Voting with Endogenous Information Acquisition: Experimental Evidence*

Sourav Bhattacharya† John Duffy‡ Sun-Tak Kim§

December 23, 2016

Abstract

The Condorcet jury model with costless but informative signals about the true state of the world predicts that the efficiency of group decision-making increases unambiguously with the group size. However, if signal acquisition is made an endogenous and costly decision, then rational voters have disincentives to purchase information as the group size becomes larger. We investigate the extent to which human subjects recognize this trade-off between better information aggregation and greater incentives to free-ride in a laboratory experiment where we vary the group size, the cost of information acquisition and the precision of signals. We find that the theory predicts well in the case of precise signals. However, when signals are imprecise, free-riding incentives appear to be much weaker as there is a pronounced tendency for subjects to over-acquire information relative to equilibrium predictions. We rationalize the latter finding using a quantal response equilibrium that allows for risk aversion.
JEL Codes: C72, D72, D81.
Keywords: Voting, Condorcet Jury Model, Information Aggregation, Endogenous Information Acquisition, Experimental Economics.

---

*We thank the editor, Marco Battaglini, two anonymous referees and participants at various conferences and workshops for helpful comments and suggestions on earlier drafts. We gratefully acknowledge funding from a U.S. National Science Foundation Doctoral Dissertation Grant, #SES-1123914 awarded to SunTak Kim.

†Department of Economics, Royal Holloway, University of London. Email: sourav.bhattacharya@rhul.ac.uk

‡Corresponding author, Department of Economics, University of California, Irvine. Email: duffy@uci.edu

§Department of Economics, National Taiwan University. Email: sunkim@ntu.edu.tw

1 Introduction

Condorcet's jury theorem (Condorcet 1785) asserts that if a group of individuals have common preferences over some binary outcome (e.g., convicting the guilty or acquitting the innocent) and are given independent and informative private signals about the true state of the world (e.g., "guilt" or "innocence") then, under majority rule, the correct outcome is more likely to be achieved as the group size of voters is increased. Feddersen and Pesendorfer (1997) have shown that this result is robust to strategic or insincere voting, where voters may rationally vote against their private information; even if voters vote strategically against their signals, they do so in an optimal way so that information aggregation continues to improve as the group size increases. An implication of these results for optimal voting mechanisms is that, under the maintained assumptions, we can always make a voting mechanism better by adding more voters. However, this result assumes that private signals about the true but unknown state of the world are costless and exogenously provided.

In this paper we study the question of endogenous information aggregation in a setting where voters must first independently decide whether to acquire a costly signal about the true state of the world prior to voting as a group whether to convict or acquit under majority rule. In particular, we present results from a laboratory experiment designed to explore how the number of players, the cost of information and the informativeness of signals matter for information aggregation by juries or committees. We believe that a laboratory experiment provides the best means of empirically evaluating the theory of voting and information aggregation with endogenous information acquisition as the laboratory allows for firm control over the number of voters, the costs and precision of information that voter receive as well as the incentives that voters face, so that the theory can be properly tested.

The basic set-up of our experiment is the Condorcet jury model in which voters have common preferences and must make a decision as a group about whether to convict or acquit a defendant based on private, informative signals about whether the defendant is guilty or innocent. A main focus of our study is how the size of the group affects the probability that it makes the correct decision (henceforth referred to as informational efficiency). Theory suggests that adding an additional individual (or voter) to the group has two opposing effects. On the one hand, since the additional individual's signal is informative – it is more likely to be correct than incorrect – efficiency will increase. We term this the *information aggregation effect*, and the content of the various versions of the Condorcet Jury Theorem is that when voters are exogenously endowed with private, independent but informative signals about the state of the world, this effect ensures that arbitrarily large groups can reduce the likelihood of error in the group decision without bound, thus improving informational efficiency. However, when the acquisition of information (signals) is a costly choice, then as the group size increases, each individual has a lower incentive to acquire information. This countervailing *free-riding effect* works to reduce informational efficiency. Thus, when information is endogenously chosen and costly, the overall effect of group size on informational efficiency depends on the tradeoff between the information aggregation effect and the free riding effect. Persico (2004) and Koriyama and Szentes (2009) show the existence of an upper bound on the optimal group size in Condorcet jury environments with costly information acquisition.

These theoretical papers provide us with testable hypotheses that we evaluate in our laboratory experiment. In particular, increases in the group size should result in an increase in informational efficiency when information is informative and freely available. However, if information acquisition is costly, informational efficiency should only increase up to a certain group size before falling off and for large enough group sizes, reaching the minimum efficiency level. Depending on the model parameterization, all voters may have an incentive to acquire information up to a certain group size, but beyond that group size rational voters play a mixed strategy with regard to information acquisition, and for a large enough group sizes, rational voters should refuse to acquire any information at all. Thus, the theory puts an upper bound on the optimal group size and one purpose of our experiment is to determine whether this upper bound really matters among the laboratory subjects who are asked to make a decision about the purchase of costly information prior to voting. In addition to increasing the group size, we also vary the cost of information acquisition and the precision of the signal process.

To preview our results, we find that if signals are costly and noisy (but informative), the free-riding effect on information acquisition that is predicted to become dominant as the group size increases is actually rather weak, so that the information aggregation effect associated with a larger group size tends to dominate and thus welfare is generally increasing with the group size, counter to theoretical predictions. On the other hand, consistent with theoretical predictions, we find that if signals are costly and perfectly informative, then there is a drop in welfare as the group size increases in line with theoretical predictions. We then consider several explanations for why the group size effect is not as strong in the noisy signal environment as compared with the perfect signal environment.

Specifically, we first consider whether subjects might simply be coordinating on asymmetric equilibria as opposed to the symmetric equilibria that we focus on. We find, however, that these two different types of equilibria are not sufficiently distinct from one another to provide a meaningful explanation. We then consider several different behavioral explanations for our findings including 1) that subjects may approach the game in decision-theoretic rather than game-theoretic terms thereby ignoring free-riding considerations; 2) that behavior reflects noisy best responses so that a quantal response rather than a Nash equilibrium is the appropriate benchmark for analysis and finally, 3) that subjects are risk averse with regard to uncertain money payoffs (rather than risk neutral as the theory presumes), and this risk aversion leads them to over-acquire information in the noisy signal environment. We conclude that a quantal response equilibrium with risk averse preferences provides a compelling explanation for why behavior departs from theoretical predictions in the noisy signal environment.

The rest of the paper is organized as follows. Section 2 discusses related literature. Section 3 presents the theoretical model and equilibrium predictions. Section 4 describes our experimental design and in section 5 we state our research hypotheses with numerical predictions under the parameterizations used in the experiment. Section 6 presents our main experimental findings in comparison with theoretical explanations and we also evaluate the various behavioral explanations for why, in certain treatments, information acquisition departs from theoretical predictions. Finally, section 7 concludes with a summary of our main findings and some suggestions for future research.

2 Related Literature

The theory of endogenous information acquisition in the Condorcet jury model begins with Persico (2004) and Martinelli (2006). Persico (2004) observed that if agents must first decide whether to acquire private noisy information that is then aggregated to reach a collective decision, then the information acquisition decision is properly viewed as a free-rider problem with the result that information acquisition will generally be less than the social optimum under a given voting rule. An implication of this observation is that for any given signal precision and voting rule there will exist an optimal committee size, and in contrast to the standard Condorcet Jury Theorem, larger committees will not always be welfare-improving. Martinelli (2006, 2007) studies endogenous, costly and noisy information acquisition but considers the case where the signal precision is the choice variable, with more precise signals being more costly. Martinelli shows that if the marginal cost of the signal precision is zero at the lowest level of precision, then voters acquire some information even in large electorates and that the voting outcome can be (under certain assumptions) asymptotically efficient. Mukhopadhaya (2003) and Koriyama and Szentes (2009) also explore the Condorcet Jury model under endogenous information acquisition and show that larger than optimal committee sizes do lead to social welfare losses relative to smaller committee sizes, but that these losses might not be so great. Gerardi and Yariv (2008) show that the optimal voting mechanism is in general not ex-post efficient; distortions have to be introduced to ensure that agents have incentives to acquire information. Oliveros (2013) adds abstention and heterogeneity in intensity of preferences to Martinelli's endogenous information choice setting and shows that those acquiring more precise information do not necessarily abstain less often.

The first experimental studies of the Condorcet jury model all studied environments where informative signals were *exogenously* and freely provided to voters (there was no information acquisition decision): Guarnaschelli et al. (2000), Ali et al. (2008), Battaglini et al. (2010), Goeree and Yariv (2011), Bhattacharya, et al. (2014) and Anderson et al. (2015). These studies focus on the extent of strategic versus sincere voting under a variety of different conditions: different voting rules (majority or unanimity), timing assumptions (simultaneous or sequential), committee durations (ad hoc/one-shot or standing/extended), asymmetrically informed voters, preplay communication, compulsory versus voluntary voting (abstention), and differential costs associated with incorrect group decisions. More recent experimental studies by Großer and Seebauer (2016) Elbittar et al. (2017) and Mechtenberg and Tyran (2016) use the Condorcet jury model to explore the consequences of allowing *endogenous* costly information acquisition prior to voting as we do in this paper.[1] Großer and Seebauer (2016) study costly information acquisition by groups of size 3 or 7 and focus on whether compulsory rather than voluntary voting (where abstention is allowed), provides greater incentives for voters to acquire information (it does). Elbittar et al. (2017) explore endogenous information acquisition under a voluntary voting mechanism focusing on the extent to which the voting rule, majority or unanimity, matters for information acquisition and participation in voting. Mechtenberg and Tyran (2016) study costly information acquisition prior to voluntary voting (abstention allowed) where voters can first petition to hold a vote or defer the decision to an expert. They report that when subjects can first petition to hold a vote, information acquisition is greater than when the vote is exogenously imposed. By contrast with these three studies, we consider only the majority rule, compulsory voting setting (no abstention) where we focus not only on the effect of changes in group size, but also on changes in the cost of acquiring information (signals) about the true state of the world as well as the precision of those signals. Our paper is the only one in this literature to consider cases where subjects can acquire, at a fixed cost, noisy or perfect signals of the true state of the world. Our design thus enables a more complete assessment of the comparative statics implications of group size, information cost, and signal precision for endogenous, costly information acquisition all under the simple majority rule, compulsory voting mechanism.

[1] We only became aware of these other studies after we had begun working on this project.

3 The Model

Our experiment implements the standard Condorcet Jury model with the addition of an endogenous, costly information acquisition stage that takes place prior to the voting stage as in Persico (2004) and Martinelli (2006, 2007). Within this environment, we consider the comparative statics implications of varying the information acquisition cost, c, the group size, N, and the signal precision, x.

In all of our experimental settings (or "treatments") a group consisting of an odd number, N, of individuals faces a choice between two alternatives, labeled R (Red) and B (Blue). The group's choice is made in an election decided by majority rule, that is, the alternative, R or B that receives more than $N/2$ votes is the group's decision. It is common knowledge among voters that there are two equally likely states of nature, ρ and β, i.e., all voters have the common prior $\Pr[\rho] = \Pr[\beta] = .5$. Alternative R is commonly known to be the better choice in state ρ while alternative B is commonly known to be the better choice in state β. Specifically, in state ρ each group member earns a payoff of $M > 0$ if R is the alternative chosen by the group and 0 if B is the chosen alternative. In state β the payoffs from R and B are reversed. Formally, we have:

$$\begin{aligned} U(R|\rho) &= U(B|\beta) = M, \\ U(R|\beta) &= U(B|\rho) = 0. \end{aligned}$$

Prior to the voting decision, each individual may acquire a costly private signal regarding the true, binary state of nature. This signal can take on one of two values, either r or b. The probability of receiving a particular signal depends on the true state of nature. Specifically, each subject choosing to acquire a signal receives a conditionally independent signal where

$$\Pr[r|\rho] = \Pr[b|\beta] = x.$$

Voters who do not acquire a signal have no more information about the true state of the world than the initial common prior that the two states are equally likely.

We suppose that $1/2 < x \leq 1$ so that signals are informative but possibly noisy. More precisely we will consider cases where $1/2 < x < 1$, so that the signal is noisy but informative as well as cases where $x = 1$, and the signal (if purchased) is perfectly informative. The latter case eliminates fundamental uncertainty about the true state of the world so that the voter only faces strategic uncertainty as to the information acquisition choices of other voters.

Given that $x > 1/2$, signal r is associated with state ρ while signal b is associated with state β (we say that r is the correct signal in state ρ while b is the correct signal in state β). It can be easily checked that when the signal precision is *symmetric* and the two priors are equal, the posterior probabilities that signals are matched with the correct states are the same in both states and are given by the signal precision parameter x:

$$\Pr[\rho|r] = \Pr[\beta|b] = x.$$

In our setting, each individual must decide whether or not to acquire a private signal *at a fixed cost* of $c > 0$. If an individual acquires a private signal then her payoff is $U(A|\omega) - c$, where A is the group decision outcome and ω is the state of nature (i.e., payoffs are either $M - c$ or $-c$, depending on the correctness of the group decision). If an individual does not acquire a private signal her payoffs are given by $U(A|\omega)$.

Having specified the preferences and information structure of the model, we next discuss the equilibrium strategies and outcomes for the voting games we study in our experiment. We initially restrict attention to symmetric equilibria in weakly un-dominated strategies as such equilibria would seem to be the most relevant given the information that is available to subjects in our experiment and our use of anonymous random matching to form groups (as detailed later in section 4). In particular, we require that in equilibrium (i) all voters of the same signal type play the same strategies and (ii) no voter uses a weakly dominated strategy. We will discuss later the possibility that subjects coordinate on other asymmetric equilibria, but our design involves a choice of parameters such that a symmetric equilibrium in weakly un-dominated strategies always exists and this symmetric equilibrium is unique.

3.1 Symmetric Equilibrium Predictions

The strategy of a voter consists of three elements: (1) a probability σ of buying costly information, (2) a probability v_s of voting sincerely conditional on buying a signal $s \in \{r, b\}$, and (3) a probability of voting for each alternative conditional on not buying information.

Since we use a setting that is symmetric across alternatives (equal priors in favor of either state, equal signal precision, x, in each state, and simple majority rule), we focus on equilibria where voters vote for each alternative with equal probability if they do not buy information. Moreover, it is easy to show that in any equilibrium, voters must vote sincerely conditional on buying information, i.e., $v_b^* = v_r^* = 1$. Thus, in this setting, the symmetric equilibrium is characterized by σ^*, the equilibrium probability of information acquisition alone. In special cases, this probability may be 0 or 1, i.e., we may have a pure strategy equilibrium. We will sometimes denote the pure action of information acquisition ($\sigma = 1$) by σ_1 and the pure action of not acquiring information ($\sigma = 0$) by σ_0. Our focus on symmetric equilibria, where voters with the same private information use the same strategy, follows the common practice in the literature.

Under costly information acquisition, there may exist multiple equilibria (including asymmetric equilibria) where individuals acquire information with positive probability ($\sigma^* > 0$).[2] As noted earlier, we always choose our parameter values such that the voting game in our experiment has a unique *symmetric* equilibrium.

Given sincere voting conditional on information acquisition, the expected utility from acquiring information is given by

$$U(\sigma_1) = \frac{M}{2} \{Pr[\rho|r] \Pr[Piv|\rho] + Pr[\beta|b] \Pr[Piv|\beta]\} - c \tag{1}$$

and the expected utility from not acquiring information is

$$U(\sigma_0) = \frac{M}{2} \left\{ \frac{1}{2} \Pr[Piv|\rho] + \frac{1}{2} \Pr[Piv|\beta] \right\}. \tag{2}$$

Suppose the probability of information acquisition is $\sigma \in [0, 1]$. Then, the ex-ante likelihood of a voter voting for the correct alternative is

$$z_\sigma = x\sigma + \frac{1}{2}(1 - \sigma). \tag{3}$$

Notice that we must have $\Pr[Piv|\rho] = \Pr[Piv|\beta] = p_\sigma$ (say). Therefore, we have

$$\left.\begin{array}{c} U(\sigma_1) = Mxp_\sigma - c \\ U(\sigma_0) = \frac{M}{2} p_\sigma \end{array}\right\} \text{ where}$$

$$p_\sigma \;=\; \Pr(Piv|\omega) = \binom{N-1}{\frac{N-1}{2}} [z_\sigma(1 - z_\sigma)]^{\frac{N-1}{2}}. \tag{4}$$

The information acquisition part of the strategy σ depends on the sign of $U(\sigma_1) - U(\sigma_0)$, which turns out to be a comparison of the net benefit of information acquisition conditional on being pivotal with the normalized cost. In particular, $\sigma^* \geq 0$ if and only if

$$\left(x - \frac{1}{2}\right) p_\sigma \geq \frac{c}{M} \tag{5}$$

and $\sigma^* = 1$ if the inequality is strict. Notice that the net benefit of information acquisition is itself a function of σ.

[2]Since subjects are randomly matched to form a different group in each round of our experiment (which will be explained in detail in the next section on experimental design), we doubt that subjects could find a way to coordinate on play of an *asymmetric* equilibrium. However, we address the question of whether our subjects coordinated on asymmetric equilibria later on in section 6.3.

Again, the solution value σ^* is then used for the calculation of informational efficiency. It is given by the formula

$$W = \sum_{k=\frac{N+1}{2}}^{N} \binom{N}{k} z_\sigma^k (1 - z_\sigma)^{N-k}. \tag{6}$$

4 Experimental Design

We consider three main treatment variables: 1) the group size, N 2) the information cost, c and (3) the signal precision, x. We adopt a between subjects experimental design so that in each session, subjects only make decisions under a single set of the three treatment variables.[3]

The experiment was computerized and was presented to subjects as an abstract group decision-making task using neutral language that avoids any direct reference to voting, elections, jury deliberation, etc., so as not to trigger some other possible (non-theoretical) motivations for voting (e.g., civic duty, the sanction of peers, etc.). A sample of the written instructions that were given to subjects and read aloud prior to the start of each experimental session are provided in Appendix A.

Each session consisted of $2N$ inexperienced subjects and 25 rounds. At the start of each of these 25 rounds, subjects were randomly and anonymously matched into two groups of size N and this random formation of N-member groups at the start of each round was made public knowledge in the written instructions.[4] One of the two groups of size N was randomly assigned to the red jar (state ρ) and the other was assigned to the blue jar (state β), thus fixing the true state of nature for each group and ensuring that we have an equal number of ρ and β states. No subject knew which jar was assigned to her group. Subjects *did* know that it was equally likely that their group was assigned to either the red or the blue jar at the start of each round, that is, we took care to implement this common prior belief among the subjects.

Both the red and blue jars held 10 ball each. The red jar was known to contain a fraction x of *red* balls (signal r) and a fraction $1 - x$ of *blue* balls (signal b) while the blue jar was known to contain a fraction x of *blue* balls and a fraction $1-x$ of *red* balls. We fixed the signal precision at either $x = 0.7$ or at $x = 1$ in a given session, and these

[3]That is, in each session, the group size N, information cost c, and signal precision x, are fixed over all rounds of the session.

[4]Following the terminology of Ali et al. (2008), we use an "ad hoc committee" design. Our intention was to disrupt repeated game dynamics that could arise under the alternative fixed-match "standing committee" design, which could enable coordination on *asymmetric* equilibria or other collusive outcomes. Our random, ad hoc committee design is intended to make the *symmetric* equilibrium predictions of the model as salient as possible.

signal precisions were made public knowledge in the written instructions. We thus implemented symmetric signal precisions so as to facilitate subjects' understanding of equilibrium strategies in the compound decision making environment that we study. In addition, as we have previously noted, symmetric signal precisions rule out strategic (insincere) voting under the majority rule compulsory voting mechanism that we employed.

The sequence of moves in each round of our experiment was as follows. First, each subject had to decide whether or not to pay the fixed and known per round cost, $c > 0$, to draw a ball from the jar randomly assigned to her group for that round. If a subject decided to pay the cost and draw a ball, this was then done virtually in our computerized experiment; subjects clicked on one of 10 balls on their decision screen and the color of their chosen ball was privately revealed to them. In each round and for each subject, the assignment of colors to the 10 ball choices the subject faced were made randomly according to whether the jar the subject was drawing from was the red jar (in which case x percent of the balls were red and $1-x$ percent were blue) or the blue jar (in which case x percent of the balls were blue and $1-x$ percent were red). While each subject observed the color of the ball she had drawn, she did not observe the color of the balls drawn by any other subject or the color of the jar from which she had drawn a ball. The group's common and publicly known objective was to correctly determine the jar, "red" or "blue", that had been assigned to their group.

Subjects who chose not to draw a ball had to wait until other group members (if any) finished drawing a ball. Subjects were seated at computer workstations with privacy dividers so they could not observe the choices of other subjects. After the information acquisition decision was made and any voters who had chosen to acquire information had drawn their ball and observed its color, play proceeded to making a choice between red or blue for the color of the group's jar. All N subjects, regardless of whether or not they chose to draw a ball, had to choose either red or blue for their guess of the color of their group's jar for that round, that is, voting was compulsory and abstention was not allowed. The group's decision, red or blue, was then determined by majority rule. Ties were ruled out by the fact that the group size, N, was always chosen to be an odd number so that a group's decision via majority rule was always unambiguously either red or blue. In our experiment we considered several different group sizes, $N \in \{3, 7, 13\}$; signal precisions, $x \in \{.7, 1\}$ and information acquisition costs, $c \in \{5, 8, 25\}$. As we explain in the next section, these choices for N were made so as to explore various equilibrium predictions of the theory.

Payoffs in each round were determined as follows. First, at the start of each round, each subject was endowed with c points, an amount equal to the treatment-specific cost of acquiring information in each round. If the group's decision via majority rule turned out to be correct, i.e., the group's decision was red (blue) and the jar assigned to that group was in fact the red (blue) jar, then each group member who chose to acquire information prior to voting received 100 points for the round ($M = 100$), while those group members who chose not acquire information prior to voting received $100 + c$ points for the round. If the group's decision turned out to be incorrect, then each group member who chose to acquire information prior to voting received 0 points for the round, while those group members who did not acquire information prior to voting received c points for the round. Thus, if a voter decided not to draw a ball (i.e., buy information), then she kept her endowment of c points and also earned the group-wide payoff in points as well, which depended on whether the group got the decision correct (100 points) or incorrect (0 points). Note that our awarding of an endowment of c points to each subject at the start of each round does not change any of the equilibrium predictions and it avoids the possibility that subjects earn negative payoffs in any round.[5] We vary the magnitude of the information acquisition cost $c \in \{5, 8, 25\}$; as we explain in the next section, these choices for c were made so as to explore various equilibrium and behavioral hypotheses. The parameterization of the payoff function (i.e., the value of M and c) was held constant across all rounds of any given session (i.e., as noted above, we use a between-subjects design).

At the end of each round, after all choices had been made and payoffs determined, subjects received feedback on the results of the round. If a subject had paid the cost to buy a signal, she was reminded of the private signal (red or blue) that she received prior to voting in the round. All subjects were reminded of their "choice" (i.e., their vote) for red or blue for the color of their group's jar for the round. Subjects were then informed of the number of red and blue choices (votes) made by group members, the group's decision (red or blue) according to the majority rule, the true color of their group's jar for the round (red or blue) and whether their group's decision was "correct" or "incorrect". Subjects were also informed of their payoff in points for the round, which was one of four possible values: 0, c, 100 or $100 + c$. Subjects were further informed at the end of each round about the number of subjects in their N-player group who did or did not choose to purchase information (a signal) prior to making a choice (voting) for red or blue in that round.[6]

[5] Levine and Palfrey (2007) and Bhattacharya et al. (2014) use this same design.

[6] Theoretically speaking, such ex-post information on the number of group members acquiring/not acquiring information prior to voting is irrelevant in our one-shot, random-matching (ad hoc committee) design. However, in an initial pilot experiment (not reported in this paper) where we did not provide ex-post information on the number of voters acquiring information, we observed over-acquisition of information relative to equilibrium predictions and so we thought it might be behaviorally important (from a learning perspective) to provide such information, which we did in *all* of the sessions reported on in this paper. As it turned out, information acquisition frequencies remained higher than equilibrium predictions in many (but not all) of our experimental treatments (as detailed below) even with the feedback we provided at the end of each round about the number of group members acquiring/not acquiring information in that round.

Following 25 rounds of play, the session was declared over. Subjects' point totals from all 25 rounds of play were converted into dollars at the fixed and known rate of 1 point = \$0.01 and these dollar earnings were then paid to subjects in cash and in private. In addition, subjects were awarded a \$5 cash show-up payment.

Treatment Conditions			No. of Sessions	No. of Subjects per Session	No. of Rounds per Session
N	c	x			
3	5	0.7	4	6	25
3	8	0.7	4	6	25
3	25	0.7	4	6	25
7	5	0.7	4	14	25
7	8	0.7	4	14	25
13	8	0.7	4	26	25
3	8	1.0	4	6	25
7	8	1.0	4	14	25

Table 1: Characteristics of Experimental Sessions

Table 1 summarizes our experimental design, which consists of 32 voting sessions involving 368 subjects. Subjects were recruited from the undergraduate population of the University of Pittsburgh and the experiment was conducted in the Pittsburgh Experimental Economics Laboratory using Willow, a Python framework for experimental economics.[7] No subject was allowed to participate in more than one session of this experiment. Total earnings including the \$5 show-up payment averaged around \$24 per subject for a 1 hour experiment.

5 Research Hypotheses

Table 2 shows symmetric equilibrium predictions for various combinations of our three treatment parameters: N, c and x. Note that we did not conduct sessions for all of the treatment combinations shown in Table 2 as budget constraints prevented such an exhaustive exercise. The predictions of the treatments for which we *did* collect experimental data are indicated in boldface type in Table 2; other predictions from alternative parameterizations of the model shown in non-boldface type are given for reference purposes only and serve to justify the model parameterizations (treatments) that we did choose to run in the laboratory.[8]

[7]The program we used to conduct the experiment reported in this paper is available upon request.

[8]For example, as Table 2 reveals, when $x = 0.7$, as N increases from 3 to 7 to 13, efficiency (w^*) more rapidly declines to the minimal level of 0.5 as c is increased from 5 to 8 and to as high as 25.

x=0.7	$N=3$		$N=7$		$N=13$	
	σ^*	w^*	σ^*	w^*	σ^*	w^*
5	**1**	**0.784**	**0.6693**	**0.773**	0	0.5
8	**1**	**0.784**	**0**	**0.5**	**0**	**0.5**
25	**0**	**0.5**	0	0.5	0	0.5
x=1	$N=3$		$N=7$		$N=13$	
	σ^*	w^*	σ^*	w^*	σ^*	w^*
$c=5$	0.8944	0.992	0.5621	0.955	0.356	0.912
8	**0.825**	**0.978**	**0.447**	**0.902**	0.2359	0.810

Notes: σ^* indicates the equilibrium rate of information acquisition and w^* indicates equilibrium (informational) efficiency.

Table 2: Symmetric Equilibrium Predictions

Based on the equilibrium predictions shown in Table 2, we formulate three research hypotheses concerning the effect of our three treatment variables on the frequency of information acquisition (and hence on the frequency of a group's making correct decisions - the informational efficiency of group decision-making always moves in the same direction as the rate of information acquisition, as Table 2 reveals).

H1. Group size effect: For any fixed (positive) information cost and signal precision $(c, x) \in \{5, 8, 25\} \times \{0.7, 1\}$, the frequency of information acquisition, σ^*, decreases toward zero or remains at zero as we increase the group size from $N = 3$ to $N = 7$ to $N = 13$.

To see the free riding effect, notice that for any given σ, x and c, as N increases, the pivot probability p_σ, (4), *decreases*, leading to a drop in the benefit of information acquisition (left hand side of inequality (5)). The two competing effects of changing group size on efficiency are easy to see from expression (6). On the one hand, the free riding effect depresses the equilibrium value of σ (and therefore z_σ), which depresses W if N were held constant. On the other hand, the information aggregation effect predicts that, for any strictly positive σ (and therefore, $z_\sigma > \frac{1}{2}$), an increase in N raises W. While these effects work in opposite directions, notice that for any c, x and σ, for a large enough N, the pivot probability, p_σ, will eventually be small enough that the benefit of information acquisition is lower than the cost. Thus, for a large enough group size, the free riding effect dominates so that $\sigma^* = 0$, and, as a result, the likelihood of a correct decision is no better than the prior so that $W = 0.5$.

Thus, there is less reason to study variations in c for large enough N, e.g., $N = 13$.

H2. Cost effect: For any fixed group size and signal precision $(N, x) \in \{3, 7, 13\} \times \{0.7, 1\}$, the frequency of information acquisition, σ^*, remains constant or decreases toward zero as we increase the information acquisition cost from $c = 5$ to $c = 8$ to $c = 25$.

The effect of information cost is straightforward: the higher is the cost of information acquisition, the less likely are individuals to acquire information, a prediction that follows directly from expression (5). In the case where $N = 3$, a cost of $c = 8$ is the highest (integer) value for which all N players rationally acquire information in the symmetric equilibrium of the majority rule voting game. Thus we chose to compare the case of $c = 8$ with the case of $c = 5$. Note that for the cost effect we have data for two cases: 1) $N = 3$ and $x = 0.7$, where we vary c from 5 to 8, and 2) $N = 7$ and $x = 0.7$ where we also vary c from 5 to 8. We further note that there might be some salience issue. For example, theoretically speaking, with $x = 0.7$, an information cost of $c = 8$ should be large enough to dissuade voters from acquiring any information when N is sufficiently large, 7 or higher. Behaviorally speaking, subjects may feel that such a cost level (8 points) is not sufficiently large compared with the benefit level from a correct group decision (100 points), and therefore they may continue to acquire information with a positive frequency. For this reason we also explore a very large cost of $c = 25$ in the $N = 3$ and $x = 0.7$ treatment. This large, $c = 25$, information cost treatment was also added to address the extent to which players perceive they are playing a game with others, as we discuss in further detail later in section 6.4.1.

H3. Signal precision effect: For a given group size and information cost $(N, c) \in \{3, 7, 13\} \times \{5, 8, 25\}$, the frequency of information acquisition can either decrease or increase with an increase in the signal precision from $x = 0.7$ to $x = 1$.

As we increase the signal precision, there are again two effects that work against each other. On the one hand, a more precise signal will induce individuals to invest in information with a higher frequency holding the pivot probability constant. More precisely, fixing p_σ, an increase in x increases the left hand side of the inequality in expression (5) making information acquisition more likely. On the other hand, a better quality of information makes an individual's vote less likely to be pivotal since those who have acquired the more precise signal are now more likely to vote for the correct alternative. More precisely, from expression (3), an increase in x results in an increase in z_σ and from expression (4), an increase in z_σ leads to a decrease in p_σ, which from (5) makes information acquisition less likely. Overall, whether voters acquire information with a higher or lower frequency in equilibrium will depend on which of these two effects is dominant. Note that if subjects are purely decision-theoretic and don't fully understand the strategic interactions associated with the collective decision problem, then only the first effect is at play and the frequency of information acquisition should increase with an increase in the signal precision. However, if subjects do reason game-theoretically, then the equilibrium consequence of a more precise signal is a lower frequency of information acquisition in small groups, e.g., $N = 3$, but a higher frequency of information acquisition in larger groups, e.g., $N = 7$ – see Table 2 for the precise predictions. For the signal precision effect, we have data for two cases: 1) $N = 3$, $c = 8$ and $x = 0.7$ or $x = 1$ and 2) $N = 7$, $c = 8$ and $x = 0.7$ or $x = 1$.

6 Experimental Results

We discuss our experimental findings at both the aggregate and the individual level. We then consider the possibility that subjects coordinated on asymmetric equilibria and we conclude with an analysis of several behavioral explanations for our findings.

6.1 Aggregate Data

Table 3 reports on the average frequencies of information acquisition and informational efficiency in each of the four sessions of each treatment combination (x, N, c) as well as over all sessions of each treatment. Using the data reported in this table, we first address the group size effect, [H1]. Fixing $x = 0.7$ and $c = 0.5$, Table 3 reveals that the mean frequency of information acquisition increases as N is increased from 3 to 7, rising from .695 to .767 but this difference is not statistically significant according to non-parametric Mann-Whitney tests using the session-level averages ($p > 0.10$).[9] By contrast, the theory predicts a movement in the *opposite* direction from a frequency of information acquisition of 1 to .669 as N is increased from 3 to 7. Next, keeping $x = 0.7$ but now fixing c at a higher value, $c = 8$, Table 3 indicates that the mean frequency of information acquisition decreases slightly as N is increased from 3 to 7 to 13, from .582 to .514 to .463, respectively. These differences are again not statistically significant according to pairwise tests ($p > 0.10$). Nevertheless, for the $c = 8$ treatment, the observed decline in the mean frequency of information acquisition as N is increased *is* consistent with the theory, though the magnitude of the acquisition frequencies departs substantially from theoretical predictions: the theory predicts that the frequency of information acquisition declines from 1 when $N = 3$ to 0 when N is 7 or 13.

[9]Unless otherwise noted, in what follows, reported p-values are from non-parametric Mann-Whitney tests using independent session-level averages.

Treatment Conditions	Variable	Session Number 1	2	3	4	Overall Average	Theoretical Prediction
$x = 0.7$, $N = 3$, $c = 5$	$\hat{\sigma}$	0.547	0.760	0.640	0.833	0.695	1.000
	$\hat{w}$	0.620	0.760	0.700	0.660	0.685	0.784
$x = 0.7$, $N = 3$, $c = 8$	$\hat{\sigma}$	0.600	0.353	0.740	0.633	0.582	1.000
	$\hat{w}$	0.680	0.620	0.660	0.620	0.645	0.784
$x = 0.7$, $N = 3$, $c = 25$	$\hat{\sigma}$	0.340	0.087	0.240	0.420	0.272	0.000
	$\hat{w}$	0.540	0.480	0.520	0.700	0.560	0.500
$x = 0.7$, $N = 7$, $c = 5$	$\hat{\sigma}$	0.640	0.826	0.746	0.857	0.767	0.669
	$\hat{w}$	0.760	0.800	0.860	0.840	0.815	0.773
$x = 0.7$, $N = 7$, $c = 8$	$\hat{\sigma}$	0.340	0.751	0.360	0.603	0.514	0.000
	$\hat{w}$	0.580	0.820	0.700	0.740	0.710	0.500
$x = 0.7$, $N = 13$, $c = 8$	$\hat{\sigma}$	0.442	0.617	0.388	0.406	0.463	0.000
	$\hat{w}$	0.700	0.820	0.780	0.740	0.760	0.500
$x = 1$, $N = 3$, $c = 8$	$\hat{\sigma}$	0.887	0.833	0.707	0.833	0.815	0.825
	$\hat{w}$	1.000	1.000	0.960	1.000	0.990	0.978
$x = 1$, $N = 7$, $c = 8$	$\hat{\sigma}$	0.514	0.529	0.646	0.537	0.556	0.447
	$\hat{w}$	0.920	0.940	0.980	0.960	0.950	0.902

Notes: $\hat{\sigma}$ indicates the observed frequency of information acquisition and $\hat{w}$ indicates the observed informational efficiency.

Table 3: Information Acquisition and Efficiency by Treatment: Averages Over All Rounds of Each Experimental Session and Over All Sessions of Each Treatment

Remarkably, as Table 3 further reveals, support for the group size effect is much more clearly evident when signals are perfectly precise, i.e., in the case where $x = 1$. Fixing $x = 1$ and $c = 8$, the mean frequency of information acquisition drops significantly from .815 when $N = 3$ to .556 when $N = 7$ ($p < 0.02$). The theoretical prediction is that the frequency of information acquisition is .825 when $N = 3$ and falls to .447 when $N = 7$. Using a two-sided Wilcoxon sign rank test on session-level observations we cannot reject the null hypothesis that the frequency of information acquisition in the $x = 1$, $N = 3$, $c = 8$ treatment differs from the theoretical prediction ($p = 0.46$), though we can marginally reject this same null hypothesis for the corresponding $N = 7$ treatment ($p = 0.07$). Hence, in the $x = 1$ case, not only do the experimental data reflect the group size effect, but in addition, the mean frequencies of information acquisition are close to or insignificantly different from the theoretical predictions for both group sizes $N = 3$ and $N = 7$. This finding may obtain because when $x = 1$ we have strictly interior predictions for σ^* whereas when $x = 0.7$, we have mostly boundary predictions, either 0 or 1.[10] In addition or alternatively, the elimination of noise in the signal may have enabled subjects to understand the free-riding effect more clearly, a conjecture that we address more fully in the next section.

At this stage, it is also instructive to look at how decision accuracy correlates with group size in the experimental data and compare that with theoretical predictions. For noisy signals ($x = 0.7$) and $c = 5$ as N increases from 3 to 7, Nash equilibrium predicts a marginal drop in efficiency from 78% to 77%. However, the experimental data show a statistically significant ($p < 0.03$) increase in efficiency from 68.5% to 81.5%. For $x = 0.7$ and $c = 8$, efficiency is supposed to be 78.4% at $N = 3$ and then drop to the minimal level (50%) for $N = 7$ and $N = 13$. However, in the experimental data, the efficiency increases from 64.5% to 71% to 76% as N increases from 3 to 7 to 13. Moreover, the change in efficiency from $N = 3$ to $N = 13$ is statistically significant ($p < 0.03$).[11] When signals are noisy, the free riding effect is weak enough that in general, it is dominated by the information aggregation effect in sharp contrast to theoretical predictions. On the other hand, when signals are perfectly informative, free riding is strong enough in the laboratory that efficiency drops with the group size in accordance with theoretical predictions. For $x = 1$ and $c = 8$, as N increases from 3 to 7, efficiency drops from 99% to 95%, which is very close to the theoretically predicted drop from 97.8% to 90%, and this drop is moderately significant ($p < 0.06$).

We next turn to the information cost effect [H2]. Fixing $x = 0.7$ and $N = 3$, Table 3 reveals that an increase in the cost of acquiring information from $c = 5$ to 8 results in a decrease in the frequency of information acquisition from .695 to .582, but this decrease is not statistically significant ($p > 0.10$). However, as we further increase the cost to $c = 25$ holding $x = 0.7$ and $N = 3$ constant, the frequency of information purchase drops more dramatically to .272 and this drop is statistically significant ($p < 0.03$ in the comparison of $c = 8$ vs. $c = 25$). Fixing $x = 0.7$, and $N = 7$, an increase in the cost of acquiring information from $c = 5$ to 8 results in a decrease in the frequency of information acquisition from .767 to .514 - theory predicts a fall from .669 to 0 - and this decrease is marginally significant ($p = 0.08$).

Finally, we consider the signal precision effect [H3]. Fixing $N = 3$ and $c = 8$, Table 3 reveals that an increase in the signal precision from $x = 0.7$ to 1 results in an increase in the mean frequency of information acquisition from .582 to .815 and this difference is statistically significant ($p = 0.04$). The theoretical prediction, by contrast, is for a decrease from 1 to .825. On the other hand, fixing $N = 7$ and $c = 8$, an increase in the signal precision from $x = 0.7$ to 1 results in a slight increase in the frequency of information acquisition from .514 to .556. This difference is not

[10] These boundary predictions cannot be assessed using simple statistical tests as the boundary frequency values of 0 or 1 can never be exceeded.

[11] However, the differences in efficiency from $N = 3$ to $N = 7$ and from $N = 7$ to $N = 13$ are not statistically significant ($p \geq 0.10$ for both comparisons).

statistically significant ($p > 0.10$). Still, the observed increase *is* consistent with the theoretical prediction, which calls for an increase in information acquisition from 0 to .447 as x is increased from 0.7 to 1. While play is close to Nash equilibrium predictions when $x = 1$, behavior is far away from equilibrium predictions when $x = 0.7$ (there is under-acquisition of information in the $N = 3$ case and over-acquisition of information in the $N = 7$ case), and it is the latter deviations that are responsible for the observed deviations from the signal precision effect (H3).

Treatment Conditions	Variable	1st 13 rds		2nd 12 rds	Overall	Prediction
$x = 0.7,\ N = 3,\ c = 5$	$\hat{\sigma}$	0.715	$>$	0.674	0.695	1.000
	$\hat{w}$	0.664	$<$	0.708	0.685	0.784
$x = 0.7,\ N = 3,\ c = 8$	$\hat{\sigma}$	0.590	$>$	0.573	0.582	1.000
	$\hat{w}$	0.606	$<^{**}$	0.688	0.645	0.784
$x = 0.7,\ N = 3,\ c = 25$	$\hat{\sigma}$	0.276	$>$	0.267	0.272	0.000
	$\hat{w}$	0.519	$<$	0.604	0.560	0.500
$x = 0.7,\ N = 7,\ c = 5$	$\hat{\sigma}$	0.768	$>$	0.766	0.767	0.669
	$\hat{w}$	0.779	$<$	0.854	0.815	0.773
$x = 0.7,\ N = 7,\ c = 8$	$\hat{\sigma}$	0.523	$>$	0.503	0.514	0.000
	$\hat{w}$	0.692	$<$	0.729	0.710	0.500
$x = 0.7,\ N = 13,\ c = 8$	$\hat{\sigma}$	0.473	$>$	0.452	0.463	0.000
	$\hat{w}$	0.789	$>$	0.729	0.760	0.500
$x = 1,\ N = 3,\ c = 8$	$\hat{\sigma}$	0.817	$>$	0.813	0.815	0.825
	$\hat{w}$	1.000	$>$	0.979	0.990	0.978
$x = 1,\ N = 7,\ c = 8$	$\hat{\sigma}$	0.584	$>$	0.527	0.556	0.447
	$\hat{w}$	0.952	$>$	0.948	0.950	0.902

Notes: $\hat{\sigma}$ indicates the observed frequency of information acquisition and $\hat{w}$ indicates observed (informational) efficiency. $***$, $**$, and $*$ indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 4: Information Acquisition and Efficiency Over Time. Averages from All Sessions of a Given Treatment: First 13 Rounds, Second 12 Rounds and Overall (All 25 rounds)

Table 4 reports on the average frequencies of information acquisition and efficiency over the first 13 rounds versus the final 12 rounds of each treatment (using data from all sessions) in an effort to assess whether there is any evidence for learning over time. In particular, we examined whether there was a significant upward, "<", or downward, ">" trend in the overall mean frequencies of information acquisition or informational efficiency in the first 13 rounds as compared with the last 12 rounds. As Table 4 reveals, the evidence for learning is weak; there is a very slight decrease in the mean frequencies of information acquisition ($\hat{\sigma}$) from the first to the second half of *all* sessions. Sometimes this decrease is in the direction of equilibrium prediction, i.e., in treatments $(x, N, c) \in \{(0.7, 3, 25), (0.7, 7, 5), (0.7, 7, 8), (0.7, 13, 8), (1, 7, 8)\}$ while in other instances it is not, i.e., in treatments $(x, N, c) \in \{(0.7, 3, 5), (0.7, 3, 8), (1, 3, 8)\}$. However, none of these decreases in the frequency of information acquisition are statistically significant using Mann Whitney tests on session-level averages. Despite the slight decrease in the mean frequency of information acquisition, the mean level of informational efficiency is, in most treatments, increasing slightly from the first to the second half of sessions, though with a single exception, these trends are also statistically insignificant. Overall, we conclude that the evidence for learning is weak.

Summarizing, using session level means, we find that the experimental evidence departs from the precise point predictions of the theory and there is mixed support for the theory's comparative statics predictions as identified in our Hypotheses H1-H3. The theory appears to perform best in terms of the match between point predictions and the experimental data in the case where $x = 1$ so that signals are perfectly informative. We next turn to exploring individual subject behavior in some detail so as to determine whether our aggregate data analysis (using session-level means) may be masking any larger behavioral differences across treatment conditions.

6.2 Individual Behavior

Figure 1 shows cumulative distributions of the frequency of information acquisition over all 25 rounds of our experiment using pooled data from various combinations of treatments where the signal precision is fixed at $x = 0.7$. Figure **??** shows the same type of cumulative frequency distributions for various combinations of treatments where the signal precision is either $x = 0.7$ or $x = 1$. In these figures, the left-most intercept of the cumulative frequency indicates the percentage of subjects who *never* chose to purchase information, 0 of 25 rounds.

Consider first the case where the signal precision is fixed at $x = 0.7$. When $c = 5$, the upper left panel of Figure 1 reveals that the cumulative frequency of information purchase when $N = 7$ stochastically dominates the cumulative frequency of information purchase when $N = 3$, which is completely opposite to theoretical predictions. By contrast, when $c = 8$, as shown in the upper right panel of Figure 1, an increase in group size largely follows the comparative statics prediction that information acquisition decreases as the group size gets larger. Indeed, for this case, the cumulative frequency of information purchase when $N = 3$ stochastically dominates the cumulative frequency of information purchase when $N = 13$. The bottom two panels of Figure 1 confirm that the individual distributions follow the comparative statics prediction that, holding N and x constant, increases in information cost, c, are associated with less information acquisition for groups of size $N = 3$ and $N = 7$ respectively (we

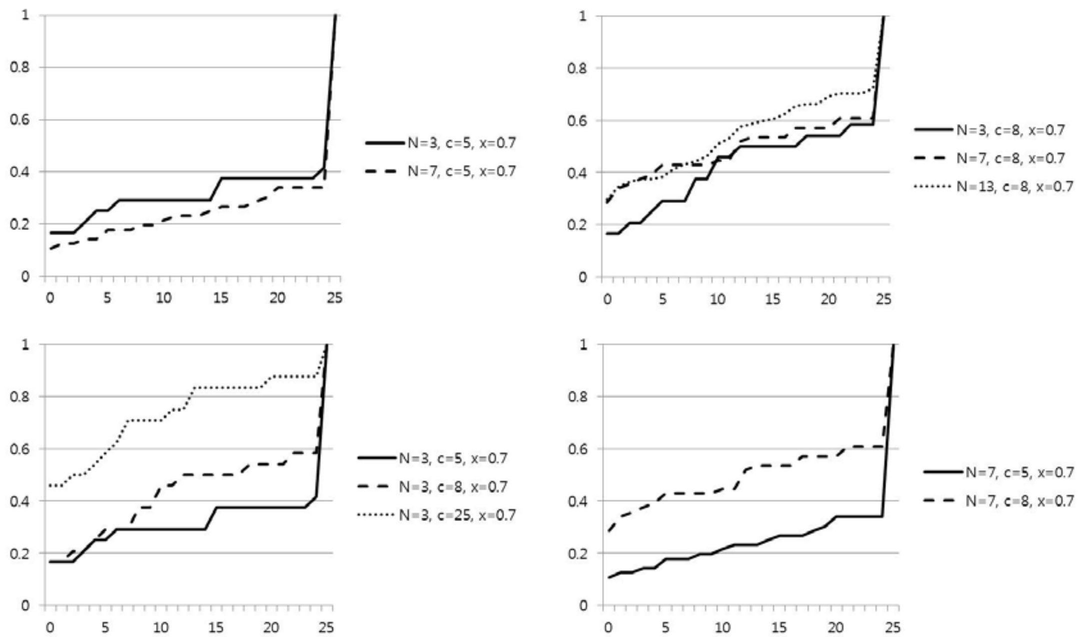

Figure 1: Distribution of the Individual Frequencies of Information Acquisition over All 25 Rounds, $x = 0.7$

administered only one cost level $c = 8$ for the larger group size $N = 13$). Indeed, for the $N = 7$ case we see clearly that the cumulative frequency of information acquisition when $c = 5$ stochastically dominates the cumulative frequency of information acquisition when $c = 8$.

The cumulative frequency distributions shown in Figure **??** enable us to examine the effect of changes in the signal precision and/or the group size on information acquisition. Here we fix $c = 8$ as this is the only cost that we considered in treatments where $x = 1$. We find that, consistent with the theory, when $x = 1$ there is more information acquisition when $N = 3$ than when $N = 7$, as seen in the upper right panel of Figure **??**. However, we also find that the comparative statics predictions of the theory concerning variations in the signal precision are not clearly found in our data for either groups of size $N = 3$ or $N = 7$. As the upper and bottom left panels of Figure **??** show, the frequency of information acquisition generally increases as we increase the level of the signal precision from $x = 0.7$ to $x = 1$, though this is more clearly evident in the case where $N = 3$ than in the case where $N = 7$. However, according to the theory, an increase in signal precision sometimes implies a decrease in the equilibrium frequency of information acquisition (e.g., from 1.00 to 0.825 when $N = 3$ and $c = 8$) and sometimes an increase (e.g., from 0.00 to 0.447 when $N = 7$ and $c = 8$), again as a consequence of competition between the

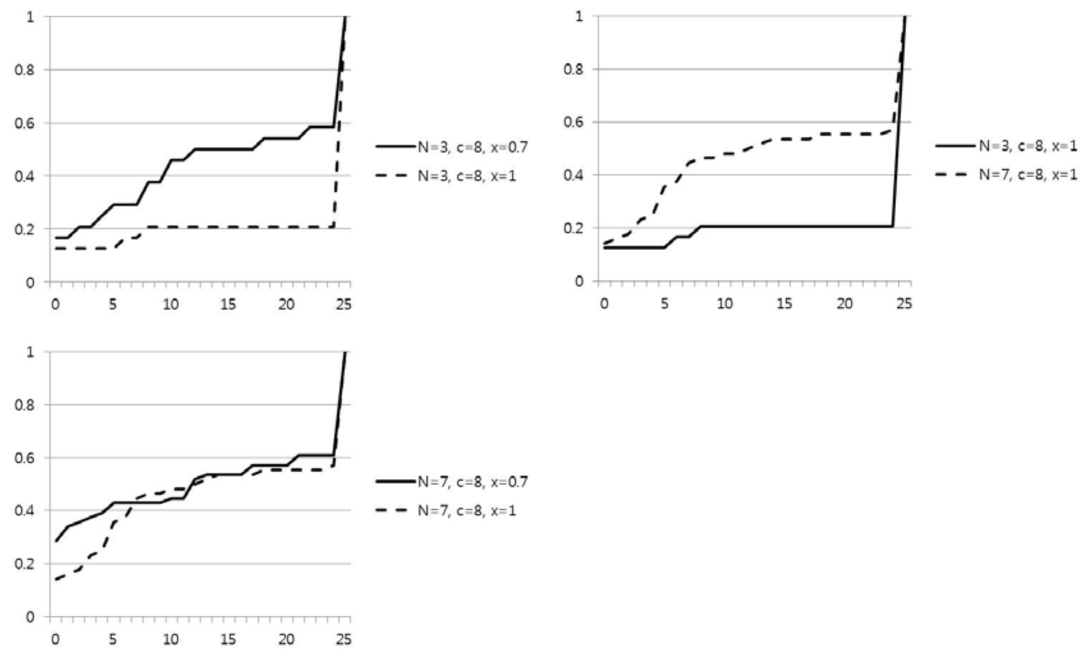

Figure 2: Comparison of Individual Distributions across Different Signal Precisions

information aggregation effect and the free-riding effect. Our finding that increases in signal precision generally lead to greater frequencies of information acquisition (for fixed N and c) suggests that our data may be better explained by decision-theoretic (as opposed to game-theoretic) considerations. From a decision-theoretic perspective, if N and c are held constant, then as the quality of information improves, it becomes more desirable to acquire such information. The proportion of subjects who behave according to this decision-theoretic approach should be large enough to sway the overall results in their favor while the game-theoretic reasoning is so subtle here that strategic (game-theoretic) subjects may fail to grasp such incentives. We will address this issue further in section 6.4.1.

In addition to non-parametric tests, we have also conducted a parametric analysis of individual information acquisition decisions. Specifically, we report on a random effects probit regression of subjects' binary decision to buy or not buy information (buy=1, not buy=0) in every round of our experiment using pooled data from all sessions of all treatments with standard errors clustered at the session level. Table 5 reports on results from several different regression model specifications.

In these regressions, "perfect" is a dummy variable set to 1 if $x = 1$ and 0 otherwise ($x = 0.7$); "round" is the round number, 1,2,...,25; N is the group size, c is the information cost; "successlag1" is a dummy variable set to 1 if the subject's group correctly guessed the jar assigned to it in the last round and 0 otherwise; "pctuniflag1" denotes the percent of the subject's group who did not buy information (were uninformed) last round (recall that this information was revealed at to all group members at the end of each round); "perfXpctuniflag1" is an interaction variable, perfect × pctuniflag1; "perXsuccesslag1" is another interaction variable, perfect × successlag1; "succXpctuniflag1" is another interaction variable, successlag1 × pctuniflag1; and "perfXsuccXpctuniflag1" is a three-way interaction variable, perfect × successlag1 × pctuniflag1. Note that the baseline model, where perfect = 0, is the $x = .7$ case.

Dependent var: buy info	(1)	(2)	(3)	(4)
constant	5.316***	4.630***	4.791***	4.860***
	(10.76)	(9.35)	(9.45)	(9.63)
perfect	0.914*	0.881	0.678	4.336***
	(2.50)	(1.82)	(1.40)	(3.33)
round	-0.019***	-0.019***	-0.019***	-0.019***
	(-5.53)	(-5.49)	(-5.47)	(-5.61)
N	-0.229***	-0.208***	-0.221***	-0.220***
	(-5.34)	(-5.04)	(-5.11)	(-5.27)
c	-0.269***	-0.194***	-0.239***	-0.250***
	(-8.91)	(-6.89)	(-8.20)	(-8.47)
successlag1	-0.245***	-0.205**	0.212	0.290
	(-3.99)	(-3.24)	(1.37)	(1.86)
pctuniflag1	-0.238	-0.496**	0.0166	0.115
	(-1.52)	(-2.77)	(0.07)	(0.46)
succXpctuniflag1			-0.799**	-0.950***
			(-2.96)	(-3.46)
perfXsuccesslag1		-0.495*	-0.312	-4.077**
		(-2.00)	(-1.23)	(-3.22)
perfXpctuniflag1		0.795*	1.042**	-3.747*
		(2.09)	(2.68)	(-2.29)
perfXsuccXpctuniflag1				5.077**
				(3.03)
Observations	9200	9200	9200	9200

Notes: t statistics in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 5: Random Effects Probit Analysis of Buy Decisions, Data From All Sessions

Table 5 reveals that across all four specifications, increases in N and c yield statistically significant *decreases* in the probability of buying information, which is consistent with the group size effect, hypothesis H1, and the cost effect, hypothesis H2. While we did not find such strong evidence for these hypotheses in the treatment-by-treatment pairwise comparisons using non-parametric tests (especially in the case of noisy signals) as reported on in the previous section, using the pooled data set and a random effects probit regression analysis, we find stronger qualitative support for these two hypotheses.

Consider next the effect of the two feedback variables, successlag1 and pctuniflag1; theory is silent on the effect of these lagged variables on decisions to acquire information since the model being tested is a one-shot, static model. We include these two feedback variables because they serve as proxies for incentives to free ride. The number of uninformed voters in the last round directly reveals the extent of free riding in that round. Whether a player's group achieved success last round is also an inducement to free ride for an individual in the sense that, if the group members play similarly this round, the outcome is again likely to be a success even if the individual free rides (so long as the individual was not pivotal in the last round). Overall, as revealed in regression specifications 1-2, we find that individuals are less likely to buy information this round if they were 1) in a group that experienced a success last round (successlag1=1) or 2) the greater was the percentage of group members who did not buy information and were thus uninformed last round, indicating that these proxies for free riding incentives have some predictive power. However, as the last two regression specifications in Table 5 make clear, the impact of these lagged outcome variables on information acquisition decisions clearly depends on whether interaction effects are allowed and whether information was perfect or not. In specifications 3-4, for the baseline case of imperfect information (perfect=0 or $x = 0.7$), the coefficient estimates on successlag1 and pctuniflag1 are not significantly different from zero; only the interaction variable, succXpctuniflag1 leads to a significant decrease in the likelihood of buying information. However, allowing for all possible interaction effects as in specification 4, we see that in the case of perfect information, both the lagged success rate and the lagged percentage of uninformed voters work to depress the likelihood of buying information, as evidenced by the significantly negative coefficients on the perfXsuccesslag1 and perfXpctuniflag1 interaction variables. We interpret these differences as indicating that incentives for free riding are much stronger for subjects in the perfect information case as compared with the imperfect information case.[12] These findings provide some explanation for why behavior in the perfect information

[12]The significantly positive coefficient on the three–way interaction variable, perfXsuccXpctuniflag1, is a little puzzling. However, one interpretation may be that players in the perfect information treatment are willing to free ride up to some point, but if they experience a success with a low percentage of informed voters, they rightly attribute that outcome to luck, and increase their willingness to buy information in the subsequent round.

case was found to be closer to the symmetric Nash equilibrium predictions than in the imperfect information case.

6.3 Asymmetric Equilibria

Thus far, we have considered the extent to which our experimental data might be rationalized by the unique *symmetric* equilibrium of the theory as described in section 3.1, i.e., by an equilibrium in which two voters with the same private information use the *same* strategy. This restriction is justified by the fact that coordinating on an asymmetric profile of strategies should be difficult in our experiment given our random and anonymous matching of subjects into groups of size N in every round.

In this section we consider a different natural restriction on the set of equilibrium strategies that leads to asymmetric voting strategies in some cases. Specifically, we now require equilibrium strategies to satisfy two criteria : first, each subject's strategy with respect to information acquisition should be pure, i.e., $\sigma_i^* \in \{0, 1\}$, and second, the voting part of their strategy should be symmetric across alternatives. In this setting, it is easy to establish that those who acquire information ($\sigma_i^* = 1$) must vote according to their signal, while those who do not acquire information ($\sigma_i^* = 0$) must vote randomly. Thus, equilibrium under this new set of restrictions is simply characterized by the number of voters who acquire information and vote sincerely represented by $k^* \in \{0, 1, ...N\}$. The other $N - k^*$ voters set $\sigma^* = 0$ and vote randomly. If $k^* = 0$ or $k^* = N$, the equilibrium strategies are type-symmetric, but if $0 < k^* < N$, we have an *asymmetric* equilibrium.[13]

In order to characterize the (potentially asymmetric) equilibrium with a pure information acquisition strategy, consider the profile where k out of N voters acquire information (and vote sincerely) and the remaining $N - k$ do not (and vote randomly) and denote the probability of a tie under such a profile by $p(k, N)$.[14]

Remark 1 *In the environment (x, N, c), there is an equilibrium where k^* voters acquire information (and vote sincerely) and $N - k^*$ do not (and vote randomly) if and*

[13] We consider only asymmetric equilibria of this type. A complete characterization of asymmetric equilibria for our model is beyond the scope of this paper.

[14] The formula for $p(k, N)$ is:

$$p(k,N) = \begin{cases} \sum_{t=0}^{\frac{N-1}{2}} \left[\binom{k}{t} x^t (1-x)^{k-t}\right] \left[\binom{N-k}{\frac{N-1}{2}-t} \left(\frac{1}{2}\right)^{N-k}\right] & \text{if } 0 < x < 1, \\ \binom{N-k}{\frac{N-1}{2}-k} \left(\frac{1}{2}\right)^{N-k} & \text{if } x = 1. \end{cases}$$

$$\text{where } \binom{a}{b} = 0 \text{ if } a < b$$

Notice that for a fixed N, $p(k, N)$ is strictly decreasing in k.

only if

$$p(k^* - 1, N - 1) \geq \frac{c}{M\left(x - \frac{1}{2}\right)} \geq p(k^*, N - 1) \qquad (7)$$

Moreover, the value of k^* that satisfies the above condition is unique.

See Appendix B for a proof.

In environments where the symmetric equilibrium value for σ^* is strictly interior, i.e., $\sigma^* \in (0, 1)$, pure information acquisition strategies induce *asymmetric* equilibria with the property that $\frac{k^*}{N}$ approximates (but can differ from) σ^*.[15] To explore whether these asymmetric equilibria might better explain our experimental data, we focus on the three cases where there are interior symmetric equilibria $(x, N, c) = \{(.7, 7, 5), (1, 3, 8), (1, 7, 8)\}$ where the ratios k^*/N are: $5/7 = .71$, $2/3 = 0.67$ and $3/7 = 0.43$, respectively, and the corresponding values for σ^* are 0.67, 0.83, and .045, respectively. It is interesting to note that the equilibrium k^* is that value of k which produces a search frequency $\frac{k^*}{N}$ closest to σ^* in every case. We first verified that in these three cases, the uninformed are voting randomly while the informed are voting sincerely according to their private signals.[16]

We next perform a simple check of whether voters best respond to some proxy for the threshold value, k^*, the number of voters who are expected to acquire information in an asymmetric equilibrium. The proxy we use is the number of voters among the other $N-1$, who had acquired information in the last round, k_{t-1}^{-i}. In our experiment, subjects received feedback at the end of each round as to how many in their group of size N acquired information in that round. Thus, if k_{t-1}^{-i} voters acquired information last round and are again expected to do so in the current round, then a subject's best response is to buy information if $k_{t-1}^{-i} < k^*$ and not buy information if $k_{t-1}^{-i} \geq k^*$. One would therefore expect the probability of information acquisition to be decreasing in $k_{t-1}^{-i} - k^*$. We test this prediction for the three environments where we have interior solutions for k^* using again a random effects probit regression analysis of individual decisions to buy information with standard errors clustered on session level observations. Here, each regression uses only data from one of the three treatments with interior asymmetric equilibria. The only explanatory variable, in addition to a constant term, is the difference variable, $k_{t-1}^{-i} - k^*$, where k^* is determined by treatment conditions: $k^* = 5, 2, 3$ for the (x, N, c) treatments $(.7, 7, 5)$, $(1, 3, 8)$, $(1, 7, 8)$, respectively. The results are reported in Table 6.

[15] In environments where the symmetric equilibrium values for σ^* are corner solutions, i.e., $\sigma^* \in \{0, 1\}$, k^*/N will also be symmetric and correspond to the same corner solution (as in 5 of our 8 treatments), so we don't focus on these cases.

[16] Using data from all sessions of all treatments, we find that the null hypothesis that uninformed voters randomize between R and B cannot be rejected ($p > .05$). Among informed voters, sincere voting (voting according to the signal received) is close to or equals 100 percent in each treatment of our experimental treatments.

Dependent var: buy info	(1)	(2)	(3)
$(x, N, c) =$	$(.7, 7, 5)$	$(1, 3, 8)$	$(1, 7, 8)$
Constant	2.663***	1.646***	2.764**
	(7.21)	(3.65)	(2.98)
$k_{t-1}^{-i} - 5$	-0.0987		
	(-1.09)		
$k_{t-1}^{-i} - 2$		-0.597	
		(-1.09)	
$k_{t-1}^{-i} - 3$			-0.129*
			(-2.54)
Observations	1400	600	1400

Notes: t statistics in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

Table 6: Regression Evidence on Best Response Play According to Asymmetric Equilibrium Predictions. Three Treatments with Interior Asymmetric Equilibria.

The regression results reported in Table 6 confirm the negative relationship between the difference variable $k_{t-1}^{-i} - k^*$ and information acquisition decisions in each of the three treatments. However, the coefficient on this difference variable is only significantly negative for the treatment $(x, N, c) = (1, 7, 8)$ (column 3) involving perfect signals, where the difference between $\frac{k^*}{N}$ and the symmetric equilibrium prediction σ^* is the smallest of the three cases examined. Given this very small difference, it seems likely that the significantly negative coefficient in this case simply reflects players playing a best response to the *symmetric* equilibrium prediction, especially since we have already found evidence that behavior in the $x = 1$ treatments is very close to symmetric equilibrium predictions. We conclude that there is not much support for the notion that our data are better rationalized by this class of asymmetric equilibria, which, for most of our treatments, coincide with the unique symmetric equilibrium. Having considered the explanatory power of both symmetric and asymmetric equilibrium predictions, we next turn to examining some behavioral models in an attempt to further rationalize our experimental data.

6.4 Behavioral Models

As we have seen, in some of our experimental treatments subjects under-acquire information relative to the symmetric Nash equilibrium prediction, e.g., the case where $N = 3$, $c = 8$ and $x = 0.7$. On the other hand, we often see that with a single change of a treatment variable we move from under- to over-acquisition of information as for example in the case where $N = 7$, $c = 8$ and $x = 0.7$. In this section we present several possible explanations for the observed over- or under-acquisition of information in our experimental data.

6.4.1 Decision-theoretic rather than strategic thinking

Suppose that subjects under-weight or dismiss completely the strategic interaction that is involved in the collective action voting game. As an extreme case, let us suppose that subjects perceive the game to be one where $N = 1$ and so in effect, they are lone decision-makers and thus they always view themselves as being pivotal.[17] If $N = 1$, then it is rational to acquire information at the fixed cost c so long as $M(x - 1/2) \geq c$ (this follows from expression (5) with $p_\sigma = 1$) and to not acquire information otherwise. In our parameterization, we have $M = 100$. Thus for our $x = 0.7$ treatment, it becomes rational to acquire information if $c \leq 20$, while for our $x = 1$ treatment it is rational to acquire information so long as $c \leq 50$.[18] These cost thresholds are satisfied for *all* of our treatments, with the sole exception of the $x = 0.7$, $c = 25$ treatment, and indeed, that is why we chose to implement that particular treatment. In that treatment, both the decision-theoretic and game-theoretic incentives are perfectly aligned and so one might expect that subjects would *never* acquire information in that setting. Note that while the characterization of subjects as decision theorists can explain *over*–acquisition of information in our $x = 0.7$ treatments with $c < 20$ it cannot explain *under*–acquisition of information as in the $N = 3$, $x = 0.7$ treatments where $c = 5$ or $c = 8$. More generally we note that over- or under-acquisition of information may be an unavoidable finding in settings where the equilibrium point predictions of the theory imply, respectively, 0 and 100 percent frequencies of information acquisition, as is indeed the case in several of our treatments.

[17]This assumption finds some support in Duffy and Tavits (2008) who elicited the beliefs of voters prior to their participation in an experimental voting game and found that many voters greatly overestimated the pivotality of their voting choice, though such miss-perceptions tended to decrease with experience.

[18]Hence for fixed c, information acquisition becomes more attractive as x increases, as discussed at the end of section 6.2.

Treatment	Type	1st 13 rds		2nd 12 rds	Overall
$x = 0.7$, $N = 3$, $c = 5$	NB	0.167	$<$	0.250	0.167
	S	0.250	$>$	0.125	0.250
	AB	0.583	$<$	0.625	0.583
$x = 0.7$, $N = 3$, $c = 8$	NB	0.167	$<$	0.292	0.167
	S	0.416	$>^*$	0.250	0.416
	AB	0.417	$<$	0.458	0.417
$x = 0.7$, $N = 3$, $c = 25$	NB	0.500	$<$	0.542	0.458
	S	0.375	$>$	0.292	0.417
	AB	0.125	$<$	0.166	0.125
$x = 0.7$, $N = 7$, $c = 5$	NB	0.107	$<$	0.143	0.107
	S	0.232	$>$	0.161	0.232
	AB	0.661	$<$	0.696	0.661
$x = 0.7$, $N = 7$, $c = 8$	NB	0.286	$<$	0.393	0.286
	S	0.321	$>$	0.178	0.321
	AB	0.393	$<$	0.429	0.393
$x = 0.7$, $N = 13$, $c = 8$	NB	0.317	$<$	0.366	0.298
	S	0.375	$>$	0.317	0.423
	AB	0.308	$<$	0.317	0.279
$x = 1$, $N = 3$, $c = 8$	NB	0.125	$=$	0.125	0.125
	S	0.083	$=$	0.083	0.083
	AB	0.792	$=$	0.792	0.792
$x = 1$, $N = 7$, $c = 8$	NB	0.179	$<$	0.250	0.143
	S	0.393	$>$	0.304	0.428
	AB	0.428	$<$	0.446	0.429

Notes: NB refers to subjects who *never buy* information over the sample period, S, to subjects who *switch* between buying and not buying information (at least once) over the sample period, and AB, to subjects who *always buy* information over the sample period. $***$, $**$, and $*$ indicate significance at the 1%, 5%, and 10% levels, respectively.

Table 7: Proportions of Different Subject Types Over Time: Averages from All Sessions of a Given Treatment: First 13 Rounds, Second 12 Rounds and Overall (All 25 rounds)

In an effort to address the extent to which subjects might be ignoring strategic considerations and acting as decision-theorists, we classified each subject based on their information acquisition decisions. Specifically, we classified each subject according to one of three distinct types: 1) those who *never buy* (NB) information; 2) those who *switch* (S) at least once between buying and not buying information; and 3) those who *always buy* (AB) information. Table 7 shows the proportion of subjects who fall into these three classifications for each treatment condition (x, N, c) over all 25 rounds (Overall) as well as for the first 13 and last 12 rounds. Using the latter two measures, we examined whether there was an trend upward, "$<$", downward "$>$", or no change, "$=$", in the proportions of a given type from the first to the second half of a session and we further tested the significance of any trend changes using Mann Whitney tests on session-level averages.

A general observation from Table 7 is that the proportion of switching (S) types declines as subjects gain experience, while the proportion of pure strategy, NB and AB types grows over time though these differences are, with a single exception, not statistically significant. Further, the growth in NB types is generally greater than the growth in AB types, and this finding explains why information acquisition declines slightly from the first to the second half of sessions as reported earlier in Table 4.

Regarding comparative statics predictions, let us first focus on the $x = 0.7$, $c = 8$ treatments. In this case, the share of AB-types in the population steadily decreases from 41.7 percent to 39.3 percent to 27.9 percent as we increase N from 3 to 7 to 13, respectively. The declines from N=3 to N=7 or from N=7 to N=13 are not statistically significant according to Mann-Whitney tests on session level data, ($p > 0.10$ for both tests). The game-theoretic equilibrium prediction is for 100 percent AB-types when $N = 3$ and a drop-off to 0 percent AB-types (and 100 percent NB-types) when $N = 7$ or $N = 13$. By contrast, the decision-theoretic prediction is for 100 percent AB-types in all three of these treatments. The steady but more gradual decline in AB-types as N is increased as reported in Table 7 suggests that decision costs, as opposed to strategic, group-size considerations alone may be playing a role in the behavior of some of our AB-type subjects.

Consider next the case where $x = 0.7$ and $N = 7$ and c is varied from 5 to 8. The game-theoretic equilibrium prediction is that subjects should acquire information on average 66.9 percent of the time when $c = 5$, but should never acquire information when $c = 8$. By contrast, the decision-theoretic prediction again calls for 100 percent AB-types in both of these treatments as c is always less than 20. Table 7 reveals that there is indeed a much larger percentage of AB-types when $c = 5$ (66.1 percent) than when $c = 8$ (39.3 percent) and this difference is statistically significant ($p < 0.05$) though the percentage of AB-types remains strictly greater than 0 in violation of game-theoretic equilibrium prediction but consistent with the notion that some subjects may be acting as decision-theorists.

Similarly, in the case where $x = 0.7$ and $N = 3$, we see a steady decline in the frequency of AB-types as the cost, c, increases from 5 to 8 to 25, however, only the latter decline in AB types as c increases from 8 to 25 is statistically significant ($p = 0.052$).

This finding on type differences helps us to understand an earlier finding in connection with Table 3, where fixing $x = 0.7$ and $N = 3$, we found a statistically insignificant decrease in the frequency of information acquisition as c was increased from 5 to 8 but a statistically significant drop in the frequency of information acquisition as c was further increased from 8 to 25. However, these results remain inconsistent with the pure game-theoretic prediction of 100 percent AB-types when $c = 5$ or $c = 8$ and a decline to 0 percent AB-types when $c = 25$. Note that in the last case, the decision-theoretic model predictions coincide with the game-theoretic equilibrium predictions, and so there are also inconsistencies with using the decision-theoretic approach to characterize subject behavior; in particular, when $c = 25$, the observed frequency of AB types is *not* 0, as both the game theoretic equilibrium and decision-theoretic models predict.

Finally, consider the case where $x = 1$ and $c = 8$. In this case, the game–theoretic equilibrium predictions are closer to matching the distribution of subject types than the decision-theoretic predictions. In particular, when $N = 3$, the game-theoretic equilibrium prediction is for 82.5 percent of subjects to acquire information, while when $N = 7$ the prediction is for 44.7 percent of subjects to acquire information. The decision-theoretic prediction is for all subjects to acquire information in both of these treatments as c is always less than 50. Table 7 reveals that the frequency of AB-types falls from 79.2 percent when $N = 3$ to 42.9 percent when $N = 7$ (a statistically significant decrease, $p = 0.017$) instead of remaining constant at 100 percent as would be consistent with the decision-theoretic approach.

Summarizing, the evidence on individual behavior suggests that when $x < 1$, the player population could be characterized as a mixture of game-theoretic and decision-theoretic player types; decision-theoretic reasoning can account for over-acquisition of information in all but one of our treatments (the one where $c = 25$), though not *under*–acquisition of information as is often observed in our treatments where $N = 3$. By contrast, when $x = 1$, the distribution of player types is more closely aligned with game-theoretic equilibrium predictions as opposed to decision-theoretic predictions. The latter finding suggests that subjects may compensate for the greater noise in the imperfect signal ($x = 0.7$) treatments and the associated strategic uncertainty about the information held by others by ignoring strategic considerations altogether and acting more like decision-theorists.

6.4.2 Quantal response equilibrium

A second possible explanation for why the frequency of information acquisition is at odds with theoretical predictions is that the experimental environment in which voters are operating is a noisy one and so the appropriate way to model behavior is using a noisy best response function that conditions on the actual distribution of subjects' decisions and allows subjects to make mistakes. The idea of finding equilibria that comprise mutual best responses to the empirical distribution of actual and possibly noisy behavior, as opposed to the theoretical ideal has been formalized as the concept of a quantal response equilibrium (QRE) by McKelvey and Palfrey (1995). In this section we estimate the QRE predictions for information acquisition using our experimental data and we compare these with the observed frequencies of information acquisition as well as with the Nash equilibrium predictions.

In the QRE model, we calculate the information acquisition choice probabilities as quantal response functions of the expected payoffs. Given the slope λ of the logistic quantal response function, the information acquisition choice strategy of a subject can be written as:

$$\sigma(\lambda) \quad = \quad \frac{1}{1 + \exp[-\lambda\{U(\sigma_1) - U(\sigma_0)\}]} \tag{8}$$

where, as before, σ_1 means "acquire information," while σ_0 means "do not acquire information." Here, λ is understood to measure the "degree of rationality" of the subjects, with $\lambda = 0$ corresponding to random information acquisition choice behavior as in that case, $\sigma(\lambda = 0) = \frac{1}{2}$. As $\lambda \to \infty$, the QRE estimates of $\sigma(\lambda)$ converge to the rational choice predictions of the model.

The likelihood function we maximize is given by:

$$\mathcal{L}(\lambda) = \sigma(\lambda)^{\sigma_1}\,[1 - \sigma(\lambda)]^{\sigma_0} \tag{9}$$

In all instances, we use pooled data from all sessions of a single treatment condition, (x, N, c) in maximizing the above likelihood function. We focus on the $x = 0.7$ treatments only since the noisy signal setting is where we observe the greatest departure of the experimental data from theoretical predictions.[19]

In addition to treatment-by-treatment estimations, we also estimate a quantal response parameter λ_p using pooled data from all sessions of all treatments where $x < 1$. In that case, we maximize the following likelihood function

$$\mathcal{L}(\lambda_p) = \prod_j \sigma(\lambda_p)^{\sigma_1^j}\,[1 - \sigma(\lambda_p)]^{\sigma_0^j} \tag{10}$$

where the index j runs through all treatments; i.e., each index j represents a fixed treatment of (x, N, c). The results of our maximum likelihood estimation are reported in Table 8.

The QRE estimates for the individual treatments, $(\hat{\sigma})$, generally provide a good fit to the observed frequencies of information acquisition (Observed) and thus a means of

[19]QRE estimates for the $x = 1$ case are available from the authors on request.

Precision	Group Size	Cost	$\hat{\lambda}$	$\hat{\sigma}$	$\hat{\lambda}_p$	$\hat{\sigma}_p$	Observed	Predicted
$x = 0.7$	$N = 3$	$c = 5$	9.7	0.694	3.3	0.573	0.695	1
		$c = 8$	11.3	0.582	3.3	0.526	0.582	1
		$c = 25$	3.3	0.269	3.3	0.269	0.272	0
	$N = 7$	$c = 5$	∞	0.669	3.3	0.508	0.767	0.669
		$c = 8$	0	0.500	3.3	0.461	0.514	0
	$N = 13$	$c = 8$	1.7	0.463	3.3	0.431	0.463	0

Notes: $\hat{\lambda}$ and $\hat{\sigma}$ are estimated for each treatment with fixed (x, N, c) while $\hat{\lambda}_p$ and $\hat{\sigma}_p$ are estimated using pooled data from all $x = 0.7$ treatments. The mean squared error (MSE) between the pooled estimates and the experimental data is found to be 0.01479.

Table 8: Quantal Response Equilibrium: Maximum Likelihood Estimates

rationalizing our experimental data as the play of noisy best responses by the subjects in our experiment. Note in particular, that the QRE approach, unlike the view of subjects as decision-theorists, can account for both the over-acquisition of information and the under-acquisition of information that is observed in our experiment; for example the under-acquisition of information in the $N = 3$, $x = 0.7$ treatments where $c = 5$ or $c = 8$. The pooled QRE estimates $\hat{\sigma}_p$ provide a slightly poorer fit to the observed information acquisition choice frequencies but nevertheless continue to capture the observed variations in information acquisition (above and below the equilibrium predictions) across all of our different treatments. We note that the mean squared error (MSE) between the pooled estimates and the observed frequencies of information acquisition is 0.0148, indicating a very good fit.[20] We note further that the pooled rationality parameter estimate, $\hat{\lambda}_p = 3.3$, indicates a rather low level of overall game theoretic rationality by the subjects in our experiment.

6.4.3 Risk aversion

A third possible explanation for why information acquisition decisions are at odds with theoretical predictions is that we have assumed that agents are risk neutral with regard to uncertain money earnings. This assumption can be relaxed by allowing agents to be risk averse with respect to uncertain monetary payoffs. In our context,

[20]The MSE is calculated by summing the squared errors between observed information acquisition frequencies and the QRE estimates using the pooled data over all treatments, and then dividing the resulting sum of squares by the number of included treatments; $MSE = \sum_j (\sigma_o - \hat{\sigma}_p)^2/6$, where j represents each treatment with fixed (x, N, c), σ_o is an observation in our data, and $\hat{\sigma}_p$ is an estimate from the pooled data from all treatments.

risk aversion is equivalent to the marginal cost of information acquisition depending on the voting outcome: the subject perceives this marginal cost to be lower if the group decision is correct.

For convenience of exposition, in this section we assume that a voter's monetary payoff, m, takes the value 1 or 0 if she acquires information (depending on whether the group decision is correct or incorrect), and respectively, takes the value $1 + c$ or c if she does not. We further assume that the money amount, m, earned by a subject leads to a utility, $u(m)$. The risk-neutral case which has been studied in the main section of the paper is the special case where $u(m) = m$. We model risk aversion with the assumption that $u(\cdot)$ satisfies $u' > 0$ and $u'' < 0$ for all $m > 0$. To capture our present parameterization, we assume that the utility function is CARA, i.e.,

$$u(m) = \frac{1 - \exp(-\alpha m)}{1 - \exp(-\alpha)},$$

where $\alpha > 0$ is the parameter that captures the extent of risk aversion. In the limit, the $\alpha = 0$ case captures risk neutral behavior.[21] We now turn to the question of how the optimal behavior of a risk-averse subject may differ from that of a risk neutral one.

As before, a strategy consists of two elements: the probability of information acquisition, σ, and the probability of sincere voting v_s conditional on obtaining signal $s \in \{r, b\}$. It is easy to establish that risk averse agents vote sincerely in equilibrium, i.e., $v_b^* = v_r^* = 1$. We continue to assume that R and B are picked with equal probability in the absence of a signal. Thus, the information acquisition probability, σ, is what determines equilibrium behavior.

Under risk-neutrality, a voter has to focus only on the pivotal event p_σ, given by equation (4). However, with more general utility functions, voters have to take into account both pivotal and *non-pivotal* events. Conditioning on the other $N-1$ voters' votes not ending up in a tie, the probability of a correct decision is

$$q_\sigma = \sum_{k=\frac{N+1}{2}}^{N-1} \binom{N-1}{k} z_\sigma^k (1 - z_\sigma)^{N-1-k},$$

where z_σ is the probability of a random voter voting in favor of the correct alternative, and is given by equation (3).

The respective expected utilities from acquisition and non-acquisition of informa-

[21]The denominator allows us to normalize $u(0) = 0$ and $u(1) = 1$ for all α, just like the risk-neutral case.

tion are now given by

$$U(\sigma_1) = [u(1) - u(0)] [p_\sigma x + (1 - p_\sigma)q_\sigma] + u(0)$$
$$U(\sigma_0) = [u(1 + c) - u(c)] \left[p_\sigma \frac{1}{2} + (1 - p_\sigma)q_\sigma\right] + u(c)$$

If there is some $\sigma^* \in (0, 1)$ for which $U(\sigma_1) = U(\sigma_0)$, then σ^* is an equilibrium. Otherwise, we have a corner solution in the usual way.

At this stage, it is useful to study how the net benefit from acquisition of information $U(\sigma_1) - U(\sigma_0)$ depends on the extent of risk aversion, as captured by the parameter α in the CARA specification of the utility function. In doing so, we keep the belief (p_σ, q_σ) about others' strategies fixed.

The net benefit of information acquisition, $U(\sigma_1) - U(\sigma_0)$, can be written as

$$\left\{[u(1) - u(0)] [p_\sigma x + (1 - p_\sigma)q_\sigma] - [u(1 + c) - u(c)] \left[p_\sigma \frac{1}{2} + (1 - p_\sigma)q_\sigma\right]\right\} - \{u(c) - u(0)\} \quad (11)$$

When $u(m) = m$, we have $[u(1) - u(0)] = [u(1 + c) - u(c)] = 1$. It is easy to check that the first term of (11) now reduces to $p_\sigma \left(x - \frac{1}{2}\right)$ and the second term reduces to c, giving us condition (5) in section 3.1.

If we allow risk aversion, there are two opposing effects on incentives to acquire information. First, for any $\alpha > 0$, the marginal gain in utility from obtaining the correct group decision is higher when information is acquired than when it is not:

$$u(1) - u(0) > u(1 + c) - u(c).$$

Moreover, this difference in marginal gains increases with α, raising the first term in the expression (11). On the other hand, as a subject becomes more risk averse, she values more the cost saved by *not* acquiring information in the event of a wrong group decision, i.e., the second term also increases in α. Typically, the first effect dominates for lower values of α and the second effect dominates for higher values of α. As a result, $U(\sigma_1) - U(\sigma_0)$ will typically first increase and then decrease with increases in the degree of risk aversion α. Therefore, it is in general unclear whether risk aversion leads to more or less information acquisition. Importantly for our purposes, risk aversion may help to explain why we sometimes observe too much and sometimes too little information acquisition compared with risk neutrality.

To explore the extent to which risk-aversion might help to explain our experimental findings, we build upon the QRE model of the previous section 6.4.2, and add to it the normalized CARA utility function, $u(m) = \frac{1-\exp[-\alpha m]}{1-\exp[-\alpha]}$ in place of the risk neutral

$u(m) = m$ assumption.[22] Using this CARA specification for $u(\cdot)$, we estimate the information acquisition choice strategy of a subject using the logistic specification:

$$\sigma(\alpha, \lambda) = \frac{1}{1 + \exp[-\lambda\{U(\sigma_1; \alpha) - U(\sigma_0; \alpha)\}]}$$

where $U(\cdot; \alpha)$ is understood to be calculated as expected utility based on the above risk-averse preference $u(m)$. We then maximize the likelihood function,

$$\mathcal{L}(\alpha, \lambda) = \sigma(\alpha, \lambda)^{\sigma_1} [1 - \sigma(\alpha, \lambda)]^{\sigma_0} \tag{12}$$

using the data from each treatment, (x, N, c). As before, we also maximize

$$\mathcal{L}(\alpha_p, \lambda_p) = \prod_j \sigma(\alpha_p, \lambda_p)^{\sigma_1^j} [1 - \sigma(\alpha_p, \lambda_p)]^{\sigma_0^j} \tag{13}$$

using the pooled data from all of our treatments where $x < 1$.[23] The results from maximum likelihood estimation of the QRE model with CARA preferences are reported in Table 9. The top part of this table reports the unpooled, treatment-by-treatment estimation results while the bottom part reports the pooled estimation results.

The estimates using the pooled data indicate some improvement in the fit of the QRE model with risk aversion relative to the model that assumes risk neutral preferences. In particular, we note that the mean squared error (MSE) for the QRE with CARA preferences is slightly smaller, MSE=0.0126 (vs. 0.0148 for the risk neutral case) by incorporating the additional risk-aversion parameter α into the model. While the absolute change in the MSE is not very large, the percentage decrease in the MSE from including risk aversion, 14.94%, is considerable. We further note that the pooled rationality parameter estimate, $\hat{\lambda}_p = 9$, is now higher than that found for the QRE model without risk aversion (where $\hat{\lambda}_p = 3.3$).

Finally, we note that the estimates of the risk aversion parameter, $\hat{\alpha}$, are generally quite plausible with the pooled estimate, $\hat{\alpha}_p = 1.1$, indicating a moderate degree of risk aversion among our subjects. Nevertheless, as it turns out the allowance for some amount of risk aversion yields a significantly better fit of the QRE model to the experimental data. In particular, according to a likelihood ratio (LR) test, the difference between the unrestricted QRE model which allows for risk–averse preferences and the restricted QRE model with risk-neutral preferences ($\alpha = 0$) is highly significant (*LR Stat* $= -2 \ln l = 54.81 \gg \chi^2_{.001} = 10.83$ with $d.f. = 1$).[24] Thus we can easily reject the null hypothesis (H_0) that the subjects' preferences are risk-neutral ($p < 0.001$).

[22]See also Goeree et al. (2002, 2003).

[23]Again, estimates for the $x = 1$ treatments are available upon request.

[24]The degree of freedom is one here as we have only one restriction on risk-aversion parameter that $\alpha = 0$.

Precision	Group Size	Cost	$\hat{\lambda}$	$\hat{\alpha}$	$\hat{\sigma}$	Observed	Predicted1	Predicted3
$x = 0.7$	$N = 3$	$c = 5$	28	3.1	0.695	0.695	1	1
		$c = 8$	59	2.4	0.582	0.582	0.999	1
		$c = 25$	7	0.3	0.272	0.272	0	0
	$N = 7$	$c = 5$	180	1.7	0.767	0.767	0.969	0.669
		$c = 8$	1	1.2	0.500	0.514	0	0
	$N = 13$	$c = 8$	4	3	0.463	0.463	0	0
Precision	Group Size	Cost	$\hat{\lambda}_p$	$\hat{\alpha}_p$	$\hat{\sigma}_p$	Observed	Predicted2	Predicted3
$x = 0.7$	$N = 3$	$c = 5$	9	1.1	0.602	0.695	1	1
		$c = 8$	9	1.1	0.541	0.582	1	1
		$c = 25$	9	1.1	0.233	0.272	0	0
	$N = 7$	$c = 5$	9	1.1	0.522	0.767	0.996	0.669
		$c = 8$	9	1.1	0.464	0.514	0	0
	$N = 13$	$c = 8$	9	1.1	0.432	0.463	0	0

Notes: $\hat{\lambda}$, $\hat{\alpha}$ and $\hat{\sigma}$ are estimated for each treatment with fixed (x, N, c) while $\hat{\lambda}_p$, $\hat{\alpha}_p$ and $\hat{\sigma}_p$ are estimated using pooled data from all $x = 0.7$ treatments. The MSE between the pooled estimates and the experimental data is found to be 0.01258. Predicted1 is the Nash equilibria with treatment-specific risk-aversion parameter $\hat{\alpha}$; Predicted2 is the Nash equilibria with pooled risk-aversion parameter $\hat{\alpha}_p$; and Predicted3 is the Nash equilibria with risk-neutral preferences ($\alpha = 0$).

Table 9: Quantal Response Equilibrium with Risk-Aversion: Maximum Likelihood Estimates

Summarizing, of the three behavioral models we have considered in section 6.4, the QRE model with risk averse preferences with respect to uncertain money amounts is the one that can best rationalize our experimental finding of both over- and under-acquisition of costly information prior to voting.[25]

[25] In an earlier draft, we also considered a fourth behavioral model of subjective beliefs equilibrium (SBE) due to Elbittar et al. (2017). That model supposes that some fraction of voters hold biased prior beliefs about the true state of the world that depart from the correct 0.5 prior. We found that we could only apply this model to cases where $x = .7$, since in the perfect information case ($x = 1$), biased prior beliefs don't matter for posterior. For the $x = 0.7$ case we found that the SBE estimates provided a fit to the data that was better than QRE model but not as good as the QRE model with risk aversion.

7 Conclusion

We have designed and reported on an experiment examining the effects of group size, information cost and signal precision on information acquisition decisions made prior to committee or jury voting decisions. Our experiment, building upon the work of others, is the first to systematically explore the comparative statics implications of changes in group size, information cost and signal precision for information acquisition and efficiency under a compulsory, majority rule voting mechanism. A comparison of the behavior of subjects in our experimental setting with the comparative statics predictions of the theory is important for understanding the extent to which subjects appreciate the tradeoff between better information aggregation (from greater information acquisition) and information free-riding that the theory emphasizes.

Our experimental findings suggest that there is mixed support for the comparative statics predictions of the rational choice theory of endogenous information acquisition and voting. In particular, when signals are noisy ($x = 0.7$), there is only weak evidence for a group size effect where the frequency of information acquisition decreases as the group size increases due to the eventual dominance of the free-riding effect over the information aggregation effect. However, when signals are precise ($x = 1$) there *is* strong evidence for a group size effect; we conjecture that subjects better comprehend the free-riding problems of larger group sizes when signals are precise as compared with when they are imprecise. With regard to the effect of information cost, we find that higher costs reduce the frequency of information acquisition just as rational choice theory predicts, but we do not observe the often sharp, corner solution point predictions of the theory, e.g., where σ^* goes from 1 to 0 as the information cost, c, is steadily increased.

We observe that in most of our treatments, subjects are over-investing in costly information, hence the extent of free-riding is not as large as predicted and consequently, efficiency is not decreasing so rapidly with increases in the group size. Many subjects appear to be ignoring strategic considerations and acting as lone decision-theorists as evidenced by the significant percentages of subjects who always buy information in all 25 rounds, even in settings where in the rational choice equilibrium, no voter should buy information in any round. If subjects incorrectly perceive the setting to be one where $N = 1$, then it can be rational to buy information whenever $M(x - 1/2) \geq c$, a condition that holds in most of our treatments thus enabling a decision-theoretic rationale for always acquiring information. We conducted one treatment where the information cost was very large so that the inequality did not hold. In that treatment we found the lowest mean level of information acquisition across all of our treatments, 27.17%, but this frequency of information acquisition was still greater than the game-theoretic and decision-theoretic prediction of zero information acquisition.

The characterization of some subjects as decision theorists can explain over-acquisition of information, but it cannot explain the under-acquisition of information that we sometimes also observe, e.g., in our $N = 3$, $x = 0.7$ treatments. To explain both patterns of behavior with regard to information choices, we first considered a noisy best response or *quantal response* equilibrium analysis which is generally useful in explaining the more gradual changes in subject behavior relative to the sharp point predictions of rational choice theory. Using the QRE approach we find that we can successfully account for the phenomena of both over- and under-acquisition of information across all of our experimental treatments albeit with a rather low rationality parameter estimate for λ. We further show how augmenting the QRE approach to allow for risk averse voters can aid in rationalizing both over- and under-acquisition of information relative to the theoretical equilibrium predictions that assume risk neutral agents. When we allow for risk averse subjects in our QRE estimation, we find that this unrestricted model fits the data better than does the restricted model allowing for risk neutral subjects only. We also find a modest improvement in the mean squared error between the observed information acquisition frequencies and the QRE estimates allowing for risk aversion and we find confirmatory evidence that our experimental subjects are modestly risk averse.

Our experiment considers only the case of compulsory, majority rule voting with symmetric signal precisions and relatively small group sizes. We think that a promising direction for future research on group size and cost effects on endogenous information acquisition would be to consider larger information costs or larger group sizes (e.g., as in an internet-based voting experiment) especially in settings with perfectly precise signals, as the $x = 1$ setting is the one where we have found the closest correspondence between the theory and the experimental findings. It would also be of interest to adapt the model so that subjects could purchase more precise signals at a higher cost. Finally, it would be of interest to study endogenous information acquisition in the case where voters can freely communicate with one another following the information acquisition stage and receipt of any signals, but prior to voting, in which case free-riding considerations might become even more pronounced relative to the no-communication environment that we study in this paper. We leave these extensions to future research.

References

Ali, S.N., Goeree, J.K., Kartik, N., Palfrey, T.R., 2008. Information aggregation in standing and ad hoc committees. Amer. Econ. Rev. 98, 181-186.

Anderson, L.R., Holt, C.A., Sieberg, K.K., Oldham, A.L., 2015. An experimental study of jury voting behavior, in: Schofield, N., Caballero, G.(Eds.), The Political Economy of Governance: Institutions, Political Performance and Elections. Springer, Berlin, pp. 157-178.

Battaglini, M., Morton, R.B., Palfrey, T.R., 2010. The swing voter's curse in the laboratory. Rev. Econ. Stud. 77, 61-89.

Bhattacharya, S., Duffy, J., Kim, S., 2014. Compulsory versus voluntary voting: an experimental study. Games Econ. Behav. 84, 111-131.

Condorcet, M., 1785. Essai sur l'application de l'analyse la probabilité des décisions rendues à la pluralité des voix. De l'imprimerie Royale, Paris.

Duffy, J., Tavits, M., 2008. Beliefs and voting decisions: a test of the pivotal voter model. Amer. J. Polit. Sci. 52, 603-618.

Elbittar, A., Gomberg, A., Martinelli, C., Palfrey, T.R., 2017. Ignorance and bias in collective decision: theory and experiments. Forthcoming, J. Econ. Behav. Org.

Feddersen, T., Pesendorfer, W., 1997. Voting behavior and information aggregation in elections with private information. Econometrica 65, 1029-1058.

Gerardi, D., Yariv, L., 2008. Information acquisition in committees. Games Econ. Behav. 62, 436-459.

Goeree, J.K., Holt, C.A., Palfrey, T.R., 2002. Quantal response equilibrium and overbidding in private-value auctions. J. Econ. Theory 104, 247-272.

Goeree, J.K., Holt, C.A., Palfrey, T.R., 2003. Risk averse behavior in generalized matching pennies games. Games Econ. Behav. 45, 97-113.

Goeree, J.K., Yariv, L., 2011. An experimental study of collective deliberation. Econometrica 79, 893-921.

Großer, J., Seebauer, M., 2016. The curse of uninformed voting: an experimental study. Games Econ. Behav. 97, 205-226.

Guarnaschelli, S., McKelvey, R.D., Palfrey, T.R., 2000. An experimental study of jury decision rules. Amer. Polit. Sci. Rev. 94, 407-423.

Koriyama, Y., Szentes, B., 2009. A resurrection of the Condorcet jury theorem. Theor. Econ. 4, 227-252.

Levine, D.K., Palfrey, T.R., 2007. The paradox of voter participation? A laboratory study. Amer. Polit. Sci. Rev. 101, 143-158.

McKelvey, R.D., Palfrey, T.R., 1995. Quantal response equilibria for normal form games. Games Econ. Behav. 10, 6-38.

Martinelli, C., 2006. Would rational voters acquire costly information? J. Econ. Theory 129, 225-251.

Martinelli, C., 2007. Rational ignorance and voting behavior. Int. J. Game Theory 35, 315-335.

Mechtenberg, L., Tyran, J-R., 2016. Voter motivation and the quality of democratic choice. Working paper, https://ssrn.com/abstract=2869787

Mukhopadhaya, K., 2003. Jury size and the free rider problem. J. Law Econ. Org. 19, 24-44.

Oliveros, S., 2013. Abstention, ideology and information acquisition. J. Econ. Theory 148, 871-902.

Persico, N., 2004. Committee design with endogenous information. Rev. Econ. Stud. 71, 165-191.

Appendix A: Experimental Instructions

In this Appendix we provide the instructions used in the $N = 7$, $c = 5$, $x = 0.7$ treatment of our experiment. Instructions for the other treatments are similar.

Instructions

Welcome to this experiment in the economics of decision-making. Funding for this experiment has been provided by the National Science Foundation. We ask that you not talk with one another for the duration of today's session.

For your participation in today's session you will be paid in cash at the end of the experiment. Different participants may earn different amounts of money. The amount you earn depends partly on your decisions, partly on the decisions of others, and partly on chance. Thus it is important that you listen carefully and fully understand these instructions before we begin. There will be a short comprehension quiz following the reading of these instructions which you will all need to complete before we can begin the experimental session.

The experiment will make use of the computer workstations, and all interaction among you will take place through these computers. You will interact anonymously with one another and your data records will be stored only by your ID number; your name or the names of other participants will not be revealed at any time during today's session or in any write-up of the findings from this experiment.

Today's session will involve 14 subjects and 25 rounds of a decision-making task. In each round you will view some information and make a decision. Your decision together with the decisions of others determine the amount of points you earn each round. Your dollar earnings are determined by multiplying your total points from all 25 rounds by a conversion rate. In this experiment, each point is worth 1 cent, so 100 points = $1. Following completion of the 25th round, you will be paid your total dollar earnings plus a show-up fee of $5.00. Everyone will be paid in private, and you are under no obligation to tell others how much you earned.

Specific details

At the start of each and every round, you will be randomly assigned to one of two groups, the R (Red) group or the B (Blue) group. Each group will consist of 7 members. All assignments of the 14 subjects to the two groups of size 7 at the start of each round are equally likely. Neither you nor any other member of your group or the other group will be informed of whether they are assigned to the R or B groups until the end of the round.

Imagine that there are two "jars", which we call the red jar and the blue jar. Each jar contains 10 balls; the red jar contains 7 red balls and 3 blue balls while the blue jar contains 7 blue balls and 3 red balls. The red jar is always assigned to the R (Red) group and the blue jar is always assigned to the B (Blue) group. However, recall that you do not know which group (Red or Blue) you have been assigned to; that is, you don't know the true color of your group's jar. Furthermore, your assignment to the R or B group is randomly determined at the start of every round.

To help you determine the jar that has been assigned to your group for the round, you and each member of your group can decide whether or not you want to independently choose one ball from your group's jar and privately observe the color of that ball. You face this decision on the first decision screen for each round where you are asked: Do you want to draw a ball? If you click on no, then you can get additional points as will be explained in detail below, however, in that case you will not have any more information about the jar that has been assigned to your group; all you will know is that there is a 50 percent chance your group is assigned to the red jar and a 50 percent chance your group is assigned to the blue jar. If you click on yes, then you will be shown 10 different balls that you can choose. The balls are numbered 1 to 10. You must then click on one of the 10 balls. When you are satisfied with your choice click the OK button. After doing so you will be privately informed of the color of that ball. You will not be informed about whether other members of your group chose to select a ball, or how many members of your group chose to select a ball (until the end of the round), nor will they learn whether you chose to select a ball. You will also not be told the color of the balls drawn by any other members of your group who chose to draw balls, nor will they learn the color of the ball you chose, and it is possible for members of your group to draw the same ball as you do or any of the other 9 balls as well. Each member in your group who chooses to draw a ball selects one ball on their own and only observes the color of his/her own ball. However, all members of your group (Red or Blue), if they decide to choose a ball, will choose a ball from the *same* jar that contains the same number of red and blue balls. Recall again that if you are choosing a ball from the red jar, that jar contains 7 red balls and 3 blue balls while if you are choosing a ball from the blue jar, that jar contains 7 blue balls and 3 red balls.

After all group members have decided whether or not to draw a ball and those choosing to draw a ball have chosen their ball and observed its color, all group members will face a second decision screen where they will be asked to make a choice about the color of the jar that has been assigned to their group. Specifically, all group members, regardless of whether or not they have chosen to draw a ball, will face a choice between RED or BLUE for the color of the jar that has been assigned to their group. Those who chose to draw a ball will be reminded on this second decision screen of the color of the ball they have drawn. But all group members, even those who did not choose to draw a ball must choose whether the jar assigned to their group is BLUE or RED by clicking on either the blue or the red buttons.

Your group's decision depends on the individual member decisions. Your 7-member group's decision is RED if 4 or more of the members of your group (a majority) choose RED and your group's decision is BLUE otherwise, that is, if 4 or more of your group members (a majority) choose BLUE.

Suppose you selected to draw a ball (and selected RED or BLUE). If your group's decision (via majority rule) is the same as the true color of the jar that is assigned to your group, then the group decision is CORRECT, and you and every member of your group earns 100 points from the group's correct decision. If your group's decision is different from the true color of your group's jar, then the group decision is INCORRECT, and you and every member of your group will earn 0 points from the group's incorrect decision.

Suppose you selected not to draw a ball. Then you get an additional 5 points for the round. In other words, if your group's decision is the same as the true color of the jar that is assigned to your group, then you will earn 105 points from the group's correct decision. If your group's decision is different from the true color of your group's jar, then you will earn 5 points from the group's incorrect decision. Thus, by choosing to draw a ball to be further informed of the true color of the jar that is assigned to your group, you give up an additional 5 points for the round.

If the final (25th) round has not yet been played, then at the start of each new round you will again be randomly assigned to one of two groups of size 7. One group, Group R, will be assigned to the red jar and the other group, Group B will be assigned to the blue jar. Again, no one will know to which group or jar they have been assigned. Each group member will have the opportunity to privately decide whether or not to draw a new ball from your group's jar and observe its color (your decision to draw a ball in the previous round doesn't affect your decision for the current round), and then to choose between BLUE or RED. In other words, the group you are in will change from round to round.

Following completion of the final, 25th round, your points earned from all 25 rounds will be converted into cash at the rate of 1 point = 1 cent. You will be paid these total earnings together with your $5 show-up payment in cash and in private.

Questions?

Now is the time for questions? If you have a question about any aspect of these instructions, please raise your hand and an experimenter will answer your question in private.

Quiz

Before we start today's experiment we ask you to answer the following quiz questions that are intended to check your comprehension of the instructions. The numbers in these quiz questions are illustrative; the actual numbers in the experiment may be quite different. Before starting the experiment we will review each participant's answers. If there are any incorrect answers we will go over the relevant part of the instructions again.

1. I will be assigned to the same group, R or B in every round. Circle one: True False.

2. I must draw a ball from my group's jar in every round. Circle one: True False.

3. If you decide to draw a ball from your group's jar and the color of the ball you have actually drawn is red, then the color of your group's jar is also red. Circle one: True False.

4. The red jar contains _____ red balls and _____ blue balls. The blue jar contains _____ red balls and ______ blue balls.

5. Consider the following scenario in a round. 4 members of your group choose RED.

   a. What is your group's decision? ________

   b. If the jar of balls your group was drawing from was in fact the RED jar and if you have drawn a ball from the jar, how many points do you earn? ________

   c. If the jar of balls your group was drawing from was in fact the BLUE jar and if you have drawn a ball from the jar, how many points do you earn? ________

   d. If the jar of balls your group was drawing from was in fact the RED jar and if you have *not* drawn a ball from the jar, how many points do you earn? ________

   e. If the jar of balls your group was drawing from was in fact the BLUE jar and if you have *not* drawn a ball from the jar, how many points do you earn? ________

Appendix B: Proof of Remark 1

Proof. To see why, note that from condition (7), it is optimal to (not to) acquire information if the pivot probability is greater (less) than $\frac{c}{M\left(x-\frac{1}{2}\right)}$. In a candidate equilibrium where k^* voters acquire information and $N-k^*$ do not, the pivot probability faced by a voter who acquires information is $p(k^*-1, N-1)$ and that faced by one who does not acquire information is $p(k^*, N-1)$. Finally, strict monotonicity of $p(k, N)$ in its first argument guarantees the uniqueness of k^*. ■