round of elections.[18] Knowing these probabilities, we can compute the probability that party A gets a votes, candidate B gets b votes, and so on, out of n-1 voters using the following multinomial probability:

$$f(a, b, c, d, e) = \frac{(n-1)!}{a!b!c!d!e!} \cdot p_A^a \cdot p_B^b \cdot p_C^c \cdot p_D^d \cdot p_E^e$$

where $a + b + c + d + e = n - 1$. Given a distribution of votes in the previous election, there is then a function that assigns seats to each of the five parties (taking into account the district magnitude and the seat allocation formula, in our case Sainte-Laguë for the M=2 and M=3 treatments). We denote the seats assigned to party k given a particular vote distribution as $\Omega_k(a, b, c, d, e)$. The expected utility of a subject with preferences Î¸ when voting for party A can then be expressed as follows:

$$E\left(u_\theta(\text{vote for party A}) = \sum_{(a,b,c,d,e)\in V} \left(\sum_{k=A,B,C,D,E} \Omega_k(a+1, b, c, d, e) \cdot \theta_k \cdot f(a, b, c, d, e) \right)\right.$$

where θ is the type of the voter (the utility he assign to each party) and V is any possible combination of votes for the five parties so that they add up to (n-1).

In our experimental setting, we restrict our attention to situations in which being sophisticated is solely driven by the probability of being pivotal. There are, however, many motivations that may lead voters to vote non sincerely in real elections. Voters may use their votes to communicate their preferences (**?**). They may anticipate the best candidate to be elected given other citizens' votes (**???**). Or, they may try to influence the coalition formation in the legislature (**?????**).

[18]In the first round of elections, just after preferences have been redrawn, there is no previous round on which to condition the voting decision so the sophisticated vote coincides with the sincere one: voting for most preferred party.

6.4. **Estimation Procedure.** Suppose we want to find θ which maximizes $P(\mathbf{X}|\theta)$, $\theta = (\alpha, \beta, \pi)$. Say that we have an unobserved random variable Z, with a given realization z (WLOG, for simplicity let us assume that $z = z^{soph}$, and that π is the unconditional probability that an action is of type 'sophisticated'. Also for simplicity we will omit subscript 't'). Then

$$\hat{\theta}_0 \equiv \underset{\theta}{\text{argmax}}\ logP(\mathbf{X}|\theta) = \underset{\theta}{\text{argmax}}\ log \left\{ \sum_z P(\mathbf{X}|z,\theta)P(z|\theta) \right\} \tag{1}$$

In **?** it is shown that

$$\max_\theta\ log \left\{ \sum_z P(\mathbf{X}|z,\theta)P(z|\theta) \right\} \geq \max_\theta \sum_z P(z|\mathbf{X},\theta_\tau)\ log\ P(\mathbf{X}, z|\theta) \tag{2}$$

where θ_τ is the set of estimated parameters at iteration t. That is, given a set of estimates θ_τ, $\sum_z P(z|\mathbf{X},\theta_\tau)\ log\ P(\mathbf{X}, z|\theta)$ is bounded above by $logP(\mathbf{X}|\theta)$. Therefore, given an unobserved vector z and an educated initial guess θ_0, maximizing the RHS of (2) will yield us $\hat{\theta}$, the best approximation to $\hat{\theta}_0$.
Since in our case

$$\begin{aligned} &P_i(\mathbf{X}_i|z_i,\theta) = \left(f_i^{soph}\right)^{z_i} \left(f_i^{pty}\right)^{(1-z_i)} \\ &P_i(z_i|\theta) = \pi \\ &P_i(z_i|\mathbf{X}_i,\theta_\tau) = \hat{z}_i(\mathbf{X}_i, \alpha_\tau, \beta_\tau, \pi_\tau) \qquad \text{and} \\ &P_i(\mathbf{X}_i, z_i|\theta) = \left(\pi f_i^{soph}\right)^{z_i} \left((1-\pi) f_i^{sin}\right)^{(1-z_i)} \end{aligned}$$

we have that the best approximation to

$$\underset{\alpha,\beta,\pi}{\text{argmax}} \quad log\mathcal{LIK}(\mathbf{X}|z,\alpha,\beta,\pi) = \sum_i^N log\left\{\pi\left(f_i^{soph}\right)^{z_i} + (1-\pi)\left(f_i^{sin}\right)^{(1-z_i)}\right\}$$

is given by

$$\begin{aligned} &\underset{\alpha,\beta,\pi}{\text{argmax}} \quad \sum_{i=1}^N \hat{z}_i(.)log\left(\pi f_i^{soph}\right) + (1-\hat{z}_i(.))log\left((1-\pi)f_i^{sin}\right) = \\ &\underset{\alpha,\beta,\pi}{\text{argmax}} \quad \sum_{i=1}^N \hat{z}_i(.)log\left(\pi\right) + \hat{z}_i(.)log\left(f_i^{soph}\right) + (1-\hat{z}_i(.))log\left(1-\pi\right) + (1-\hat{z}_i(.))log\left(f_i^{sin}\right) \end{aligned}$$

which is what one actually maximizes.

6.5. **The M-STEP in detail.** We will omit subscript τ for ease of notation. At iteration τ, in order to get estimates of $\{\alpha, \beta, \pi\}$, we need to solve

$$\begin{aligned} &\underset{\alpha,\beta,\pi}{\text{argmax}} \sum_{i=1}^N \{\hat{z}_i log\pi + (1-\hat{z}_i)log(1-\pi)\} + \sum_{i=1}^N \hat{z}_i \left\{\sum_j y_{ij} log(P_{ij}|z_i = 1)\right\} \\ &+ \sum_{i=1}^N (1-\hat{z}_i)\left\{\sum_j y_{i1} log(P_{ij}|z_i = 0)\right\} \end{aligned} \tag{3}$$

Recall that

$$\begin{aligned} &(P_{ij}|z_i = 0) = \frac{e^{u_{ij}\alpha+\delta_j}}{\sum_{j=1}^J e^{u_{ij}\alpha+\delta_j}} \qquad (P_{ij}|z_i = 1) = \frac{e^{v_{ij}\beta+\delta_j}}{\sum_{j=1}^J e^{v_{ij}\beta+\delta_j}} \\ &\delta_j = [\delta_1, \delta_2, ..., \delta 5] \text{ with } \delta_1 = [0,0,0,0,0] \text{ for the base outcome} \\ &\text{and } 1_{y_{ij}} = 1 \text{ if } i \text{ votes for } j,\ 0 \text{ otherwise.} \end{aligned} \tag{4}$$

FOCs yield:

$$
\begin{aligned}
&[\pi] \Rightarrow \hat{\pi} = \frac{\sum_i^N \hat{z}_i}{N} \\
&[\alpha] \Rightarrow \frac{\partial log\mathcal{LIK}}{\partial \alpha} = \sum_i^N 1_{y_{ij}} \times \hat{z}_i \times (\Psi_{ij}|z_i = 1) \times v_{ij} + \\
&\qquad \sum_i^N 1_{y_{ij}} \times (1 - \hat{z}_i) \times (\Psi_{ij}|z_i = 0) \times u_{ij} \stackrel{!}{=} 0 \\
&[\beta_{s,j}] \Rightarrow \frac{\partial log\mathcal{LIK}}{\partial \beta_{s,j}} = -\sum_i^N 1_{y_{ij}} \times \hat{z}_i \times (P_{ij}|z_i = 1) \times x_i^a - \\
&\qquad \sum_i^N 1_{y_{ij}} \times (1 - \hat{z}_i) \times (P_{ij}|z_i = 0) \times x_i^a \stackrel{!}{=} 0
\end{aligned}
$$

where

$$
\begin{aligned}
(\Psi_{ij}|z_i = 1) &= \frac{\sum_j e^{v_{ij}\alpha + \delta_j} \times v_{ij}}{\sum_j e^{v_{ij}\alpha + \delta_j}}, \qquad \delta_j = [0, 0, ..., 0] \text{ for the base outcome} \\
(\Psi_{ij}|z_i = 0) &= \frac{\sum_j e^{u_{ij}\alpha + \delta_j} \times u_{ij}}{\sum_j e^{u_{ij}\alpha + \delta_j}},
\end{aligned}
$$

Note: Since the likelihood conditional on any type is always strictly positive, then we find that even when a vote is obviously sincere, the algorithm can only assign it a probability of being sincere which is arbitrarily close to one, but never one (and, in some cases, given the available information, is not necessarily close to one). Moreover, the probability assigned to each single action depends not only on all the information we have on that action (embedded in the likelihood function), but also on the average of all the probabilities assigned to all actions in the previous round. Therefore, even if in one particular round a voter cast a vote that was unambiguously sincere, it is likely that the probability assigned by the algorithm converges to *but is not* one. for each action, the probability that it is of type z is a weighted average of the likelihoods conditional on that action being of all possible types (sincere, sophisticate, frontrunner). More specifically, these probabilities are defined at each iteration r as follows by the EM algorithm :

$$
\begin{aligned}
&\text{Pr(action a}_{it} = \text{type } z) \text{ in iteration } r \equiv \\
&\hat{z}_{it}(r) = \frac{\Pi_z \cdot \text{Likelihood a}_{it} \mid \text{a}_{it}\text{=type } z}{\Pi_{sin} \cdot (\text{Lik. a}_{it} \mid \text{a}_{it}\text{=sin}) + \Pi_{sin} \cdot (\text{Lik. a}_{it} \mid \text{a}_{it}\text{=soph.}) + \Pi_{FR} \cdot (\text{Lik. a}_{it} \mid \text{a}_{it}\text{=FR})} \\
&z = \{\text{sincere, sophisticated, frontrunner}\} \\
&\Pi_z = \sum_i^N \sum_t^T \hat{z}_{it}(r-1)
\end{aligned} \tag{5}
$$