Does Mandatory Voting Lead to More Informed Voting?

Authors:
Jacob Meyer[a]
Lucas Rentschler[b]

September 2023

[a]Department of Economics Cornell University jm2363@cornell.edu
[b]Department of Economics and Finance Utah State University and the Center for Growth and Opportunity lucas.rentschler@usu.edu

* We thank the Department of Economics and Finance at Utah State University for generous financial support.

The Center for Growth and Opportunity at Utah State University

Abstract

We develop and experimentally test a model of voter information acquisition in nonpartisan elections, both with and without abstention. We theoretically demonstrate that allowing for abstention can increase information acquisition, provided the cost of information is not too low. Our experimental data find that voters are less responsive than predicted to increases in the cost of information. As a result, the cost required to yield higher levels of informedness when abstention is allowed is higher than predicted. Our data are well explained by agent quantal response equilibrium, which accounts for the fact that uninformed voters are often observed to vote, even when it is not rational to do so.

Keywords: Mandatory voting; Information acquisition; Costly voting

JEL Classification Numbers: C70; C90; D72; D80

1 Introduction

Citizens often rely on the party affiliation of candidates when voting (Bonneau and Cann, 2015; Rahn, 1993; Schaffner and Streb, 2002). Indeed, Rahn states: "The most powerful cue provided by the political environment is the candidate's membership in a particular political party.... The cue provided by the party label is simple, direct, and...consequential in shaping individuals' perceptions and evaluations of political candidates."

Yet there are many elections in which the ballots do not provide this basic information. These so-called nonpartisan elections often involve choosing city council members, county prosecutors, judges, natural resource managers, or sheriffs. By withholding partisan cues, the belief is that voters are more likely to acquire relevant information about candidates and make informed decisions.[1]

However, Schaffner and Streb (2002) find that only a small proportion of voters in such elections are willing to express preferences and that many of these voters are "guessing." In practice, voter participation in these elections is low relative to more high-profile races (Bowler, Donovan, and Happ, 1992; Feddersen and Pesendorfer, 1999; Ghirardato and Katz, 2006; Schaffner, Streb, and Wright, 2001). This is especially interesting since a voter's probability of affecting electoral outcomes is significantly higher in most nonpartisan elections.

Low turnout and an uninformed electorate could (and perhaps do) lead to decreased government accountability and a weak connection between voter preferences and implemented policy. One potential solution, which has been adopted in many countries, is to enact mandatory voting laws.[2] Thus, understanding the effect of such laws is of practical importance.

Would mandating voter participation increase the number of voters who become informed about political issues? We address this question by developing a theory of information acquisition, which we then test in a laboratory experiment.

[1]For example, Lim and Snyder Jr (2015) find that nonpartisan judicial election outcomes are driven by candidate quality rather than party affiliation.

[2]Currently, 21 countries have such laws in place (see https://www.cia.gov/the-world-factbook/field/suffrage/). In most countries these laws are enforced by imposing fines on those who fail to show up to the polls, but some go as far as to suspend passport privileges (Brazil) or freeze bank accounts (Bolivia).

In our model we focus on whether voters will pay to acquire the information that is costlessly available in partisan elections: party affiliation.[3] We establish that there is a symmetric perfect Bayesian equilibrium (PBE), regardless of whether voting is mandatory. In equilibrium, the model predicts that willingness to pay to learn candidate preferences is increasing in extremity of preferences, since more extreme citizens have more to lose if their preferred candidate loses.[4]

Interestingly, a mandatory voting law will increase informedness if information is cheap relative to the cost of voting. If the cost of information is high relative to the cost of voting, allowing for abstention increases informedness. The intuition behind this reversal is as follows. When abstention is permitted, the *effective* cost of information includes the cost of voting, since only informed voters would incur this cost. This lowers the level of informedness if voting is not mandatory. However, since uninformed citizens would not vote unless required to do so, the pool of expected votes is smaller relative to mandatory voting. This increases the probability of being the pivotal voter, effectively increasing the value of information. When the cost of being informed is low, the former effect dominates, while the latter effect dominates when the cost of information is sufficiently high.

Our laboratory experiments closely mirror our theoretical model. We vary the ability to abstain on a between-subject basis, and the cost of information on a within-subject basis. In a first series of experiments, we consider a low cost of information in which PBE predicts that mandatory voting will result in a more informed electorate, as well as a medium cost in which PBE predicts that allowing for abstention will result in a more informed electorate.

We find that information acquisition is higher under mandatory when the cost of information is low, in keeping with the PBE prediction. However, we also observe uninformed subjects voting when not required to do so. This, in effect, reduces the value of information when abstention is permitted, since the probability of being the pivotal voter is reduced. As such, we find that allowing abstention did not result in a significant difference in informedness for the medium cost.

[3]Note that our model is somewhat more general than the frame of nonpartisan elections and would be relevant in any situation in which a voter must incur a cost to learn which candidate is closer to her own preferences, provided there is one candidate on each side of the relevant spectrum.

[4]This is consistent with the existing literature. See, for example, Gersbach (1992); Ortoleva and Snowberg (2015); Matějka and Tabellini (2017).

To explain this behavior, we turn to agent quantal response equilibrium (AQRE).[5] This model accounts for the systematic errors we observe at both the information acquisition and voting stages of the game. Most importantly, this alternative model is able to account for the uninformed votes we observe, and thus the finding that increasing the cost of information did not result in a more informed electorate with abstention.

We further noted that a *prediction* of the AQRE model is that if the cost of information were further increased, we would then observe a reversal in informedness. In response, we ran a second series of experiments, which are identical to the first series except that we replace the medium cost of information with a high cost. That is, we suggested an alternative model on the basis of our initial data and then tested the implications of this alternative model.

As the AQRE predicted, the high cost of information led to a higher level of informedness when abstention was permitted. While this certainly does not allow us to conclude that AQRE explains behavior, it is suggestive. We view our data as a single observation in the literature that considers QRE/AQRE models in voting games. We believe that the case for these models is compelling.

Our results have important implications for nonpartisan elections. Specifically, when learning about candidates' policy platforms is sufficiently costly, mandatory voting laws can decrease voter informedness, since allowing for abstention increases the probability of being a pivotal voter, and thus the value of information. However, our experimental data suggest that this effect is substantially diluted since uninformed voters often opt to vote, despite it being irrational to do so. As such, in nonpartisan elections, mandatory voting is likely to lead to a more informed electorate unless the cost of becoming informed is quite high.[6]

Our results also have implications for the experimental literature on committee decision-making. In this literature, players are often assumed to have identical preferences and are incentivized to reach a group decision that matches some unobserved state of the world. Elbittar et al. (2020) and Großer and Seebauer (2016) study environments with endogenous information acquisition that allow abstention. Consistent

[5]Quantal response equilibrium is often used to explain behavior in voting games. See, for example, Großer and Palfrey (2019); McKelvey and Patty (2006); Tyszler and Schram (2016); Bhattacharya, Duffy, and Kim (2017); Bassi (2015); Levine and Palfrey (2007); Elbittar, Gomberg, Martinelli, and Palfrey (2020); Goeree and Holt (2005).

[6]Our study focuses on information acquisition as the outcome variable of interest, rather than analyzing welfare effects. We aim to add to the literature by providing novel insight as to how mandatory voting affects information acquisition, consciously abstracting from any welfare considerations.

with our data, they observe uninformed voting, which Großer and Seebauer (2016) call the "curse of the uninformed voter." Significantly, they observe such over-voting when voting is costless, whereas we observe it in an environment where voting costs are strictly positive. Thus, our findings suggest that this is a relatively robust phenomenon, which would be a fruitful avenue for future research.

1.1 Related Literature

Previous research has found that the introduction of mandatory voting may increase turnout rates (Hoffman, León, and Lombardi, 2017), but it is less clear if it will lead to a more informed electorate. Other studies also find empirical evidence for a causal effect of citizen informedness on turnout decisions (Lassen, 2005; Feddersen and Pesendorfer, 1999; Strömberg, 2004), but evidence of causality in the other direction is scant.

Many studies have examined information acquisition in small group elections. A number of these studies show that uninformed citizens may prefer to abstain from voting. Others find that moderate citizens are less likely to become informed voters or to vote at all.

Feddersen and Pesendorfer (1999) model an election with costless voting and asymmetric information, finding that some of the less-informed voters always prefer to abstain. They also find that level increases in aggregate citizen informedness result in higher abstention rates for both the informed and uninformed. Battaglini, Morton, and Palfrey (2008) give experimental evidence that aggregate turnout will be positively correlated with the number of informed voters.[7] In a theoretical environment, Oliveros (2013) shows that information acquisition and turnout decisions may in fact be uncorrelated for some voters.

Gersbach (1992) finds that the median voter is the least likely to purchase information in an election. Ortoleva and Snowberg (2015) show theoretically and empirically that moderate citizens have lower turnout rates compared to more ideologically extreme citizens. Matějka and Tabellini (2017) develop a theoretical model of voter information acquisition in a multi-dimensional policy election, finding that moderate voters invest less in information acquisition. They also point out that including party affiliation could lead to increased voter attention. Bhattacharya et al. (2017) study costly endogenous information acquisition in Condorcet settings as they vary group size and information precision.

[7] In a related paper Morton, Piovesan, and Tyran (2019) show that when voting is mandatory and some voters are biased, a majority voting rule can lead to poor group outcomes.

They find that subjects acquire information as predicted when information is precise, but with noisy information, subjects over-acquire information. They use quantal response equilibrium to rationalize this finding.

Other studies have compared voting laws with and without abstention, mostly with an eye to the welfare effects of mandatory voting. Börgers (2004) shows that in a costly voting environment with symmetric voter preferences and that allow for abstention, Pareto dominates mandatory voting in terms of welfare. Ghosal and Lockwood (2009) extend this work and find that mandatory voting may be welfare enhancing when facing uncertainty. Krasa and Polborn (2009) develop an alternative model that shows mandatory voting can be welfare enhancing when the support groups for candidates are sufficiently different. Krishna and Morgan (2012) study mandatory voting in a Condorcet setting and find that allowing for abstention is strictly welfare enhancing.

We are aware of only one study that examines the intersection of mandatory voting and endogenous information acquisition. Großer and Seebauer (2016) present a theory and an experiment to examine endogenous information acquisition in a Condorcet setting. In their theoretical and experimental study, voters have identical preferences and are able to acquire a costly signal about the state of the world before participating in a costless binary election. The researchers find that mandatory voting increases information acquisition relative to the case where abstention is permitted. They also observe marked levels of uninformed voting when abstention is allowed, contrary to equilibrium predictions.

2 Theory

There are $n \in \mathbb{N}$ risk-neutral voters participating in an election, where for simplicity we assume that n is odd. Each voter has a type that is an $i.i.d.$ draw of a random variable with continuously differentiable density (distribution) $v(\cdot)$ $(V(\cdot))$ on $[0,1]$. Voter i's type is denoted as v_i and is private information.

There are two candidates, A and B, who are running for office. Each candidate also has a type, which is private information. Candidate A's (B's) type is denoted γ_A (γ_B) and is independently drawn from a distribution with density $a(\cdot)$ $(b(\cdot))$ with support on $\left[0, \frac{1}{2}\right]$ $\left(\left[\frac{1}{2}, 1\right]\right)$. Voters are not initially able to distinguish which candidate is which but can each pay $c \geq 0$ to do so.

Each voter first learns her type, and then all voters simultaneously decide whether or not to pay to learn candidate identities. After this information is observed by those who opted to pay c, voters simultaneously make their voting decisions. Voting costs $\varepsilon \geq 0$. The candidate with the most votes wins, with ties broken by fair randomization. The policy implemented, γ, corresponds to the winning candidate's type.

The payoff function for voter v_i is

$$\pi_i(K, v_i, \gamma, e_i) = K - U\big(|v_i - \gamma|\big) - e_i c - w_i \varepsilon \tag{1}$$

where K is a positive constant, and $-U\big(|v_i - \gamma|\big)$ is a continuous, twice differentiable, weakly concave function with a single peak at $v_i = \gamma$. The indicator variable e_i is equal to 1 if voter v_i opts to learn candidate identities. We consider voting schemes with and without abstention, and the indicator variable w_i is equal to 1 if the voter casts a vote.

In the interest of brevity, we add two assumptions. First, we restrict attention to environments in which v is symmetric around $\frac{1}{2}$. We also assume that $\left|\frac{1}{2} - E(\gamma_A)\right| = \left|E(\gamma_B) - \frac{1}{2}\right|$. Each of these assumptions can be relaxed without substantively changing our results. Adding them allows us to focus on equilibria, which involve symmetric behavior of voters with types equidistant from $\frac{1}{2}$.

2.1 Equilibrium

As is typical in the literature, we begin by assuming agents are sequentially rational and consider perfect Bayesian equilibria (PBE). In section 4.3, we examine the possibility that players noisily best respond to their beliefs in an agent quantal response equilibrium.

We first consider the voting stage of the game. When voting is mandatory, informed voters will vote for the candidate whose expected policy is closest to their type, while the uninformed voters will effectively randomize by arbitrarily selecting one of the two indistinguishable candidates. When abstention is permitted, a rational voter never prefers to vote when uninformed, as the candidates are indistinguishable, and voting costs $\varepsilon \geq 0$. Additionally, abstaining when informed cannot be part of an equilibrium, since knowing candidate identities does not have any value if a voter would not vote when informed. Thus, when abstention is permitted, only informed types vote.

The difference in equilibrium voting behavior across voting schemes introduces an important tradeoff, which underpins much of the following results. On the one hand, allowing abstention can increase the probability that a vote is decisive (since only the informed types vote), increasing the value of information. On the other hand, making voting mandatory decreases the *effective* cost of information from $c + \varepsilon$ to c, since voters have no choice but to incur ε. As we show below, this tradeoff leads to higher (lower) information acquisition under mandatory voting when c is sufficiently low (high).

We now turn attention to the information acquisition decision, taking as given the equilibrium voting behavior described above. Notably, the value of learning candidate identities is higher for more extreme types, who will experience a higher loss than moderate types if the candidate on the opposing "side" wins the election. That is, the stakes of the election are higher for more extreme types, and thus the value of information is increasing in the extremity of type. In equilibrium, this leads to a cutpoint information acquisition strategy in which only voters with types weakly more extreme than the cutpoints become informed.[8] Fixing the strategies of all other voters, voter v_i will pay to acquire information if

$$E[\pi_i | v_i, e_i = 1] - E[\pi_i | v_i, e_i = 0] \geq 0. \tag{2}$$

Cutpoint types are defined by the pair of voter types who are perfectly indifferent to acquiring information, assuming that all other voters follow the cutpoint information acquisition strategy.

Under mandatory voting, if an interior cutpoint pair exists, the corresponding values of v_i must satisfy the following equalities for $\delta \in \{1, -1\}$:

$$\delta \cdot \frac{1}{2} \left(\int_{\frac{1}{2}}^{1} U\big(|v_i - x|\big) b(x)\, dx - \int_{0}^{\frac{1}{2}} U\big(|v_i - x|\big) a(x)\, dx \right) \binom{n-1}{\frac{n-1}{2}} \frac{1}{2}^{n-1} = c. \tag{3}$$

The left-hand side of equation (3), which we denote as $F(v_i)$, is v_i's expected marginal benefit if her vote is pivotal (i.e., voter v_i's vote is a tie-breaker, resulting in her preferred candidate being elected), weighted by the probability that her vote is pivotal. Note that the probability of being pivotal is constant. As long as the cost of information is less than $\tilde{c}^m \equiv F(0)$, there is a single voter type that satisfies this equation for each $\delta \in \{1, -1\}$, which establishes the existence and uniqueness of this equilibrium under mandatory voting.

[8] Related cutpoint equilibrium strategies are found in the theoretical literature on voting and information acquisition. See, for example, Martinelli (2006), Martinelli (2007), Großer and Palfrey (2014), and Großer and Palfrey (2019).

All voters with types in the interior interval formed by the cutpoint pair will remain uninformed, while those with types more extreme will become informed.

Proposition 1. *When voting is mandatory, there exists a unique symmetric PBE in which each voter v_i follows a cutpoint information acquisition strategy in which*

$$e_i = \begin{cases} 0 & \text{if } \; v_i \in (v_l^m, v_r^m) \\ 1 & \text{if } \; v_i \notin (v_l^m, v_r^m). \end{cases} \tag{4}$$

In this equilibrium, informed voters vote for their preferred candidate, and uninformed voters effectively randomize by arbitrarily selecting one of the two indistinguishable candidates.

Proof. See Appendix A.

A similar result follows when abstention is allowed. However, when voting is costly, uninformed voters will not vote in equilibrium. This means that the probability of being the pivotal voter will depend on the cutpoint types. We restrict attention to voter type distributions in which the probability of being pivotal is decreasing as the cutpoint types get closer to $\frac{1}{2}$. Intuitively, this simply means that the probability of being the pivotal voter decreases as more voter types participate in the election.

If an interior cutpoint pair exists when abstention is possible, the two relevant types must satisfy the following equality for $\delta \in \{1, -1\}$:

$$\begin{aligned} &\delta \cdot \frac{1}{2}\bigg(\int_{\frac{1}{2}}^{1} U\big(|v_i - x|\big) b(x)\ dx - \int_{0}^{\frac{1}{2}} U\big(|v_i - x|\big) a(x)\ dx \bigg) \times \\ &\sum_{t=0}^{\frac{n-1}{2}} \sum_{s=0}^{1} \binom{n-1}{2t+s} \binom{2t+s}{t+s} \big(1 - 2V(v_l^a)\big)^{n-1-s-2t} V(v_l^a)^{2t+s} \quad = c + \varepsilon. \end{aligned} \tag{5}$$

We denote the left-hand side of equation (5) as $G(v_i, v_l^a)$, which is again v_i's expected marginal benefit if her vote is pivotal, weighted by the probability of being the pivotal voter. For $c + \varepsilon \leq \tilde{c}^a \equiv \lim_{v_i \to 0^+} G(v_i, v_i)$, there is a single voter type that satisfies this equation for each $\delta \in \{1, -1\}$, which establishes the existence and uniqueness of this equilibrium when abstention is allowed. All voters with types in the interior interval formed by the cutpoint pair will remain uninformed, while those with types more extreme will become informed.

Proposition 2. *When abstention is possible and $\varepsilon > 0$, there exists a unique symmetric PBE in which each voter v_i follows a cutpoint strategy in which*

$$e_i = \begin{cases} 0 & \text{if } v_i \in (v_l^a, v_r^a) \\ 1 & \text{if } v_i \notin (v_l^a, v_r^a). \end{cases} \tag{6}$$

In this equilibrium, informed voters vote for their preferred candidate, and uninformed voters do not vote.

Proof. See Appendix A.

When $\varepsilon = 0$, there is a continuum of equilibria, since uninformed voters are indifferent between voting or not. We focus attention on the unique limiting equilibrium as $\varepsilon \to 0^+$, in which uninformed voters opt not to vote.

2.2 Comparing Voting Schemes

The cutpoint strategies described in Propositions 1 and 2 allow for equilibria where none, some, or all of the electorate will choose to become informed about candidate preferences, depending on the magnitudes of c and ε. If these parameters are such that at least some voters are willing to purchase information in either voting scheme, we can compare the equilibrium cutpoints to determine if allowing for abstention results in a more informed electorate.[9] The costs of both voting and information acquisition will determine the relationship between the cutpoints in the two schemes.

Denote the maximum equilibrium willingness to pay for information for the most extreme voter types ($v_i \in \{0, 1\}$) when $\varepsilon = 0$ as $\tilde{c}^m$ for mandatory voting, and $\tilde{c}^a$ when abstention is allowed. In both voting schemes, when $\varepsilon = 0$ and $c \in [0, \tilde{c}^a]$, a subset of the population will always choose to become informed when abstention is permitted. Thus, when abstention is possible, the expected number of voters who cast votes will be weakly less than the population total, increasing the probability of voter i's vote would be pivotal. For $c \in (0, \tilde{c}^a]$ the interval of uninformed voters under mandatory voting is a proper subset of the corresponding interval when abstention is permitted: $[v_l^m, v_r^m] \subset [v_l^a, v_r^a]$. If $c > \tilde{c}^a$ or $c = 0$, then

[9] Our discussion focuses on informedness as the outcome of interest. The ex-ante welfare implications of equilibrium behavior across voting schemes are straightforward in our model. Mandatory voting always reduces welfare when $\varepsilon > 0$, since all citizens are required to bear cost c, without changing the expected election outcome.

allowing for abstention has no effect on informedness: $[v_l^m, v_r^m] = [v_l^a, v_r^a] = [0, 1]$. Thus, mandatory voting will never result in a more informed electorate when the cost of voting is zero.

Lemma 1. *When* $\varepsilon = 0$, $[v_l^a, v_r^a] \subseteq [v_l^m, v_r^m] \, \forall \, c \geq 0$.

Proof. See Appendix A.

When voting is sufficiently costly, many voters will abstain when permitted to do so, in order to avoid incurring ε. In equilibrium, these voters also remain uninformed. Specifically, when $\varepsilon > \tilde{c}^a - \tilde{c}^m$, allowing for abstention will never increase voter informedness. Even when the cost of information is small or zero, the relatively high cost of voting results in a smaller proportion becoming informed than if voting is mandatory.

Lemma 2. *When* $\varepsilon > \tilde{c}^a - \tilde{c}^m$, $[v_l^m, v_r^m] \subseteq [v_l^a, v_r^a] \, \forall \, c \geq 0$.

Proof. See Appendix A.

In situations where ε is positive but small, allowing for abstention may increase or decrease voter informedness relative to mandatory voting. When c is low relative to ε, allowing for abstention will decrease voter informedness because voters will prefer to avoid ε, whereas ε does not affect information acquisition choices under mandatory voting. However, when c is large relative to ε, the increased probability of being pivotal when abstention is possible will lead to increased informedness. As a result, whether or not abstention results in a more informed electorate depends on the value of c when ε is small.

Specifically, given $\varepsilon \in \left(0, \tilde{c}^a - \tilde{c}^m\right]$, there exists a unique cost of information $\bar{c}$, such that for $c = \bar{c}$ information acquisition will be the same under either mandatory or optional voting laws. For $c \in [0, \bar{c})$, mandatory voting laws will result in a more informed electorate, and for $c \in (\bar{c}, \tilde{c}^a - \varepsilon]$, allowing for abstention will result in a more informed electorate. For $c > \tilde{c}^a - \varepsilon$, no voters will become informed under either voting scheme.

Proposition 3. *When* $\varepsilon \in \left(0, \tilde{c}^a - \tilde{c}^m\right]$, $\exists \, \bar{c}$ *such that* $[v_l^m, v_r^m] \subset [v_l^a, v_r^a]$ *for* $c \in \left[0, \bar{c}\right)$, *and* $[v_l^a, v_r^a] \subseteq [v_l^m, v_r^m]$ *for* $c \geq \bar{c}$.

Proof. See Appendix A.

2.3 Hypotheses

As Proposition 3 demonstrates, varying c can lead to interesting differences in equilibrium behavior across voting schemes. These differences are the primary motivation for our experimental design. For concreteness, we explicitly list the predictions we are primarily interested in evaluating.

First, when the cost of voting is positive, uninformed types do not vote when abstention is permitted. Of course, under mandatory voting, uninformed types must always vote, although their votes are random.

These predictions regarding voting behavior have important consequences for information acquisition. In particular, because ε is paid by all types when voting is mandatory, but is only paid by the informed types when abstention is permitted, the effective equilibrium cost of information is c under mandatory voting, and $c + \varepsilon$ when abstention is permitted. When the cost of information is low ($c < \bar{c}$), this results in more information acquisition under mandatory voting.[10] However, as c increases, so does the probability of being the pivotal voter when abstention is permitted. At $\bar{c}$, these effects balance each other, so that equilibrium information acquisition is the same with or without abstention. For $c \in \left(\bar{c}, \tilde{c}^a - \varepsilon\right]$, the increased effective cost of information under abstention has a smaller effect on information acquisition than the increased probability of being a pivotal voter, and allowing for abstention increases voter informedness.[11]

Thus, the second prediction of interest is that increasing c can reverse the ranking of information acquisition across voting schemes: when $c < \bar{c}$, information acquisition is higher under mandatory voting, and that, when $c > \bar{c}$, allowing abstention increases information acquisition.

The third prediction of interest concerns the set of types that become informed. Since more extreme voter types have more to lose, their value of information is higher. In equilibrium, provided c is not too high, this means that there is an interval of moderate types that remain uninformed, and that only the more extreme types outside this interval become informed.

[10] This holds as long as the cost of voting is not too high: $\varepsilon \in \left(0, \tilde{c}^a - \tilde{c}^m\right)$.

[11] Once c gets large enough, there is no information acquisition. Specifically, for any $c > \tilde{c}^a - \varepsilon$ no voters become informed under either voting scheme.

3 Experimental Design

In each experimental session, 15 subjects participated in a voting game for 50 rounds. The subjects were randomly sorted into groups of 5, and these groups were randomly re-matched in each of the 50 rounds. Identities were anonymous within and across rounds.

In each round, each subject was assigned a type v_i that was an *i.i.d.* draw from a discrete uniform distribution on $\{1, 2, ..., 100\}$. Types were private information.

Each group was tasked with selecting either Option A or Option B via majority vote (ties were broken randomly). The per-subject cost of voting was 2. The selection of Option A (B) would result in a policy, γ, that was an *i.i.d.* draw from a discrete uniform distribution on $\{1, 2, ..., 50\}$ $(\{51, 52, ..., 100\})$. The realized policy associated with each Option was not known to subjects when they voted. Further, the two Options were indistinguishable unless a subject paid $c > 0$.

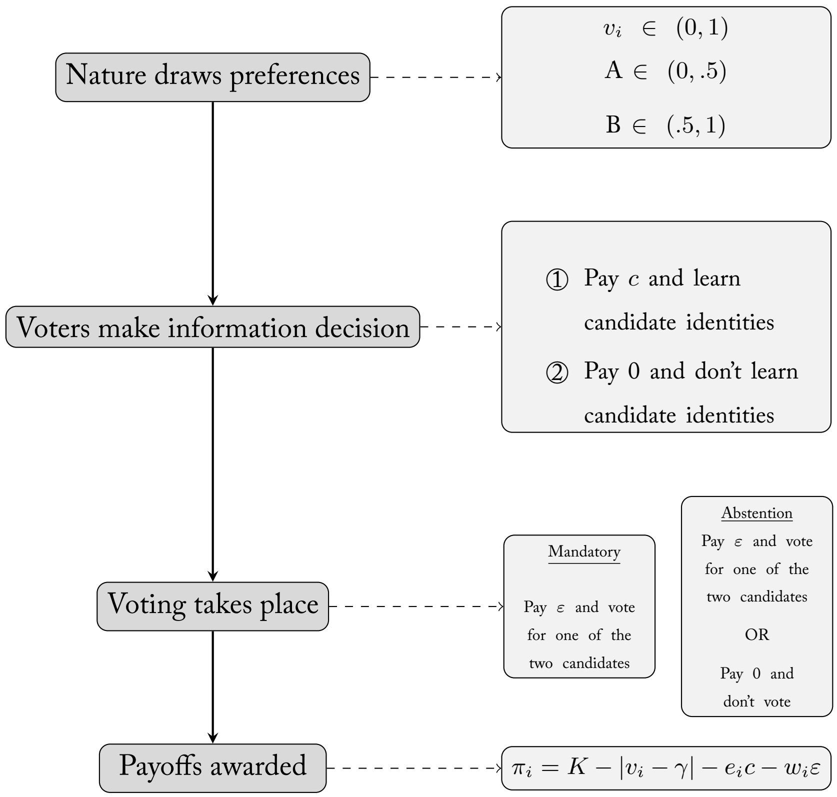
Each subject's payoff for the round was

$$\pi_i = 100 - |v_i - \gamma| - ce_i - 2w_i \tag{7}$$

where e_i is equal to one if a subject paid to get information, and w_i is equal to one if a subject voted.

After learning their types, subjects simultaneously decided whether or not to pay to observe Option identities and then proceeded to the voting stage. If a subject chose not to pay, then the Options were both labeled as "Option ?" To ensure that Option identities could not be inferred, the order of the Options was randomized in each round, and this was common knowledge. If a subject paid to learn Option identities, then the corresponding labels were revealed during the voting stage. After votes were cast, all subjects observed which Option won the election, as well as the resulting γ. Each subject also observed her own payoff for the round. The sequence of events in each round is displayed in Figure 1. At the end of the 50 rounds, 20 rounds were randomly chosen for payment.

Figure 1. Voting Game Sequence



At the voting stage, players can vote between the two candidates under mandatory voting and have the additional option to abstain under the abstention treatment.

We employed a 2×3 design. Whether abstention was permitted was varied on a between-subject basis. The cost of learning Option identities was either 3, 9, or 15, and was varied on a within-subject basis. We refer to these as the low, medium, and high-cost treatments. When abstention was permitted, we refer to treatments as AL, AM, and AH, when c was low, medium, and high, respectively. Similarly, when voting was mandatory, we refer to treatments as ML, MM, and MH, when c was low, medium, and high.

The cost of learning Option identities was constant within the first and second half of each session. In all sessions, subjects faced a cost of information $c = 3$ in one half of the experiment, and $c = 9$ or $c = 15$ in the other. We varied the order of this within-subject variation to control for order effects, and this is perfectly balanced across the sessions.

We ran a total of 24 sessions. In 12 of these, abstention was permitted, with half varying c between low and medium, and the others varying c between low and high. Voting was mandatory in the remaining 12 sessions, and half varied c between low and medium, with the other half varying c between low and high.

In all sessions, subjects were seated at terminals with dividers and communication was not allowed. Physical copies of instructions were administered and read aloud by the experimenter.[12] After reading the instructions, the experimenter administered a quiz, and each subject was required to correctly answer every question before the session could advance.[13] Sessions lasted about an hour. At the conclusion of a session, subjects were called individually to receive payment. All subjects received \$5 as a show-up fee, in addition to their in-game earnings. Earnings throughout the experiment were denominated in points, which were exchanged for dollars at a rate of one USD per 70 points. Average total earnings was \$23.87, with a range of $[\$18.96, \$28.41]$. Participants were recruited from the student population at Utah State University. No participant had any experience with any type of voting experiment. The experiment software was coded in z-Tree (Fischbacher, 2007).

In the spirit of Roth (1994), we wish to describe the process through which we arrived at our final experimental design. Initially, we relied on the PBE predictions and ran sessions 12 in which c was varied between low and medium. The results from our initial sessions provided some support for the equilibrium predictions. However, subjects deviated from equilibrium predictions in interesting ways. First, information acquisition decisions did not follow a cutpoint strategy. While extreme types were much more likely pay to learn Option identities, they did not always do so, and moderate types frequently became informed. Second, when abstention was permitted, uninformed subjects often cast votes, despite this being costly. When considering treatment effects, we found that increasing c decreased informedness, holding the voting scheme constant. Further, for $c = 3$, informedness was higher when voting was mandatory, as predicted. However, for $c = 9$, this finding did not reverse, contrary to equilibrium predictions.

In light of these observations, we hypothesized that an agent quantal response equilibrium (AQRE) was a better explanation for our data than the PBE predictions. An important implication of this model is that when abstention is permitted, uninformed voters will sometimes vote, reducing the probability of being a

[12] See Appendix B for a copy of the instructions. The instructions in this experiment draw on the instructions used in Großer and Palfrey (2019).

[13] See Appendix C for screenshots of this quiz.

pivotal voter, and thus reducing the value of information. This then "dilutes" the effect of increasing c. We fit an AQRE model to our initial data, and noted that for $c = 9$, the level of informedness was predicted to be slightly higher when voting was mandatory. This prediction is consistent with our data (and counter to the PBE predictions). We also noted that this model predicted that informedness would flip for higher levels of c.

Rather than merely note that an alternative model was a better fit for our data, we decided to test this particular implication of our preferred model explicitly. To determine whether sessions with a higher c would yield sufficient power, we performed a power analysis that accounted for the observed noisiness of behavior using an AQRE, following Woods (2023).[14] This power analysis relied on predictions from an AQRE model with estimates of λ obtained from our initial data. These predictions suggested that we would observe a more informed electorate with abstention when $c = 15$ and that running an additional 12 sessions with c varied between $c = 3$ and $c = 15$ would yield sufficient power to test this prediction.[15]

3.1 PBE Point Predictions

The PBE predictions for each of the six treatments are summarized in Table 1. In all treatments in which abstention is permitted, uninformed types are predicted not to vote. Behavior at the information acquisition stage is described by reporting the interior interval of types that are predicted not to become informed. Note that these intervals reflect that only extreme types are predicted to become informed, as they stand to lose more than moderate types when the "wrong" Option wins more votes. Further, note that for both voting schemes, the width of the interval of uninformed types is increasing in c. Table 1 also reports, for each treatment, the probability that a vote will be pivotal, and the ex-ante expected payoff of a participant in a given round. Note that probability of being pivotal is increasing in c when abstention is permitted, and is constant when voting is mandatory.

[14] See Appendix D and Appendix E for discussions of our AQRE results and AQRE power analysis, respectively.

[15] These additional sessions doubled the amount of data in which $c = 3$. While the power analysis did not call for this additional data, we wanted to ensure that sessions were as comparable as possible. Also of note, this power analysis found that our tests were vastly underpowered when $c = 9$. The number of observations needed to be sufficiently powered to test the predictions of the AQRE model was orders of magnitude higher.

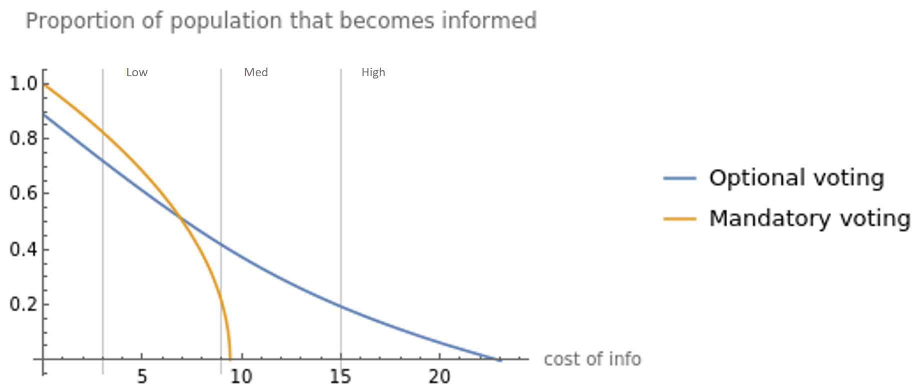
Table 1. PBE Predictions

	$[v_l^*, v_r^*]$	$P(informed)$	$P(pivotal)$	$E[\pi_{\text{round}}]$
AH	$[10, 90]$	0.20	0.70	68.69
AM	$[21, 79]$	0.42	0.53	69.53
AL	$[37, 63]$	0.725	0.41	71.90
MH	$[1, 100]$	0	0.38	66.56
MM	$[10, 90]$	0.20	0.38	66.64
ML	$[42, 58]$	0.84	0.38	70.84

Under PBE, informed types always vote for their preferred candidate. When abstention is possible, only informed types vote. $[v_l^*, v_r^*]$ is the interval of types predicted by PBE to purchase information. PBE predictions for informedness are empirically weighted based on types realized in the experiment.

Figure 2 summarizes the equilibrium level of informedness across voting schemes as c is varied. Importantly, when $c = 3$, mandatory voting is predicted to result in a strictly more informed electorate than when abstention is permitted. When $c = 9$, the cost of information is sufficiently high that this prediction reverses. Note that both voting schemes predict that some types will become informed when $c = 9$, although this set of types is larger when abstention is permitted. When $c = 15$, no types are predicted to become informed when voting is mandatory. However, when abstention is permitted, a relatively small set of extreme types are predicted to pay.

Figure 2. Predicted Information Acquisition



Information acquisition predictions for the parameterized experiment. $v \sim U[0,1]$, $a \sim U[0,.5]$ and $b \sim U[.5,1]$, with $\varepsilon = 2$ and $n = 5$.

4 Results

In each session, subjects face one cost of information in the first half of the experiment, and another cost in the second half. We refer to the 25 rounds in which the cost of information was held constant within a session as a block. In all non-parametric tests, we take the average of all decisions within a block as an independent observation. Thus, when $c = 3$, we have 12 observations in each voting scheme. For $c \in \{9, 15\}$, we have 6 observations per voting scheme. Since c is varied on a within-subject basis, one concern is that order effects may impede our ability to make inferences about the effect of changing c within a given voting scheme. To account for this, we varied the order in which costs were presented: half the time a session started with $c = 3$, and half the time a session started with the relevant higher cost. Note that this variation in order is perfectly balanced. In addition, ex-post, we ran regressions to assess whether the order of the variation in c affected behavior. We found no evidence that our results are driven by order effects. In what follows, we report p-values from two-tailed tests. Unless otherwise noted, we employ the Wilcoxon signed-rank test for comparisons with matched pairs, and the Mann-Whitney U test for comparisons without matched pairs.

4.1 Information Acquisition

Our primary outcome of interest is the frequency with which voters pay to learn Option identities. Table 2 contains summary statistics of this outcome for all treatments, alongside the associated PBE predictions. The predictions reported correspond to the predicted behavior for the actual types realized during the experiment, as opposed to ex-ante expectations. Interestingly, behavior is less responsive to changes in the cost of information than PBE predicts. In particular, when $c = 3$, we see more information acquisition than predicted both when voting is mandatory ($p = 0.002$), and when it is optional ($p = 0.002$). However, when moving to $c = 9$, we see that informedness falls substantially less than predicted. In fact, we see more informedness than PBE predicts when voting is mandatory ($p = 0.028$) and are unable to reject the null that the level of informedness conforms to PBE predictions when abstention is permitted ($p = 0.917$). When further increasing the cost of information to $c = 15$, we find that informedness exceeds PBE predictions both when voting is mandatory ($p = 0.028$) and when abstention is allowed ($p = 0.028$).

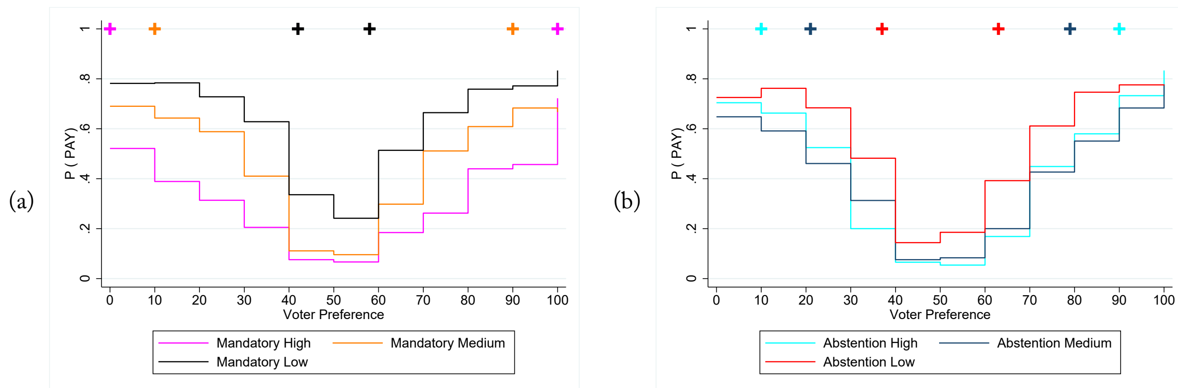
Table 2. Information Acquisition - Predicted and Observed

	$P_{\text{PBE}}(informed)$	$P_{\text{obs}}(informed)$
AH	0.20	0.42 (0.04)
AM	0.42	0.41 (0.09)
AL	0.725	0.55 (0.08)
MH	0	0.30 (0.07)
MM	0.20	0.47 (0.08)
ML	0.84	0.622 (0.06)

Means by treatment, with standard deviations in parentheses. PBE predictions are empirically weighted based on types realized in the experiment. Observed data are collapsed at the block level.

To better understand information acquisition decisions, we break them down by type. Figure 3 illustrates information acquisition choices by treatment, with types aggregated in bins of ten, as well as the relevant PBE predictions. Across all treatments, the qualitative prediction that more extreme types are more likely to become informed is borne out. However, as one might expect, subjects do not employ a cutpoint information acquisition strategy, and observed decisions are much noisier than predicted. In particular, moderate types are more informed than predicted while extreme types are less informed than predicted.

Figure 3. Predicted and Actual Information Acquisition



Data are averaged in blocks of ten by voter type. PBE cutpoint predictions are indicated by plus signs, where voter types outside the interval formed by the cutpoints are predicted to take the action indicated with probability one, and those inside the interval with probability zero.

The implications of these deviations from PBE predictions are interesting. Most significantly, while mandatory voting leads to a higher level of informedness when $c = 3$ ($p = 0.0243$), the predicted reversal of this finding when $c = 9$ is not observed. When $c = 9$, average information acquisition remains higher under mandatory voting, though the difference is not significant ($p = 0.522$). However, when c is further increased to $c = 15$, the predicted reversal emerges: informedness is higher when abstention is permitted ($p = 0.007$). As we will show below, these results can be explained by the irrational voting decisions of uninformed voters in the abstention treatments. Specifically, votes cast by uninformed types when abstention is possible reduces the probability of being the pivotal voter. Since the reason the value of information is increasing in c under abstention is an increased probability of being pivotal, these irrational votes reduce the effects of higher c.

Turning attention to the effects of changing the cost of information within a given voting scheme, we find that increases in c lead to decreases in informativeness when voting is mandatory. This is true considering an increase from $c = 3$ to $c = 9$ ($p = 0.028$), as well as when considering an increase from $c = 9$ to $c = 15$

($p = 0.004$). When abstention is permitted, we find a similar result when c increases from $c = 3$ to $c = 9$ ($p = 0.028$). However, when increasing c from $c = 9$ to $c = 15$, average informedness actually increases by about 1 percentage point, and there is no significant difference ($p = 0.873$). That is, when abstention is possible, subjects do not meaningfully respond to an increase in the price of information from $c = 9$ to $c = 15$. This is striking and is likely driven by the relatively high probability of being pivotal when c is relatively high. To see this, recall that when voting is mandatory the probability of being the pivotal voter does not depend on c and that we do see a drop in informedness when c increases from $c = 9$ to $c = 15$ in the absence of abstention.[16]

4.2 Voting Behavior

We next turn to voting decisions, since the choice to acquire information depends on beliefs regarding voting behavior. Table 3 contains summary statistics of voting behavior broken down by information acquisition choice. Recall that when abstention is permitted, PBE predicts that uninformed subjects will not vote, since doing so is costly and the two Options are indistinguishable. Further, since becoming informed is only worthwhile if a subject wishes to vote, informed subjects are always predicted to vote by PBE and to vote for the Option that, in expectation, is closest to their type.

Table 3. Summary Statistics of Observed Behavior at the Voting Stage

	$P(VC\|e_i = 1)$	$P(VW\|e_i = 1)$	$P(abs\|e_i = 1)$	$P(vote\|e_i = 0)$	$P(abs\|e_i = 0)$
AH	0.98	0.01	0.01	0.20	0.80
AM	0.96	0.03	0.01	0.18	0.82
AL	0.98	0.02	0.01	0.15	0.84
MH	0.95	0.05	0.00	1.00	0.00
MM	0.98	0.02	0.00	1.00	0.00
ML	0.96	0.04	0.00	1.00	0.00

VC indicates a vote for the "correct" Option, and VW a vote for the "wrong" Option. abs indicates a voter chose to abstain. e_i is an indicator variable that is equal to one if a subject paid to acquire information.

[16] Further research is needed to determine whether additional increases in c would yield reductions in informedness when abstention is permitted.

We observe relatively few instances in which informed voters vote for the Option that does not maximize their expected payoff (the "incorrect" Option). Similarly, informed voters seldom abstain from voting when permitted to do so. Interestingly, when abstention is possible, uninformed subjects vote fairly frequently: 15.1% of uninformed voters vote in AL, 17.6% in AM, and 20.0% in AH choose to vote.

Importantly, when uninformed subjects vote without being required to do so, the value of information under abstention is effectively diluted since uninformed votes reduce the probability of being the pivotal voter. Thus, a larger increase in c would be required to reverse the ranking of informedness by voting scheme. As discussed, this is consistent with our data in the information acquisition stage of the game.

Such "over-voting" by uninformed subjects has been observed in previous experimental studies in similar environments. In particular, Elbittar et al. (2020) and Großer and Seebauer (2016) study voting games in which voter preferences are aligned, and players are incentivized to vote in accordance with an unobserved state of the world. Before voting, subjects can purchase information that is correlated with the unobserved state. In both studies, voting by the uninformed is observed. Elbittar et al. (2020) suggest that such "over-voting" is driven by biased perceptions of the induced common prior belief regarding the state of the world. However, this explanation cannot explain our data, since such priors are not present in the games we study. Großer and Seebauer (2016) offer several other possible explanations. In particular, they argue that such "over-voting" could be explained by subject confusion regarding the negative effect such votes could have on the likelihood that the votes match the state. They also argue that uninformed voting may be driven by a desire to avoid negative reciprocity in which not voting leads to a reduction in information acquisition by other group members in future rounds (since information acquisition by any player can improve outcomes for all in their environment). Finally, they suggest that uninformed votes could be driven by perceived normative expectations regarding the social desirability of voting, such that some people have an aversion to not voting.

Only the last of these explanations could potentially explain "over-voting" in our experiment. However, if an individual level aversion to abstention drove behavior, we would expect "over-voting" would be due to a subset of subjects who consistently vote when uniformed. On the contrary, we find considerable heterogeneity: of those subjects who ever had to decide whether to vote when uninformed, 42.7% opted to vote at some point in the experiment.

Unfortunately, our experimental design does not allow us to conclude the precise driver of this "over-voting." However, the fact that such behavior is an apparently robust phenomenon across different environments makes it a promising avenue for future research. This is particularly true since there is no cost for casting a vote in both Elbittar et al. (2020) and Großer and Seebauer (2016), and we observe such voting when the cost of voting is strictly positive. Thus, we provide preliminary evidence that what Großer and Seebauer (2016) term "the curse of uninformed voting" is robust to a strictly positive cost of voting.

4.3 Agent Quantal Response Equilibrium

Our data differ substantially and systematically from the PBE predictions, which suggests that our data could be better explained by an alternative model. Such an alternative model should be able to account for three important features of our data. First, when abstention is possible, uninformed voters frequently vote, despite it being costly to do so. Second, more extreme types are more likely to become informed, but these information acquisition decisions are noisier than predicted by PBE. Third, when $c = 3$, informedness is higher when voting is mandatory, but the effect of increasing c on informedness is not as stark as predicted by PBE. In particular, an alternative model should account for the fact that the rate at which informedness declines in c is higher when voting is mandatory (so that for $c = 15$ allowing abstention results in more informedness). It should also account for the fact that these rates of decline are lower than predicted by PBE.

Quantal response equilibrium (QRE), which is frequently utilized to explain behavior in voting games, is just such an alternative model.[17] QRE generalizes Nash equilibrium by allowing for systematic errors in decision-making such that the likelihood of a particular error decreases in its costliness (McKelvey and Palfrey, 1995). Since our model involves a sequence of decisions, the appropriate variant is agent quantal response equilibrium (AQRE), which imposes a stochastic version of sequential rationality (McKelvey and Palfrey, 1998). AQRE assumes that errors at one information set are independent of the errors at another information set. In particular, at each of a player's information sets, it is assumed that a distinct "agent" noisily best responds to the player's equilibrium beliefs, which account for errors on the part of other players and by her "agents" at other information sets.

[17] Examples of voting games employing quantal response equilibrium include Großer and Palfrey (2019); McKelvey and Patty (2006); Tyszler and Schram (2016); Bhattacharya et al. (2017); Bassi (2015); Levine and Palfrey (2007); Elbittar et al. (2020); Goeree and Holt (2005).

For simplicity, and in line with the existing literature, we consider AQRE such that decisions are made using a logit response function in which a single parameter λ determines the level of noise present in a decision. Specifically, we assume that when type v_i must make a decision with possible actions $A \in \{a_1, a_2 \ldots, a_m\}$, the probability that action $a_j \in A$ is chosen is given by

$$Prob(a_j) = \frac{e^{\lambda E(\pi_i|a_j)}}{\sum_{k=1}^{m} e^{\lambda E(\pi_i|a_k)}},$$

where $E(\pi_i|a_k)$ denotes the profit of type v_i conditional on choosing action a_k. Note that if $\lambda = 0$, then all possible actions are equally likely, and as $\lambda \to \infty$, the action which maximizes expected profit is chosen with certainty. To account for the fact that the complexity of choices, strategic or otherwise, differs within our extensive form game, we consider AQRE in which λ varies across decision stages.[18] In addition, we consider AQRE in which λ is permitted to vary across treatments.

We estimate several AQRE models via maximum likelihood. Ex-ante, we committed to using the Bayesian information criterion (BIC) for model selection to ensure we are not overfitting our model. We consider a model with a single λ for all treatments and decision stages, a model in which λ varies only by voting scheme, a model in which λ varies by decision stage and voting scheme, and a model in which λ varies by treatment and decision stage. Note that we do not consider cases in which λ varies by type, as this would result in an extremely large number of free parameters.

We follow the maximum likelihood procedure described in Großer and Palfrey (2019).[19] The procedure is complicated because for a given vector of λ's, we must solve for the corresponding AQRE to evaluate the likelihood function.[20] Thus, we perform a grid search, in which we finely discretize the set of possible λ's, calculate the associated AQRE for each of these possibilities, and then evaluate the likelihood function for each of these possibilities to select the optimal value(s) for λ. Of the models, BIC is minimized when we allow λ to vary both by treatment and by decision stage. In the interest of brevity, we relegate further

[18] This difference in complexity is most obvious when comparing (a) the likelihood of an informed voter voting for the wrong candidate to (b) the likelihood of becoming informed at all. The expected marginal benefit of voting for the wrong candidate is relatively easy to assess, whereas calculating the expected marginal benefit of becoming informed requires much more sophistication. We are not the first to assume that λ may differ across decision stages. See, for example, Großer and Palfrey (2019).

[19] We thank Jens Großer for generously providing details and the associated code for this procedure. Our implementation is done in Mathematica and R, and the associated code is available on request.

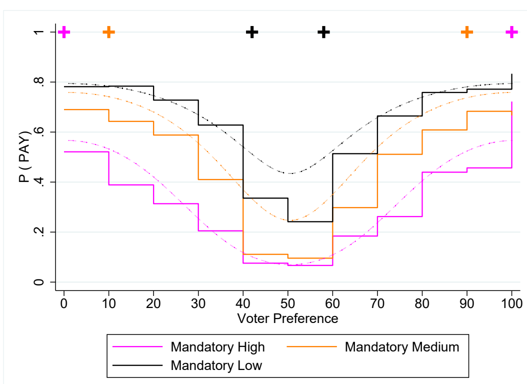
[20] Since there are 100 possible types in the game employed in our experiment, we must, for a given vector of λ's, solve for the probabilities of each possible action at each possible decision for each possible type. Focusing on a symmetric AQRE, and assuming symmetry of behavior for equally extreme types, this leaves us with a system of 100 equations when voting is mandatory and a system of 250 equations when abstention is permitted.

details of both computing the AQRE predictions and the maximum likelihood procedure for selecting parameters to Appendix D. Below, we discuss the AQRE predictions for both the information acquisition decision and the voting decision.

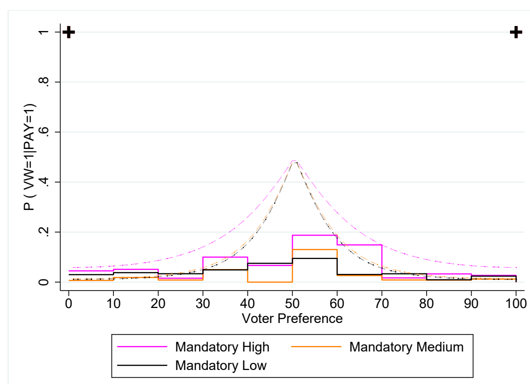
We first consider the effect of increasing the cost of information, holding the voting scheme constant. Figure 4 illustrates the observed decisions at each stage of the game when voting is mandatory, as well as the corresponding PBE and AQRE predictions, by voter type. In these plots, the AQRE predictions provide a much closer fit to our data than the PBE predictions, while also being consistent with the observed decrease in informedness as c increases.[21] It is important to note that the AQRE predictions regarding information acquisition fit the data best for extreme voter types. This is because the cost of making an error for a moderate type is much lower than for an extreme type, and the maximum likelihood procedure thus prefers parameters that fit these extremes.[22]

Figure 4. All Decisions - Mandatory

(a)



(b)



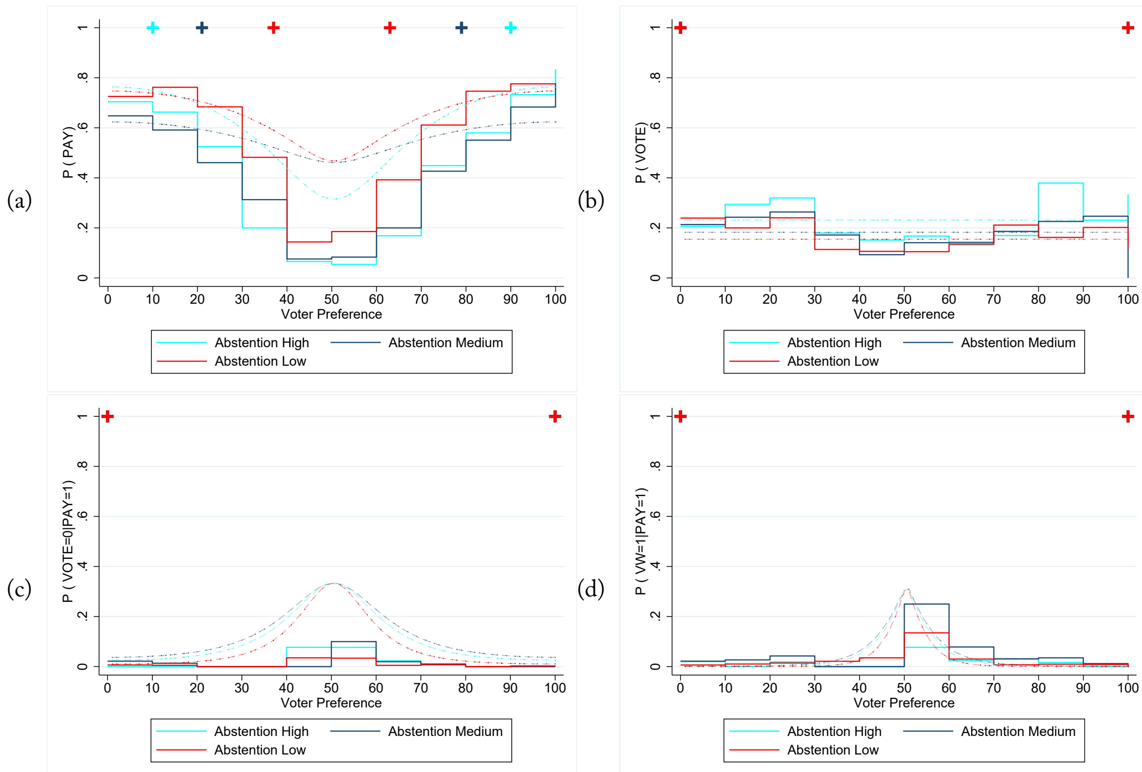
Data are averaged in blocks of ten by voter type. PBE cutpoint predictions are indicated by plus signs, where voter types outside the interval formed by the cutpoints are predicted to take the action indicated with probability one, and those inside the interval with probability zero. AQRE predictions are displayed as dotted lines. VW indicates a vote for the "wrong" Option.

Figure 5 illustrates observed decisions at each stage of the game when abstention is allowed, as well as the associated AQRE and PBE predictions. Note that the possibility of abstention means there are more potential errors to account for. In particular, informed voters can choose to abstain or to vote for the "incorrect" Option. Additionally, uninformed voters can choose not to abstain, thus incurring a cost of ε=2. AQRE can account for each of these errors. It is especially noteworthy that this AQRE model explains the

[21]Of course, the AQRE selection process nests PBE predictions, so AQRE must provide at least as good a fit as PBE.
[22]This could be corrected by allowing λ to vary by voter type. However, we felt this would introduce too many free parameters.

relatively high level of uninformed voting we observe and that such voting does not depend on voter type. Further, AQRE predictions regarding information acquisition fit our data well, especially for extreme types.

Figure 5. All Decisions - Abstention

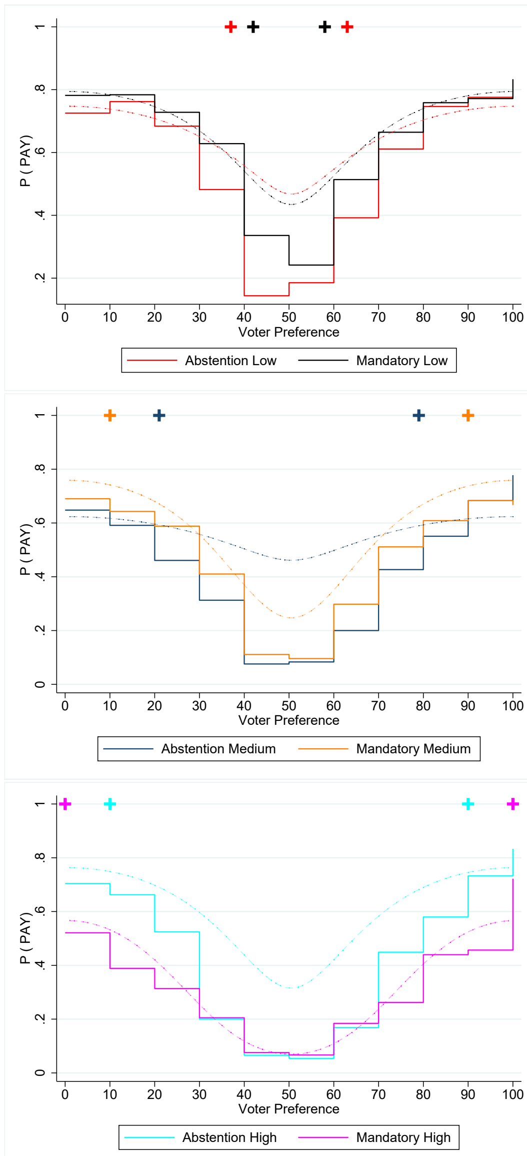


Data are averaged in blocks of ten by voter type. PBE cutpoint predictions are indicated by plus signs, where voter types outside the interval formed by the cutpoints are predicted to take the action indicated with probability one, and those inside the interval with probability zero. AQRE predictions are displayed as dotted lines. VW indicates a vote for the "wrong" Option.

We next discuss our primary effect of interest: holding c constant, how does varying the voting scheme affect behavior? This is particularly of interest since we chose to run the sessions with $c = 15$ *after* observing data with $c = 3$ and $c = 9$. Upon collecting this initial data, we found that AQRE provides a better explanation of behavior than PBE. We then noted that an AQRE (with λ's estimated from the initial data) *predicted* that informedness would be higher under abstention when $c = 15$. Thus, the second wave of data collection involves a test of an ex-ante hypothesis, which was formulated on the basis of an updated theory. Figure 6 illustrates observed behavior as well as the corresponding AQRE and PBE predictions.

Note that the AQRE predictions are consistent with the observed behavior on the tails of the distribution of voter types. Further, notice that the AQRE *predicts* that allowing abstention will increase informedness when $c = 15$, which is consistent with our findings as discussed in section 4.1. For lower levels of c, AQRE is consistent with our observation that informedness is higher when voting is mandatory.

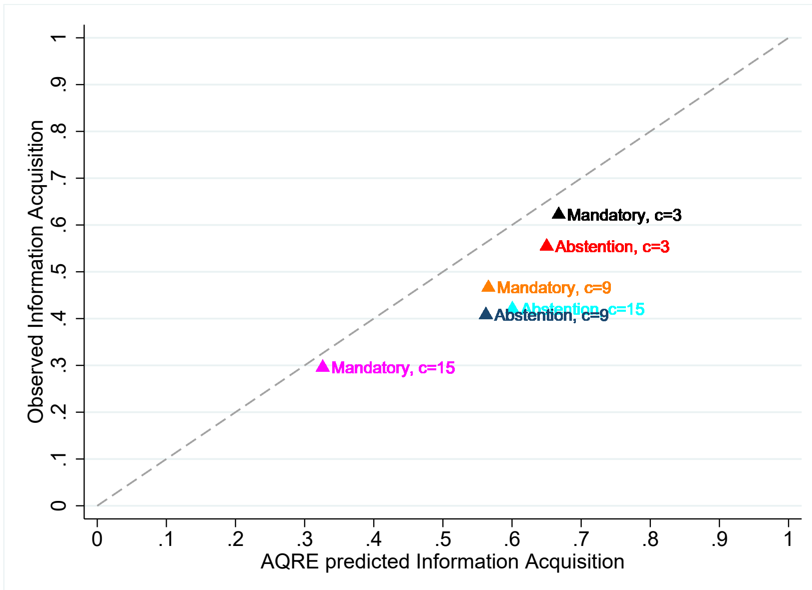
Figure 6. Information Acquisition with and without Abstention



Data are averaged in blocks of ten by voter type. PBE cutpoint predictions are indicated by plus signs, where voter types outside the interval formed by the cutpoints are predicted to take the action indicated with probability one, and those inside the interval with probability zero. AQRE predictions are displayed as dotted lines.

Overall, the AQRE model provides a much better fit for our data than the PBE predictions. This is further illustrated in Figure 7, which plots observed informedness by treatment against the AQRE predictions. Not only is AQRE able to account for the observed uninformed voting when abstention is possible, it accounts for the observed patterns of information acquisition. [23]

Figure 7. Information Acquisition - Predicted and Observed



Average information acquisition for each treatment, against AQRE predictions.

5 Conclusion

This paper investigates how mandatory voting laws affect information acquisition in nonpartisan elections. In our model, voters have single-peaked preferences over policy outcomes, which are independent draws from a common-knowledge distribution (types are private information). There are two candidates in the election, each associated with different expected policy outcomes. The outcome associated with one candidate is on the left half of the support of voter types, while the other candidate is associated with the

[23]It is important to note that we are unable to definitively conclude that AQRE is the appropriate model of behavior in these games. Rather, we are only able to state that the AQRE predictions provide a better explanation of behavior than PBE. However, we do feel that it is important to consider our data in the context of the existing experimental literature, which has often noted that QRE/AQRE is a good potential model to explain behavior in voting games. See, for example, Großer and Palfrey (2019); McKelvey and Patty (2006); Tyszler and Schram (2016); Bhattacharya et al. (2017); Bassi (2015); Levine and Palfrey (2007); Elbittar et al. (2020); Goeree and Holt (2005). We view our data as a single observation in this larger literature, and believe that the case for QRE/AQRE models in voting games is compelling.

right half of this support. Candidates are indistinguishable to voters; each voter may learn candidate identities by incurring a private cost. After the information acquisition stage, costly voting takes place. We consider voting environments with and without abstention.

Our model predicts that mandatory voting laws will increase citizen informedness when information is cheap relative to the cost of voting and decrease citizen informedness when information is relatively expensive. The primary theoretical mechanism that predicts higher informedness when abstention is allowed is the increased probability of being the pivotal voter, since in equilibrium uninformed voters abstain when permitted to do so.

To evaluate these predictions we conducted a laboratory experiment. We varied whether or not abstention was permitted on a between-subject basis and the cost of information on a within-subject basis. We initially ran experiments for only two costs of information and found that when the cost was low, that informedness was higher when voting was mandatory. However, when the cost of information was increased, we found that uninformed subjects often voted when not required to do so. This reduced the value of information under abstention, since the probability of being the pivotal voter is reduced. As a result, informedness was not significantly different across voting schemes.

These results are well explained using agent quantal response equilibrium, which accounts for the observed errors that voters make in these games. Most importantly, AQRE predicts (in line with our initial data) that uninformed voters will sometimes vote, even when abstention is permitted. AQRE also predicts that a further increase in the cost of information would result in a higher level of informedness under abstention. To assess whether our preferred model has predictive power, we ran a second set of experiments to test this hypothesis. These experiments were identical to the first series, except that we evaluated a higher cost of information. The results are in line with the AQRE predictions, which increases our confidence in our conclusion that AQRE is an appropriate prediction tool in voting games such as ours.[24]

Nonpartisan elections play a key role in determining public policy, especially at a local level. Often voters are unable to distinguish between candidates and there is no partisan cue to guide their decision. This leads to reduced election participation, which may reduce the effective accountability of these officials to voters. As such, finding ways to increase the informedness of voters in these elections is an important issue. Our results suggest that making voting mandatory in such elections may be an effective mechanism to increase

[24]Note that our data is far from conclusive and represents a single observation in the literature evaluating AQRE/QRE models in voting games. However, we believe that the evidence presented in this literature is compelling.

informedness, even when the cost of becoming informed is relatively high, since the effect of increasing the cost of information is not as stark as predicted. Indeed, the cost of information needs to be quite high before making voting mandatory would reduce the level of informedness.

Variation in observed voting schemes is relatively rare. Thus, experimental and theoretical advances can help guide policy experiments to study the effects of mandatory voting in real-world environments. This paper is a first step in this direction. Moving forward, a particularly fruitful avenue for future research would be experiments studying the effect of differences in the cost of voting itself, or allowing candidates to endogenously select platforms in an environment in which these platforms are only observed at a cost. Future work could build on this model by accounting for voter preferences for candidate quality, in addition to partisanship.

References

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

Battaglini, M., R. B. Morton, and T. R. Palfrey (2008, May). Information aggregation and strategic abstention in large laboratory elections. *American Economic Review 98*(2), 194–200.

Bhattacharya, S., J. Duffy, and S. Kim (2017). Voting with endogenous information acquisition: Experimental evidence. *Games and Economic Behavior 102*, 316–338.

Bonneau, C. and D. Cann (2015). Party identification and vote choice in partisan and nonpartisan elections. *Political Behavior 37*(1), 43–66.

Börgers, T. (2004). Costly voting. *American Economic Review 94*(1), 57–66.

Bowler, S., T. Donovan, and T. Happ (1992). Ballot propositions and information costs: Direct democracy and the fatigued voter. *Western Political Quarterly 45*(2), 559–568.

Elbittar, A., A. Gomberg, C. Martinelli, and T. R. Palfrey (2020). Ignorance and bias in collective decisions. *Journal of Economic Behavior & Organization 174*, 332 – 359.

Feddersen, T. J. and W. Pesendorfer (1999). Abstention in elections with asymmetric information and diverse preferences. *American Political Science Review 93*(2), 381–398.

Fischbacher, U. (2007). z-tree: Zurich toolbox for ready-made economic experiments. *Experimental economics 10*(2), 171–178.

Gersbach, H. (1992). Allocation of information by majority decisions. *Journal of Public Economics 48*(2), 259 – 268.

Ghirardato, P. and J. N. Katz (2006). Indecision theory: Weight of evidence and voting behavior. *Journal of Public Economic Theory 8*(3), 379–399.

Ghosal, S. and B. Lockwood (2009, Jun). Costly voting when both information and preferences differ: is turnout too high or too low? *Social Choice and Welfare 33*(1), 25–50.

Goeree, J. K. and C. A. Holt (2005). An explanation of anomalous behavior in models of political participation. *American political science Review*, 201–213.

Großer, J. and T. Palfrey (2014). Candidate entry and political polarization: An antimedian voter theorem. *American Journal of Political Science 58*(1), 127–143.

Großer, J. and T. R. Palfrey (2019). Candidate entry and political polarization: An experimental study. *American Political Science Review 113*(1), 209–225.

Großer, J. and M. Seebauer (2016). The curse of uninformed voting: An experimental study. *Games and Economic Behavior 97*, 205 – 226.

Hoffman, M., G. León, and M. Lombardi (2017). Compulsory voting, turnout, and government spending: Evidence from austria. *Journal of Public Economics 145*, 103 – 115.

Krasa, S. and M. K. Polborn (2009). Is mandatory voting better than voluntary voting? *Games and Economic Behavior 66*(1), 275–291.

Krishna, V. and J. Morgan (2012). Voluntary voting: Costs and benefits. *Journal of Economic Theory 147*(6), 2083 – 2123.

Lassen, D. D. (2005). The effect of information on voter turnout: Evidence from a natural experiment. *American Journal of Political Science 49*(1), 103–118.

Levine, D. K. and T. R. Palfrey (2007). The paradox of voter participation? a laboratory study. *American political science Review 101*(1), 143–158.

Lim, C. S. and J. M. Snyder Jr (2015). Is more information always better? party cues and candidate quality in us judicial elections. *Journal of public Economics 128*, 107–123.

Martinelli, C. (2006). Would rational voters acquire costly information? *Journal of Economic Theory 129*(1), 225–251.

Martinelli, C. (2007). Rational ignorance and voting behavior. *International Journal of Game Theory 35*(3), 315–335.

Matějka, F. and G. Tabellini (2017, September). Electoral competition with rationally inattentive voters.

McKelvey, R. and T. Palfrey (1995). Quantal response equilibria for normal form games. *Games and economic behavior 10*(1), 6–38.

McKelvey, R. and T. Palfrey (1998). Quantal response equilibria for extensive form games. *Experimental economics 1*(1), 9–41.

McKelvey, R. and J. Patty (2006). A theory of voting in large elections. *Games and Economic Behavior 57*(1), 155–180.

Morton, R. B., M. Piovesan, and J.-R. Tyran (2019). The dark side of the vote: Biased voters, social information, and information aggregation through majority voting. *Games and Economic Behavior 113*, 461–481.

Oliveros, S. (2013). Abstention, ideology and information acquisition. *Journal of Economic Theory 148*(3), 871 – 902.

Ortoleva, P. and E. Snowberg (2015, February). Overconfidence in political behavior. *American Economic Review 105*(2), 504–35.

Rahn, W. (1993). The role of partisan stereotypes in information processing about political candidates. *American Journal of Political Science 37*, 472–496.

Roth, A. E. (1994). Lets keep the con out of experimental a methodological note. *Empirical Economics 19*(2), 279–289.

Schaffner, B. and M. Streb (2002). The partisan heuristic in low-information elections. *Public Opinion Quarterly 66*(4), 559–581.

Schaffner, B., M. Streb, and G. Wright (2001). Tearns without uniforms: The nonpartisan ballot in state and local elections. *Political Research Quarterly 54*(1), 7–30.

Strömberg, D. (2004). Radio's impact on public spending. *The Quarterly Journal of Economics 119*(1), 189–221.

Tyszler, M. and A. Schram (2016). Information and strategic voting. *Experimental economics 19*(2), 360–381.

Woods, D. (2023). Improving ex-ante power analysis with quantal response simulations. *Working paper*.

A Proofs

A.1 Proof of Proposition 1

Each v_i's candidate preference is determined by

$$\begin{cases} E\left[\pi_i|v_i, \gamma \in \left[0, \tfrac{1}{2}\right)\right] - E\left[\pi_i|v_i, \gamma \in \left(\tfrac{1}{2}, 1\right]\right] > 0 & v_i \text{ prefers candidate } A \\ E\left[\pi_i|v_i, \gamma \in \left[0, \tfrac{1}{2}\right)\right] - E\left[\pi_i|v_i, \gamma \in \left(\tfrac{1}{2}, 1\right]\right] < 0 & v_i \text{ prefers candidate } B \\ E\left[\pi_i|v_i, \gamma \in \left[0, \tfrac{1}{2}\right)\right] - E\left[\pi_i|v_i, \gamma \in \left(\tfrac{1}{2}, 1\right]\right] = 0 & v_i \text{ indifferent to election outcome.} \end{cases}$$

For ease of reference, we denote voters who prefer candidate A as v_i^A, and candidate B as v_i^B.

Under mandatory voting and recalling equation (1), equation (2) for a voter v_i^A can be written as

$$E\big[U\left(|v_i^A - \gamma|\right)|v_i^A, e_i = 0\big] - E\big[U\left(|v_i^A - \gamma|\right)|v_i^A, e_i = 1\big] \geq c. \tag{8}$$

Rewritten and rearranging,

$$\begin{aligned} &\int_0^{\frac{1}{2}} U\big(|v_i^A - x|\big)a(x)\,dx \cdot \Big(P\big(\gamma \in \big[0, \frac{1}{2}\big)|e_i = 0\big) - P\big(\gamma \in \big[0, \frac{1}{2}\big)|e_i = 1\big)\Big) + \\ &\int_{\frac{1}{2}}^1 U\big(|v_i^A - x|\big)b(x)\,dx \cdot \Big(P\big(\gamma \in \big(\frac{1}{2}, 1\big]|e_i = 0\big) - P\big(\gamma \in \big(\frac{1}{2}, 1\big]|e_i = 1\big)\Big) \geq c \end{aligned} \tag{9}$$

The probability of each candidate's election depends on the information acquisition choice of voter v_i^A and the expected votes cast by all other voters v_{-i}. Suppose all voters v_{-i} follow a strategy that involves information acquisition for all types $v_{-i} \leq v_l^m$ or $v_{-i} \geq v_r^m$, for some $(v_l^m, v_r^m) \in [0, 1) \times (0, 1]$, with $v_l^m \leq v_r^m$. Types that acquire information vote for their preferred candidate, while the remaining types randomize their vote.

For simplicity, we restrict attention to symmetric distributions of v, and distributions of a and b such that $\left|\frac{1}{2} - E(\gamma_A)\right| = \left|E(\gamma_B) - \frac{1}{2}\right|$. This implies that $v_r^m = 1 - v_l^m$, and player v_i^A expects each other player to vote for candidate A with probability $\frac{1}{2}$. The probabilities in equation (9) are then straightforward to calculate, and (9) can be rewritten

$$\left(\int_{\frac{1}{2}}^{1} U\big(|v_i^A - x|\big)b(x)\ dx - \int_0^{\frac{1}{2}} U\big(|v_i^A - x|\big)a(x)\ dx\right)\left(\frac{1}{2} - \sum_{k=0}^{\frac{n-3}{2}} \binom{n-1}{k}\frac{1}{2}^{n-1}\right) \geq c$$

or equivalently

$$\frac{1}{2}\left(\int_{\frac{1}{2}}^{1} U\big(|v_i^A - x|\big)b(x)\ dx - \int_0^{\frac{1}{2}} U\big(|v_i^A - x|\big)a(x)\ dx\right)\binom{n-1}{\frac{n-1}{2}}\frac{1}{2}^{n-1} \geq c. \tag{10}$$

This is simply voter v_i^A's expected marginal benefit if her vote is pivotal (i.e., voter i's vote is a tiebreaker, resulting in her preferred candidate being chosen), weighted by the probability of being the pivotal voter.[25] For ease of notation, we denote the left-hand side of equation(10) as $F(v_i^A)$.

To find v_l^m, note that if this cutpoint exists, it is where voter $v_i^A = v_l^m$ is indifferent to purchasing information for price c. Note that $F(0) > 0$ due to the concavity of U, and $F(\frac{1}{2}) = 0$. Furthermore, via Leibniz's rule, $\frac{\partial F}{\partial v_l^m}$ is monotonically decreasing over $(0, \frac{1}{2})$:

$$\frac{\partial F}{\partial v_l^m} = \underbrace{\frac{1}{2}\left(\int_{v_l^m}^{\frac{1}{2}} \frac{\partial U}{\partial |v_l^m - x|} a(x)\, dx - \int_{\frac{1}{2}}^{1} \frac{\partial U}{\partial |v_l^m - x|} b(x)\, dx - \int_0^{v_l^m} \frac{\partial U}{\partial |v_l^m - x|} a(x)\, dx\right)}_{<0\ \forall\, v_l^m \in (0, \frac{1}{2})} \times \underbrace{\binom{n-1}{\frac{n-1}{2}}\frac{1}{2}^{n-1}}_{\in(0,1)\ \forall\, n \in \mathbb{Z}^+}. \tag{11}$$

The intermediate value theorem establishes the existence and uniqueness of an equilibrium threshold $v_i^A = v_l^m$, where $F(v_l^m)$ = c, for $c \in \big[0, \tilde{c}^m \equiv F(0)\big]$. Any voter in $[0, v_l^m]$ will pay to acquire information, and any voter in $(v_l^m, \frac{1}{2}]$ will not. An analogous proof for v_i^B yields a unique solution for v_r^m, where any voter in $[v_r^m, 1]$ will pay to acquire information, and any voter in $[\frac{1}{2}, v_r^m)$ will not.

[25] One could arrive at this result more directly by starting with equation (8) and immediately assuming symmetry of distributions. This is the course taken in the proof of Proposition 2. By delaying the assumption of symmetry, this proof remains more applicable to non-symmetric distributions.

A.2 Proof of Proposition 2

When abstention is permitted, voters choose whether or not to vote, in addition to choosing whether or not to acquire information about candidates' party affiliation. When $\varepsilon > 0$, no voter has an incentive to be uninformed and to vote, since

$$E[\pi_i|v_i, e_i = 0, w_i = 0] - E[\pi_i|v_i, e_i = 0, w_i = 1] = \varepsilon.$$

Similarly, for $c \geq 0$, no voter has an incentive to become informed and then not vote, since

$$E[\pi_i|v_i, e_i = 0, w_i = 0] - E[\pi_i|v_i, e_i = 1, w_i = 0] = c.$$

Thus, allowing for abstention, equation (2) becomes

$$E[U|v_i, e_i = w_i = 0] - E[U|v_i, e_i = w_i = 1] \geq c + \varepsilon. \tag{12}$$

Suppose all voters v_{-i} follow a strategy that involves information acquisition for all types $v_{-i} \leq v_l^a$ or $v_{-i} \geq v_r^a = 1 - v_l^a$, for some $v_l^a \in \left[0, \frac{1}{2}\right]$.[26] Types that acquire information vote for their preferred candidate, while the remaining types do not vote. If all voters v_{-i} follow this strategy, the expected utility for voter v_i^A from getting informed and voting is

$$\begin{aligned} E[U|v_i, e_i = w_i = 1] = E[U|v_i^A, e_i = w_i = 0] \; - \\ \frac{1}{2}\bigg(\int_{\frac{1}{2}}^{1} U\big(|v_i^A - x|\big) b(x)\; dx - \int_{0}^{\frac{1}{2}} U\big(|v_i^A - x|\big) a(x)\; dx\bigg) \times \\ \Bigg(\sum_{t=0}^{\frac{n-1}{2}} \binom{n-1}{2t}\binom{2t}{t}\big(1 - 2V(v_l^a)\big)^{n-1-2t} V(v_l^a)^{2t} + \\ \sum_{t=0}^{\frac{n-3}{2}} \binom{n-1}{2t+1}\binom{2t+1}{t+1}\big(1 - 2V(v_l^a)\big)^{n-2-2t} V(v_l^a)^{2t+1}\Bigg). \end{aligned} \tag{13}$$

[26]This relies on the symmetry of v and distributions of a and b such that $\left|\frac{1}{2} - E(\gamma_A)\right| = \left|E(\gamma_B) - \frac{1}{2}\right|$.

The last term in equation (13) reflects the expected marginal benefit of becoming informed and voting accordingly. The first part of the term is the expected benefit if voter i's vote is pivotal, and the last term in parentheses represents the probability of being pivotal. Note that with abstention, "pivotal" includes not only the case of being the tie-breaking voter, but also of being the tie-making voter. The expected marginal benefit is in fact the same for both cases. The probability of being pivotal varies across the cases, as evidenced by the two distinct summation terms.

Inserting equation (13) into (12), and making use of a property of binomials yields a slightly tidier expression

$$\begin{aligned}&\frac{1}{2}\bigg(\int_{\frac{1}{2}}^{1} U\big(|v_i^A - x|\big)b(x)\ dx - \int_{0}^{\frac{1}{2}} U\big(|v_i^A - x|\big)a(x)\ dx\bigg)\times \\ &\sum_{t=0}^{\frac{n-1}{2}}\sum_{s=0}^{1}\binom{n-1}{2t+s}\binom{2t+s}{t+s}\big(1-2V(v_l^a)\big)^{n-1-s-2t}V(v_l^a)^{2t+s} \quad \geq c+\varepsilon\end{aligned} \tag{14}$$

Similar to the mandatory case, this has the simple interpretation that voter v_i^A will choose to become informed if the expected marginal benefit of being pivotal, weighted by the probability of being such, exceeds the cost. For ease of reference, we denote the left-hand side of (14) as $G(v_i^A, v_l^a)$.

To find v_l^a, note that if this cutpoint exists, it is where voter $v_i^A = v_l^a$ is indifferent to purchasing information for price c, and incurring the cost ε to vote. Note that when $v_l^a = 0$, $G(v_l^a, v_l^a) > 0$ due to the concavity of U. When $v_l^a = \frac{1}{2}$, $G(v_l^a, v_l^a) = 0$. Furthermore, via Leibniz's rule, $\frac{\partial G(v_l^a, v_l^a)}{\partial v_l^a} < 0$:

$$\frac{\partial G}{\partial v_l^a} = \underbrace{\frac{1}{2}\left(\int_{v_l^a}^{\frac{1}{2}} \frac{\partial U}{\partial |v_l^a - x|} a(x)\,dx - \int_{\frac{1}{2}}^{1} \frac{\partial U}{\partial |v_l^a - x|} b(x)\,dx - \int_{0}^{v_l^a} \frac{\partial U}{\partial |v_l^a - x|} a(x)\,dx\right)}_{<0\;\forall\; v_l \in (0,\frac{1}{2})} \times$$
$$\underbrace{\sum_{t=0}^{\frac{n-1}{2}} \sum_{s=0}^{1} \binom{n-1}{2t+s}\binom{2t+s}{t+s}\big(1-2V(v_l^a)\big)^{n-1-s-2t} V(v_l^a)^{2t+s}}_{\in (0,1)\;\forall\; V(v_l)\in(0,\frac{1}{2}), n\in\mathbb{Z}^+} +$$
$$\underbrace{v(v_l^a)^n \frac{1}{2}\left(\int_{\frac{1}{2}}^{1} U\big(|v_l^a - x|\big) b(x)\,dx - \int_{0}^{\frac{1}{2}} U\big(|v_l^a - x|\big) a(x)\,dx\right)}_{>0\;\forall\; v_l\in(0,\frac{1}{2})} \times$$
$$\underbrace{\sum_{t=0}^{\frac{n-1}{2}} \sum_{s=0}^{1} \binom{n-1}{2t+s}\binom{2t+s}{t+s}\big(1-2V(v_l^a)\big)^{n-2-s-2t} V(v_l^a)^{2t+s-1}\big((2t+s) - 2V(v_l^a)(n-1)\big)}_{<0\;\forall\; V(v_l^a)\in(0,\frac{1}{2}), n\in\mathbb{Z}^+} \tag{15}$$

Recall that we assume that the probability of being a pivotal voter increases as less voter types participate in the election. This assumption allows the last bracketed term of 15, and consequently the sign of the entire expression, to be signed as strictly negative. The intermediate value theorem establishes the existence and uniqueness of a cutpoint $v_i^A = v_l^a$, where $G(v_l^a, v_l^a) = c + \varepsilon$, for $c + \varepsilon \in [0, \tilde{c}^a \equiv \lim_{v_i^A \to 0^+} G(v_i^A, v_i^A)]$.

To confirm that voters in $[0, v_l^a]$ will follow the cutpoint strategy in (6), note that for any given $v_l^a \in [0, \frac{1}{2})$, $G(0, v_l^a)$ is strictly positive, $G(\frac{1}{2}, v_l^a) = 0$, and $\frac{\partial G}{\partial v_i^A} < 0$:

$$\frac{\partial G}{\partial v_i} = \underbrace{\frac{1}{2}\left(\int_{v_i}^{\frac{1}{2}} \frac{\partial U}{\partial |v_i - x|} a(x)\,dx - \int_{\frac{1}{2}}^{1} \frac{\partial U}{\partial |v_i - x|} b(x)\,dx - \int_{0}^{v_i} \frac{\partial U}{\partial |v_i - x|} a(x)\,dx\right)}_{<0\;\forall\; v_i\in(0,\frac{1}{2})} \times$$
$$\underbrace{\sum_{t=0}^{\frac{n-1}{2}} \sum_{s=0}^{1} \binom{n-1}{2t+s}\binom{2t+s}{t+s}\big(1-2V(v_l^a)\big)^{n-1-s-2t} V(v_l^a)^{2t+s}}_{\in(0,1)\;\forall\; V(v_l^a)\in(0,\frac{1}{2}), n\in\mathbb{Z}^+}. \tag{16}$$

Thus, voters have no incentive to deviate from the cutpoint strategy, and all voters in $[0, v_l^a]$ will pay to acquire information, and those in $\left(v_l^a, \frac{1}{2}\right]$ will not. A symmetric procedure yields v_r^a, where voters in $[v_r^a, 1]$ will pay to acquire information, and those in $\left[\frac{1}{2}, v_r^a\right)$ will not.

A.3 Proof of Lemma 1

Under mandatory and optional voting, $v_l^m(c)$ and $v_l^a(c, \varepsilon)$ are implicitly defined by equations (10) and (14), yielding:

$$\frac{1}{2}\left(\int_{\frac{1}{2}}^{1} U\big(|v_l^m - x|\big)b(x)\,dx - \int_0^{\frac{1}{2}} U\big(|v_l^m - x|\big)a(x)\,dx\right)\binom{n-1}{\frac{n-1}{2}}\frac{1}{2}^{n-1} - c \;=\; 0 \qquad (10^*)$$

$$\begin{aligned}&\frac{1}{2}\left(\int_{\frac{1}{2}}^{1} U\big(|v_l^a - x|\big)b(x)\,dx - \int_0^{\frac{1}{2}} U\big(|v_l^a - x|\big)a(x)\,dx\right)\times \\ &\sum_{t=0}^{\frac{n-1}{2}}\sum_{s=0}^{1}\binom{n-1}{2t+s}\binom{2t+s}{t+s}\big(1-2V(v_l^a)\big)^{n-1-s-2t}V(v_l^a)^{2t+s} - c - \varepsilon \end{aligned} = 0 \qquad (14^*)$$

From (10*) and (14*), $v_l^m(0) = v_l^a(0,0) = \frac{1}{2}$. Furthermore, $v_l^m(\tilde{c}^m) = v_l^a(\tilde{c}^a, 0) = 0$. We now turn attention to $\frac{dv_l^m}{dc}$ and $\frac{dv_l^a}{dc}$. Total differentiation of equations (10*) and (14*) w.r.t. c yields

$$\frac{dv_l^m}{dc} = \frac{1}{(11)} \qquad \text{and} \qquad \frac{dv_l^a}{dc} = \frac{1}{(15),}$$

which can be readily compared to yield $\frac{dv_l^m}{dc} < \frac{dv_l^a}{dc} < 0$. For $\varepsilon = 0$ this implies

$$v_l^m < v_l^a \;\forall\, c \in \Big(0, max\big\{\tilde{c}^m, \tilde{c}^a\big\} = \tilde{c}^a\Big)$$

and

$$v_l^m = v_l^a = 0 \;\forall\, c \geq \tilde{c}^a.$$

These results, together with an analogous proof for the right cutpoints v_r^m and v_r^a imply that for $\varepsilon = 0$, $[v_l^a, v_r^a] \subseteq [v_l^m, v_r^m]\;\; \forall\, c \geq 0$.

A.4 Proof of Lemma 2

The introduction of a cost of voting $\varepsilon > \tilde{c}^a - \tilde{c}^m$ shifts $v_l^a(c, \varepsilon)$ by ε such that $v_l^a(0) < v_l^m(0)$, and $\tilde{c}^a - \varepsilon < \tilde{c}^m$. And, from the proof of **Lemma 1**, $\frac{dv_l^m}{dc} < \frac{dv_l^a}{dc} < 0$. This implies

$$v_l^a < v_l^m \; \forall \, c \in \Big(0, max\big\{\tilde{c}^m, \tilde{c}^a - \varepsilon\big\} = \tilde{c}^m\Big)$$

and

$$v_l^a = v_l^m = 0 \; \forall \, c > \tilde{c}^m.$$

These results, together with an analogous proof for the right cutpoints v_r^m and v_l^m imply that for $\varepsilon > \tilde{c}^a - \tilde{c}^m$, $[v_l^m, v_r^m] \subseteq [v_l^a, v_r^a] \, \forall \, c \geq 0$.

A.5 Proof of Proposition 3

The arguments in the proof of **Lemma 1** establish the existence and uniqueness of $\bar{c}$. The introduction of a small cost of voting $\varepsilon \in (0, \tilde{c}^a - \tilde{c}^m)$ shifts $v_l^a(c, \varepsilon)$ by ε, such that $v_l^a(0, \varepsilon) < v_l^m(0) = \frac{1}{2}$. From (10*) and (14*), $v_l^m(\tilde{c}^m) = v_l^a(\tilde{c}^a - \varepsilon, \varepsilon) = 0$, with $\tilde{c}^m < \tilde{c}^a - \varepsilon$. From the proof of **Lemma 1** $\frac{dv_l^m}{dc} < \frac{dv_l^a}{dc} < 0$. The intermediate value theorem guarantees existence of a unique $\bar{c} \in (0, \tilde{c}^a - \varepsilon)$, such that $v_l^m(\bar{c}) = v_l^a(\bar{c})$ and $v_l^a(c) < v_l^m(c)$ for $c \in \big[0, \bar{c}\big)$, and $v_l^m(c) \leq v_l^a(c)$ for $c \geq \bar{c}$.

A symmetric result follows for v_r^m and v_l^m, implying that when $\varepsilon \in \big(0, \tilde{c}^a - \tilde{c}^m\big)$, $\exists \, \bar{c}$ such that $[v_l^m, v_r^m] \subset [v_l^a, v_r^a]$ for $c \in \big[0, \bar{c}\big)$, and $[v_l^a, v_r^a] \subseteq [v_l^m, v_r^m]$ for $c \geq \bar{c}$.

B Instructions from experiment

Introduction

Welcome to this experiment. The following instructions will explain how you can earn money. We will go over these instructions with you.

It is important that you remain silent and do not look at other people's work. If you have any questions, or need assistance of any kind, please raise your hand and an experimenter will come to you. If you talk, laugh, exclaim out loud, etc., you will be asked to leave and you will not be paid. We expect and appreciate your cooperation.

For today's experiment, you will receive $5 for participating, plus additional earnings during the experiment that depend on your decisions and the decisions of others. Your earnings in the experiment are expressed in points. 70 points are worth $1. At the end of the experiment, your total earnings in points will be exchanged into dollars and paid to you in cash, privately.

The experiment consists of two parts, labeled Part 1 and Part 2. Each part has 25 decision rounds. We will read you the instructions for Part 1 now. After completing Part 1 we will read instructions for Part 2. After we have read the instructions, there will be time to ask clarifying questions. When we are done going through the instructions for Part 1, each of you will have to answer a few brief questions to ensure everyone understands.

Instructions Part 1

Groups

At the beginning of each round, all participants will be randomly divided into groups of 5. You will only interact with your group through the computer interface. You will not know who the other members of your group are. Within each round you will have no interaction with participants who are randomly assigned to other groups.

Groups are randomly determined *every round*, so that the participants in your group are probably not the same between rounds.

Instructions Part 1

BEST OUTCOMES

Each round the computer randomly selects a BEST OUTCOME for each participant. Each participant's BEST OUTCOME will be an integer from 1-100. Each of the values 1-100 are equally likely. Different group members will generally have different BEST OUTCOMEs.

You can think of the process of selecting the BEST OUTCOME for a participant as the computer putting 100 numbered balls in a bag, thoroughly mixing them, and drawing one out at random. To select the BEST OUTCOME for another participant, the computer puts the previously chosen ball back in the bag, mixes the balls again, and then randomly draws a ball. This process is followed for every participant in each round. BEST OUTCOMEs are independent across rounds and participants.

You receive the highest benefits if the GROUP OUTCOME equals your own BEST OUTCOME, and receive lower benefits the further the GROUP OUTCOME is from your own BEST OUTCOME. We will explain more about the GROUP OUTCOME shortly.

Instructions Part 1

GROUP OUTCOMES

In each round the GROUP OUTCOME is determined in part by the group choice, and in part by chance.

In each round, your group will choose between OPTION A and OPTION B. This is done by majority vote. In the event of a tied vote, the group choice is determined randomly, with each OPTION being equally likely. All group members are able to submit a vote, but are not required to do so. Submitting a vote costs 2 points.

If your group chooses OPTION A, the computer will randomly select a GROUP OUTCOME that is an integer between 1-50. Each of the values 1-50 will be equally likely if OPTION A is chosen.

If your group chooses OPTION B, the computer will randomly select a GROUP OUTCOME that is an integer between 51-100. Each of the values 51-100 will be equally likely if OPTION B is the chosen.

The GROUP OUTCOME selected by the computer is not affected by the GROUP OUTCOME in any other group or any other round. GROUP OUTCOMEs are selected independently.

Instructions Part 1

Information Decision

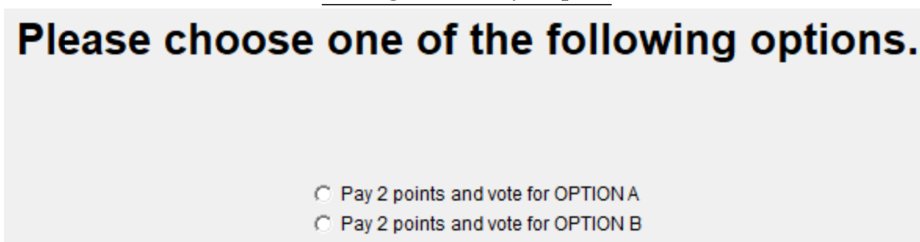
In each round, there will be two OPTIONS on the computer screen, OPTION A and OPTION B. The ordering of the OPTIONs will be randomized. That is, in one round OPTION B may be listed first, but in another round OPTION A may come first. This random ordering is determined each round, and is not affected by the ordering in any other round.

Initially, the two OPTIONs are indistinguishable. That is, the computer screen will display two identical buttons labeled OPTION ? and OPTION ?. Which one of these buttons is OPTION A, and which is OPTION B, is not revealed.

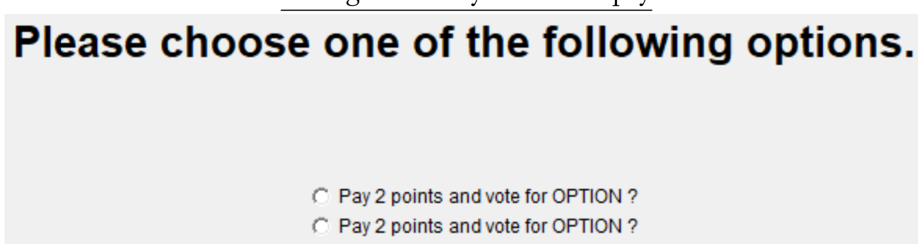
In each round, before voting, each participant may pay to learn which OPTION is OPTION A and which is OPTION B. The cost of paying is 15 points.

Example of information purchase:

Voting screen if you paid



Voting screen if you did not pay



Instructions Part 1

Decision Making Stages

Each round consists of 2 stages: Stage 1 and Stage 2.

Stage 1

The computer randomly selects the individual BEST OUTCOME for each group member. Remember that each BEST OUTCOME can be any number from 1-100, and that BEST OUTCOMES are independent from one another.

Each group member chooses whether or not to pay to learn OPTION identities in that round. The cost of learning OPTION identities is 15 points.

Stage 2

The OPTIONs are indicated on each group member's computer screen. If a group member has not paid to learn OPTION identities, the decision buttons will be labeled OPTION ? and OPTION ?. If a participant has paid to learn OPTION identities, he/she will see the true labels: one will be OPTION A and one will be OPTION B. The ordering of these OPTIONs will be randomized.

Each group member is able to submit a single vote for one of the OPTIONs, by clicking on the respective button. Submitting a vote costs 2 points. If a participant decides not to submit a vote, they indicate this by clicking the button label "Pay 0 and do not vote".

The OPTION with the most votes in your group is the group choice. In the event of a tied vote, the group choice is determined randomly, with each OPTION being equally likely.

If OPTION A is the group choice, the computer randomly determines a GROUP OUTCOME between 1-50. Each of the values 1-50 is equally likely.

If OPTION B is the group choice, the computer randomly determines a GROUP OUTCOME between 51-100. Each of the values 51-100 is equally likely.

Your Round PAYOFF

Your round PAYOFF will depend on 3 factors: the distance between your own BEST OUTCOME and the GROUP OUTCOME, whether you chose to pay to learn OPTION identities, and whether you chose to vote.

If you paid to learn the OPTION identities and decided to vote, your PAYOFF is

$$100 - |\text{YOUR BEST OUTCOME} - \text{GROUP OUTCOME}| - 15 - 2$$

If you paid to learn the OPTION identities and decided not to vote, your PAYOFF is

$$100 - |\text{YOUR BEST OUTCOME} - \text{GROUP OUTCOME}| - 15$$

If you did not pay to learn the OPTION identities, and you decided to vote your PAYOFF is

$$100 - |\text{YOUR BEST OUTCOME} - \text{GROUP OUTCOME}| - 2$$

If you did not pay to learn the OPTION identities, and you decided not to vote, your PAYOFF is

$$100 - |\text{YOUR BEST OUTCOME} - \text{GROUP OUTCOME}|$$

Example 1

At the beginning of a round, the computer selects 60 as your BEST OUTCOME.

You pay to learn OPTION identities, which costs 15 points.

You vote, which costs 2 points.

OPTION A is the group choice. The computer then randomly selects 15 as the GROUP OUTCOME

Your round PAYOFF is

$$100 - |60 - 15| - 15 - 2 = 38$$

Example 2

In the next round, you are randomly placed in a new 5-person group.

The computer selects 80 as your new BEST OUTCOME.

You choose not to pay to learn OPTION identities.

You do not vote.

OPTION B is the group choice. The computer then randomly selects 70 as the GROUP OUTCOME.

Your round PAYOFF is

$$100 - |80 - 70| = 90$$

Example 3

In the next round, you are randomly placed in a new 5-person group.

The computer selects 15 as your new BEST OUTCOME.

You choose to not pay to learn OPTION identities.

You vote, which costs 2 points.

OPTION B is the group choice. The computer then randomly selects 90 as the GROUP OUTCOME.

Your round PAYOFF is

$$100 - |15 - 90| - 2 = 23$$

Example 4

In the next round, you are randomly placed in a new 5-person group.

The computer selects 15 as your new BEST OUTCOME.

You choose to pay to learn OPTION identities.

You do not vote.

OPTION A is the group choice. The computer then randomly selects 25 as the GROUP OUTCOME.

Your round PAYOFF is

$$100 - |15 - 25| - 15 = 75$$

Selecting periods for payment

Once all 50 rounds of *the experiment* have been completed, 20 rounds will be randomly chosen for payment. Each of the 50 rounds are equally likely to be chosen for payment.

Your total earnings for the experiment are equal to the sum of the PAYOFFs in the 20 randomly chosen rounds. You will learn which of these 50 rounds have been selected for payment at the end off the experiment.

Instructions Part 1

Summary

1. In each round the computer randomly assigns each participant a BEST OUTCOME, and randomly sorts players into groups of 5.
2. In each round, group members must decide between two OPTIONs, OPTION A and OPTION B. However, group members do not know which OPTION is which. Each group member may choose to learn OPTION identities for a cost of 15 points.
3. After all members have decided whether or not to pay to learn OPTION identities, voting takes place. Every group member can vote, but is not required to do so. Voting costs 2 points.
4. The group choice is the OPTION which gets the majority of votes (ties are broken randomly). If OPTION A is the group choice, the GROUP OUTCOME will be between 1-50. If OPTION B is the group choice, the GROUP OUTCOME will be between 51-100.
5. The PAYOFF for group members who pay to learn OPTION identities and who vote is

$$100 - |\text{BEST OUTCOME} - \text{GROUP OUTCOME}| - 2 - 15$$

6. The PAYOFF for group members who pay to learn OPTION identities and who do not vote is

$$100 - |\text{BEST OUTCOME} - \text{GROUP OUTCOME}| - 15$$

7. The PAYOFF for group members who do not pay to learn OPTION identities and who vote is

$$100 - |\text{BEST OUTCOME} - \text{GROUP OUTCOME}| - 2$$

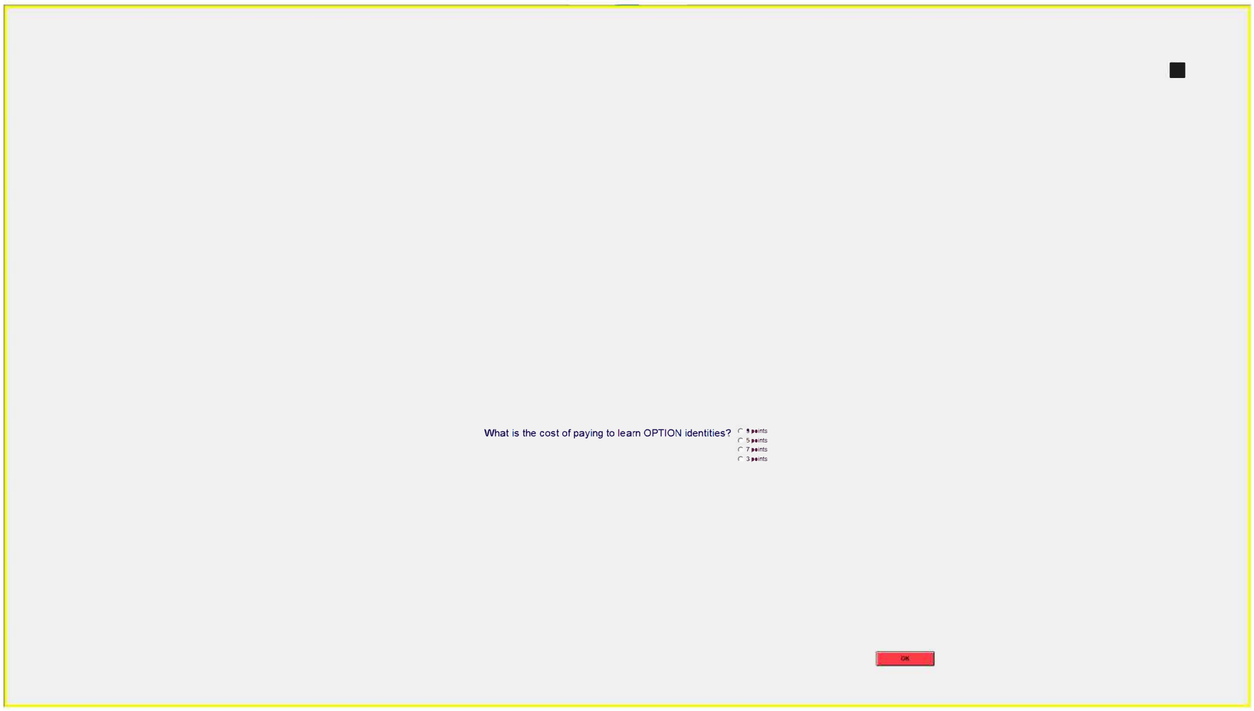
8. The PAYOFF for group members who do not pay to learn OPTION identities and who do not vote is

$$100 - |\text{BEST OUTCOME} - \text{GROUP OUTCOME}|$$

Instructions Part 2

Part 2 of the experiment has 25 decision rounds. Part 2 is the same as Part 1, except that the cost of paying to learn OPTION identities is 3 points.

C Quiz from experiment


What is the cost of paying to learn OPTION identities?
9 points
5 points
7 points
3 points
OK


What are the possible GROUP OUTCOMES if your group chooses OPTION A?
1-50
51-100
25-75
1-100
OK


Suppose that in a round, your BEST OUTCOME is 22. You then choose to pay to learn OPTION identities, for a cost of 9 points. You then vote for OPTION A, which costs 2 points. The group chooses OPTION A, and the computer randomly determines a GROUP OUTCOME of 32. What is your PAYOFF for the round (in points)?
100 - |22-32| - 9 - 2 = 79
100 - |22-50| - 9 - 2 = 61
100 - |22-32| - 2 = 88
100 - |22-32| - 9 = 81
OK


Suppose that in a round, your BEST OUTCOME is 63. You then choose to not pay to learn OPTION identities. You then vote for OPTION ?, which costs 2 points. The group chooses OPTION A, and the computer randomly determines a GROUP OUTCOME of 12. What is your PAYOFF for the round (in points)?
100 - |63-12| - 2 = 47
100 - |63-50| - 9 - 2 = 76
100 - |63-12| - 9 - 2 = 38
100 - |63-12| - 9 = 40
OK

You have successfuly completed the questions.

Please wait quietly until everyone has completed the questions
When everyone has finished, we will begin.

D AQRE Predictions

In this section, we derive the logit agent quantal response equilibrium (AQRE) for the voting games.

The AQRE predictions are found by solving a system of equations that accounts for each possible decision at each node, for each voter type. Given the symmetry in voter types and candidate types in our game, we proceed to solve for the AQRE predictions for a voter v_i^A, noting that predictions for each voter types $v_i^B = 1 - v_i^A$ will be same as for a voter type v_i^A.

D.1 Mandatory Voting

Under mandatory voting, AQRE allows for errors at the information acquisition decision, as well as the voting decision when a subject is informed.

The AQRE conditions are

$$P_i = \frac{1}{1 + e^{-\lambda VP_i}} \tag{17}$$

$$Q_i = \frac{1}{1 + e^{-\lambda VQ_i}} \tag{18}$$

where P_i represents the probability of paying to observe information about candidate identities and Q_i is the probability of voting for the candidate that does not maximize v_i's expected payoff (the "wrong" candidate) after paying to get information, for each $i \in \{1, 2, ..., 50\}$. VP_i represents a voter v_i's net expected profit from paying to learn about candidate preferences, and VQ_i is the net expected profit from voting for the "wrong" candidate after learning candidate preferences. For a voter v_i^A,

$$VP_i = \frac{1}{2} ppiv \Big(E\big[U_i|\gamma_b\big] - E\big[U_i|\gamma_a\big] \Big) - c$$
$$VQ_i = -\frac{1}{2} ppiv \Big(E\big[U_i|\gamma_b\big] - E\big[U_i|\gamma_a\big] \Big)$$

where the probability of being the pivotal voter $ppiv$ is

$$ppiv = \binom{n-1}{\frac{n-1}{2}} \frac{1}{2}^{n-1}.$$

The AQRE predictions for P_i and Q_i under mandatory voting are found for a given λ (possibly scalar or vector-valued) by simultaneously solving the 100 equations from (17) and (18). This is discussed further in section D.3.

D.2 Voting with Abstention

When abstention is allowed, AQRE accounts for errors at each stage of the game. P_i and Q_i retain the same interpretation as in the case of mandatory voting. However, now it is the case that $M_i + Q_i + R_i = 1$, where M_i is the probability of voting for the preferred candidate after learning candidate identities, and R_i is the probability of abstaining after learning candidate identities. Z_i is the probability of voting for any candidate after opting not to learn candidate identities. The AQRE conditions under optional voting are

$$P_i = \frac{1}{1 + e^{-\lambda VP_i(\cdot)}} \tag{19}$$

$$Q_i = \frac{1}{1 + e^{-\lambda(VQ_i(\cdot) - VM_i(\cdot))} + e^{-\lambda VQ_i(\cdot)}} \tag{20}$$

$$R_i = \frac{1}{1 + e^{\lambda VQ_i(\cdot)} + e^{\lambda VM_i(\cdot)}} \tag{21}$$

$$M_i = \frac{1}{1 + e^{-\lambda(VM_i(\cdot) - VQ_i(\cdot))} + e^{-\lambda VM_i(\cdot)}} \tag{22}$$

$$Z_i = \frac{1}{1 + e^{-\lambda VZ}} \tag{23}$$

where $VP_i(\cdot), VQ_i(\cdot), VM_i(\cdot)$, and VZ are a voter i's net expected profit from becoming informed, voting for the "wrong" candidate when informed, voting for the "correct" candidate when informed, and voting for any candidate when uninformed. For voter v_i^A,

$$VP_i(P_{-i}, Q_{-i}, R_{-i}, M_{-i}, Z_i) = \frac{1}{2} ppiv\Big(E\big[U_i|\gamma_b\big] - E\big[U_i|\gamma_a\big]\Big)(M_i - Q_i) - c - (M_i + Q_i)\varepsilon$$

$$VQ_i(P_{-i}, Q_{-i}, R_{-i}, M_{-i}, Z_i) = \frac{1}{2} ppiv\Big(E\big[U_i|\gamma_b\big] - E\big[U_i|\gamma_a\big]\Big)$$

$$VM_i(P_{-i}, Q_{-i}, R_{-i}, M_{-i}, Z_i) = -\frac{1}{2} ppiv\Big(E\big[U_i|\gamma_b\big] - E\big[U_i|\gamma_a\big]\Big)$$

$$VZ = -2.$$

Allowing for abstention, the probability of being the pivotal voter $ppiv$ is

$$ppiv = \sum_{t=0}^{\frac{n-1}{2}} \sum_{s=0}^{1} \binom{n-1}{2t+s} \binom{2t+s}{t+s} \big(1 - 2\tilde{P}\big)^{n-1-s-2t} \tilde{P}^{2t+s},$$

where

$$\tilde{P} = \frac{1}{2}\big[p(1-r) + (1-p)z\big]$$

and

$$p = \frac{1}{N}\sum_{i=1}^{N} P_i$$

$$r = \frac{1}{N}\sum_{i=1}^{N} R_i$$

$$z = \frac{1}{N}\sum_{i=1}^{N} Z_i$$

for $i \in \{1, 2, ..., 50\}$.

The AQRE predictions for P_i, Q_i, R_i, M_i and Z_i are found for a given λ (possibly scalar or vector-valued) by simultaneously solving the 250 equations from (19) - (23).

D.3 Selecting λ

We consider cases in which λ differs by treatment, as well as by decision stage. Specifically, we allow λ to differ between the decision to pay to learn candidate identities and the voting decision, as well as across election structure. Note that we do not allow λ to vary by type, as this would result in an extremely large number of free parameters.

Our ex-ante model selection criteria among the possible methods for calculating λ was the BIC. We solve the system of equations in (17) - (18) and (19) - (23) for a unified λ, λ's that varied by election structure, λ's that varied by decision, and interactions of these cases. Within each case, we use grid search to select the optimal λ (scalar or vector-valued) by maximum likelihood, following the procedure described in Großer and Palfrey (2019).[27]

After selecting λ by maximum likelihood, we calculate the BIC for each method. The BIC is minimized when we allow the λ to vary both by treatment and by decision. The resultant λ values are reported in Table A1.

[27] We thank Jens Großer for generously providing details and the associated code for this procedure. Our implementation is done in Mathematica and R, and the associated code is available on request.

Table A1. λ Values by Treatment and by Decision

	λ_{info}	$\lambda_{vote\|PAY=1}$	$\lambda_{vote\|PAY=0}$
AH	0.0475	0.085	0.600
AM	0.015	0.070	0.750
AL	0.030	0.110	0.850
MH	0.173	0.074	-
MM	0.124	0.114	-
ML	0.088	0.120	-

λ_{info} reports the λ at the information acquisition decision. $\lambda_{vote|PAY=1}$ reports the λ at the voting stage, conditional on paying for information, and $\lambda_{vote|PAY=0}$ conditional on remaining uninformed.

E AQRE-driven Power Calculations

Our initial experiment was predicated on power analysis that presumed Nash equilibrium predictions represented the true data generating process. Our initial experiment included treatments with two costs of information: $c = 3$ and $c = 9$. The results from these early sessions provided some support for the theoretical predictions of our model of information acquisition. However, as discussed in section 4.3, when abstention is permitted, uninformed voters often opt to (randomly) vote, contra Nash predictions. This results in smaller differences in the observed probability of being the pivotal voter across voting rules, which, in turn, reduces the incentive to become informed when abstention is permitted. Accounting for these voting "errors" becomes important when analyzing treatment effects, and when performing pre-experiment power analyses. To perform a useful power analysis, it is necessary to account for the amount of noise that is likely to be present in subject decisions, rather than simply assume that Nash equilibrium is the true data generating process. One compelling way of accounting for this noise is to assume that boundedly rational subjects will play a quantal response equilibrium (QRE).

A novel approach to power analysis that accounts for these errors using QRE predictions is provided in Woods (2021). This approach suggests using the QRE λ estimates as part of a power analysis to determine how many observations are necessary to achieve a sufficiently high probability of detecting the effect of treatment, if it exists.

We follow the simulation methods outlined in Woods (2021) to perform a power analysis for the first sessions where $c \in \{3, 9\}$. We use the λ parameters obtained from our first sessions to determine the decisions of 50,000 simulated subjects. We then randomly sample these subjects to form 1000 simulated "experiments". Each simulated experiment was structured identically to our experimental data. Each simulated experiment consisted of 6 blocks (1 block is 25 rounds) with mandatory voting, and 6 blocks where abstention was allowed. In each block, 15 simulated subjects were formed into groups of 5 and were randomly assigned a type in each round. For each simulated experiment, we performed the Mann-Whitney U test on the simulated data to compare the decision to pay for information across the voting schemes.

The results of our AQRE-driven power analysis varied across treatments. Tests for within-subject treatments, which varied the cost of information for a fixed voting structure, achieved conventional power levels of 0.80 or higher with 6 blocks per treatment (the number of blocks in our early data). The test of the between-subject treatment at $c = 3$ also achieved conventional power with 6 blocks per treatment. However, at $c = 9$, the number of required observations to achieve conventional power levels was orders of magnitude higher. Additionally, the AQRE predicted that at $c = 9$ information acquisition would be slightly higher under mandatory voting, contrary to PBE predictions. This difference is driven by noisy voting behavior, which dilutes the predicted treatment effect.

This observation leads to a new, testable hypothesis: will increasing c even further result in more informed voters under optional voting (the PBE prediction we did not find when $c = 9$)? By investigating this question, we can test whether our early results were driven by uninformed voting, as well as testing the predictive power of AQRE. That is, rather than simply fit the data ex post, we can make an ex ante prediction based on AQRE predictions and run an experimental test.

Again following Woods (2021), we conducted power analysis for $c = 15$ using the λ parameters from our early data with $c = 9$. The analysis found that at $c = 15$, 6 blocks per treatment is sufficiently powered. We gathered 6 blocks per treatment by running 12 additional sessions, varying the cost of information between $c = 3$ and $c = 15$. The experimental protocol, including balancing the ordering of within-subject treatments, was identical to our early data varying c between $c = 3$ and $c = 9$. Doing so, we maintained balance with our earlier data, which also contained 12 sessions of balanced treatments between $c = 3$ and $c = 9$.