# An Experimental Study of Collective Deliberation

Jacob K. Goeree — UZH

Leeat Yariv* — Caltech

August 17, 2010

Abstract. We study the effects of deliberation on collective decisions. In a series of experiments, we vary groups' preference distributions (between common and conflicting interests) and the institutions by which decisions are reached (simple majority, two-thirds majority, and unanimity). Without deliberation, different institutions generate significantly different outcomes, tracking the theoretical comparative statics. Deliberation, however, significantly diminishes institutional differences and uniformly improves efficiency. Furthermore, communication protocols exhibit an array of stable attributes: messages are public, consistently reveal private information, provide a good predictor for ultimate group choices, and follow particular (endogenous) sequencing.

**JEL classification:** C92, D02, D72

**Keywords:** Jury Decision Making, Deliberative Voting, Strategic Voting.

*Goeree: Chair for Organizational Design, University of Zürich, Blümlisalpstrasse 10, CH-8006, Zürich, Switzerland. Yariv: Division of the Humanities and Social Sciences, California Institute of Technology, Mail code 228-77, Pasadena, CA 91125, USA. We thank the editor, Larry Samuelson, and three anonymous referees for very helpful comments. We also thank Gary Charness, Guillaume Frechette, Dino Gerardi, John Kagel, Alessandro Lizzeri, Tom Palfrey, and Lise Vesterlund for many useful conversations and suggestions. Lauren Feiler, Salvatore Nunnari, and Julian Romero provided us with superb research assistance. We gratefully acknowledge financial support from the National Science Foundation (SES 0551014) and the European Research Council (ERC Advanced Grant, ESEI-249433).

# 1. Introduction

## 1.1 Overview

Ranging from jury decisions to political elections, situations in which groups of individuals determine a collective outcome are ubiquitous. There are two important observations that pertain to almost all collective processes observed in reality. First, decisions are commonly preceded by some form of communication among individual decision makers (such as jury deliberations, or election polls). Second, even when looking at a particular context, say U.S. civil jurisdiction, there is great variance in the type of *institutions* that are employed to aggregate private information into group decisions.[1]

The recent theoretical literature has tried to assess the potential impacts of communication on group decision processes making strong assumptions on the format of conversation (e.g., Austen-Smith and Feddersen, 2005, 2006, analyzing one shot simultaneous communication, or Gerardi and Yariv, 2007, allowing for general cheap talk). While experimental and field investigations of collective decisions progress hand in hand, there are several inherent difficulties germane to field data in the context of group deliberation. First, the prior inclinations of decision makers, the accuracy of information, etc. may suffer from both endogeneity problems as well as may be difficult to calibrate. Second, protocols of conversation are rarely obtainable. Indeed, the existing field analysis in the jury context uses either exit surveys, or mock juries.[2] Third, a controlled comparison of institutions is very difficult practically. Juries serve as a prime example in which communication is structured into the decision-making process. Even for particular types of cases, there is great institutional variance across state jurisdictions. Nonetheless, out-of-court settlements are not fully documented and may be affected by the voting rule at place, which makes for harsh empirical endogeneity problems (Priest and Klein, 1984).

[1] For example, in 30 state civil courts in the U.S., non-unanimous voting rules are employed ranging from 2/3 majority to 7/8 majority and anything inbetween. See State Court Organization 1998, U.S. Department of Justice, Office of Justice Programs, available online at: http://www.ojp.usdoj.gov/bjs/pub/pdf/sco98.pdf.

[2] For an overview of recent empirical research on deliberating juries, see Devine, Clayton, Dunford, Seying, and Pryce (2001).

The current paper reports observations from some of the first lab experiments aimed at understanding the effects of different institutions on outcomes when communication channels are available, as well as the impact of different preference distributions within a group on institutional performance. Furthermore, our design allows us to provide a characterization of the endogenous formation of communication protocols under different institutions and group preferences.

Specifically, we conducted an array of experiments emulating a jury decision-making process, in which groups of nine subjects were required to make a collective decision between one of two alternatives (a neutral version of acquittal or conviction). The returns to either alternative were randomly determined according to the realization of an underlying state (such as a guilty or innocent defendant) and each subject received a private signal about that realization (similar to the subjective interpretations of testimonies in a trial). We implemented a $3 \times 3 \times 2$ design. Namely, we varied the distribution of preferences among subjects (one distribution entailing common interests, and two allowing for different formats of heterogeneity), the institution or voting rule by which the group decision was made (simple majority, 2/3 super-majority, and unanimity), and the availability (or unavailability) of free-form communication.

Our experimental setup can be thought off as a metaphor for a wide variety of settings, including not only jury voting, but also investment decisions by corporate strategy committees, hiring and tenure decisions by university faculty, performances rated by a group of judges, and more.

There are several insights that come out of our investigation. First, without the ability to communicate, agents behave in a rather sophisticated strategic manner. Across treatments, agents vote against their private information when the informative equilibrium prescribes them to do so. While the experimental observations do not match the Bayesian Nash predictions point-wise numerically, the data do reveal the theoretically predicted comparative statics, across voting rules and across preferences. One consequence of subjects' strategic behavior is that, absent communication, the efficiency of simple majoritarian rules is greater than that emerging from voting rules requiring more consensual decisions (see, e.g., Feddersen and Pesendorfer, 1998).

The second, and possibly most important insight is that free-form communication greatly improves efficiency as well as diminishes institutional differences. The extent to which institutional differences are mitigated depends on the preference heterogeneity between individuals. In particular, when agents have shared (or homogeneous) preferences, as much of the extant strategic voting literature assumes (see below), there are no significant differences between outcomes under different voting rules when communication is available. Furthermore, groups make choices that are consistent with the welfare maximizing decisions given the available *aggregate information* in the group.

These observations have important implications. On the one hand, they help explain the great variety of institutions in what appear to be very similar contexts (such as trials of a particular type). Indeed, when the panel of decision makers can freely deliberate prior to making a collective decision, the institution in and of itself may not be crucial to outcomes. On the other hand, these results suggest that from a policy perspective, affecting the communication protocols that precede decisions can serve as a vital design instrument.

The third chief insight pertains to the characteristics of the endogenously created communication protocols. In our experiments, communication is predominantly public, nearly always truthful, and is a strong predictor of group choice. Correct decisions are associated with shorter chats and higher fractions of the conversations dedicated to information exchange. Furthermore, across all treatments, protocols are consistently comprised of two distinct phases – information sharing and aggregation of opinions.

In fact, a schematic description of the procedure subjects utilize is as follows. Subjects first share their information (truthfully and publicly), then decide collectively on the ultimate decision, and finally all vote for that option. Indeed, voting in unison is the modal outcome in almost all of our communication treatments. Naturally, this procedure explains the similarity in outcomes observed across voting rules when subjects deliberate.

## 1.2 Related Literature

A formal approach to the study of collective decision making under uncertainty originated with the work of Condorcet (1785) who considered group decision problems in which members have a common interest but differ in their beliefs about which alternative is correct. In particular, Condorcet considered a model with two possible states of the world (e.g., a defendant who is innocent or guilty) and individual group members, privately and imperfectly informed about which state applies, who vote for one of two alternatives (e.g., acquit or convict). The common interest assumption assures all group members readily agree about which alternative to pick if information is public (i.e., all share the same threshold of doubt for conviction). Differences in beliefs or preferences, however, create an information aggregation problem, making it harder for the group to reach a consensus and draw the right conclusion.

Within the context of this simple two-by-two model, generally referred to as the Condorcet Jury model, Condorcet (1785) argued that majority is an efficient voting rule to aggregate the group's scattered pieces of information. Furthermore, he concluded that under majority rule, groups make better decisions than individuals and large groups almost surely make the right choice. Condorcet derived this "Jury theorem" assuming individuals vote *sincerely*, i.e., their votes simply follow their private information.

Recent work, however, has shown that rational voters do not necessarily behave this way (see Austen-Smith and Banks, 1996; Myerson, 1998; Feddersen and Pesendorfer, 1996, 1997, 1998). Since a vote matters only when it is pivotal, a strategic agent considers the information contained in the event of being pivotal, taking into account others' strategies. In particular, Nash equilibrium strategies may involve *strategic* voting, where individuals go against their private information. Moreover, equilibrium strategies systematically vary with the voting rule.

There are two sets of conclusions this literature has produced. First, unanimity is expected to perform worse than non-unanimous voting rules. In fact, under unanimity the probability of a wrongful conviction may *increase* with jury size and is bounded away from zero as the jury size grows large. Second, as jury size becomes infinitely large, non-unanimous voting rules fully aggregate the available information and generate efficient outcomes.

The design of our experiments matches the theoretical setup of Feddersen and Pesendorfer (1998). In particular, our design allows us to test for strategic voting experimentally when communication is not available under different voting rules and different preference distributions.

Lately, there have been several papers analyzing the potential impact of communication on collective choice outcomes. Coughlan (2000) and Austen-Smith and Feddersen (2005, 2006) were among the first to point out that the availability of particular communication protocols[3] can dramatically alter collective decisions, while Gerardi and Yariv (2007) show that unrestricted communication (such as jury deliberation) renders a large class of voting rules equivalent in terms of the sets of sequential equilibrium outcomes they generate.[4] It is the latter paper that motivates the design of the experimental sessions with communication. We allow for free-form communication, study the emergent (endogenous) communication protocols, and compare the outcomes generated by different institutions.

Experimentally, there have been several recent laboratory inquiries of group decision making. Guarnaschelli, McKelvey, and Palfrey (2000) test some of the extreme Nash predictions by inspecting a jury (of size three and six) and varying the voting rule (majority and unanimity). Their data confirm the Nash prediction that unanimity rule triggers strategic voting; jurors with an innocent signal mix between acquit and convict.[5] In contrast, under majority rule voting tends to be sincere. Battaglini, Palfrey, and Morton (2010) also identify strategic voting behavior in the form of the so-called "swing-voter's curse" (Feddersen and

[3] Coughlan (2000) considers straw polls and Austen-Smith and Feddersen (2005, 2006) consider one stage simultaneous and public conversation. See also Elster (1998) for related work in other fields.

[4] Lizzeri and Yariv (2010) achieve a similar result for certain environments when considering communication protocols that entail a stage of costly information collection and a stage of collective decision. Gerardi and Yariv (2008) effectively consider communication protocols as a design instrument in a particular mechanism design setup pertaining to information acquisition within collective choice. Meirowitz (2006) considers a mechanism design problem that generates incentives for protocols to be carried out in a particular way.

[5] Ladha, Miller, and Oppenheimer (1999) provide experimental evidence for strategic voting in a related setting. Bottom, Ladha, and Miller (2002) illustrate the implications of non-Bayesian updating in the Condorcet world.

Pesendorfer, 1996). For an overview of political economy experiments, see Palfrey (2006).

Communication is specifically incorporated in Dickson, Hafer, and Landa (2008), who study the interpretation of information by subjects in a one-round protocol in which subjects (with potentially different preferences and private information) simultaneously decide whether to speak or to listen.[6]

As a summary of the extant literature, we note that the experiments described in this paper provide three important methodological innovations. Most importantly, our study constitutes a first experimental inquiry of how free-form communication affects institutional outcomes.[7] In addition, we allow for intermediate voting rules in addition to majority and unanimity rules (intermediate voting rules are surprisingly under-studied in the formal literature in view of their prevalence). Finally, our experimental treatments include juries with homogeneous and heterogeneous preferences.

### 1.3 Paper Structure

Section 2 describes the experimental design. The corresponding theoretical predictions are analyzed in Section 3. We start the description of the experimental observations in Section 4 in which we test for strategic voting. The collective outcomes generated by each institution, with and without the possibility to deliberate, are described in Section 5. A detailed analysis of the experimental communication protocols appears in Section 6. The protocols' effects on experimental juries' behavior is discussed in Section 7. Section 8 concludes.

[6] McCubbins and Rodriguez (2006) consider a completely different setup with experimental communication. Their subjects need to decide on a solution to an SAT problem (of unknown difficulty) and they allow subjects (with unknown math abilities) to communicate in one round (they can send or not one signal, and listen or not to others' signals). They show that the quality of individual decisions can decrease after such communication. In another different context, Cooper and Kagel (2005) illustrate how team communication makes groups behave more strategically as well as respond quicker to payoff changes than individuals. The effects of communication have also been studied experimentally in other settings, e.g., in partnerships as in Charness and Dufwenberg (2006), or dictator games as in Andreoni and Rao (2009).

[7] Guarnaschelli, McKelvey, and Palfrey (2000) allow for restricted communication, i.e., deliberations taking the form of a straw poll vote (as in Coughlan, 2000). They find that voters tend to expose their private information less than theory predicts and the impact on jury outcomes is small. In contrast, the free-form communication allowed for in our experiments has a dramatic effect on jury outcomes.

## 2. Experimental Design

The underlying setup of our experimental design replicates the characteristics of Condorcet's simple model. There is a "red" jar and a "blue" jar: the red jar contains seven red and three blue balls and the blue jar contains seven blue and three red balls. Throughout the paper, we use the red (blue) jar as a metaphor for a guilty (innocent) defendant. At the start of each period, subjects are randomized into a group of nine subjects (who are assigned labels 1 through 9 randomly) and one of the jars is chosen by a toss of a fair coin. Subjects receive private information and ultimately need to cast a vote pertaining to their guess of which jar had been chosen and are each paid according to their own and their (eight) fellow group members' guesses. There are four important components of our experimental design: the private information each subject gets, subjects' ability to interact, the voting rule in place, and subjects' preferences.[8]

**Information** In each period, after the jar had been selected, each of the nine jurors in a group receives an independent draw (with replacement) from the jar being used. The color of the drawn ball matches the jar's color with probability $q = 0.7$, commonly referred to as the *accuracy* of the private signal.

**Communication** In the no communication, or "no chat" treatments, subjects cast their guesses immediately after observing their private draws. In the communication or "chat" treatments, subjects can communicate with one another via a chat screen that automatically opens when subjects receive their private draws. They are able to direct their messages to a subset of their group or to the group as a whole (i.e., send a public message). Messages can take any form and communication is not restricted in time. When subjects are done chatting they cast their votes for red or blue.

**Voting rules** Once all votes have been received they are automatically tallied to determine the group outcome. The voting rule, explained to the subjects at the outset of the

[8] The experimental instructions are available at http://sites.google.com/site/jurydeliberation/.

experiment, is a threshold rule, where the red jar is the group choice if and only if at least (a pre-specified) $r$ red votes are submitted. There are three types of treatments corresponding to three different voting rules: $r = 5$ (simple majority), $r = 7$ (two-thirds majority), and $r = 9$ (unanimity).

**Preferences** Subjects' payoffs, which depend on whether the group decision matches the jar being used, vary by treatment. In the "homogeneous" treatment, subjects' preferences are completely aligned. In the "heterogeneous" treatment, subjects are randomly assigned (with equal probabilities) the role of weak-red or weak-blue partisan, which causes a misalignment in preferences. The weak-red (weak-blue) partisans are predisposed to choose the red (blue) jar, or, in other words, require stronger information favoring the blue (red) jar in order to prefer it. This misalignment is even stronger in the "partisan" treatment, where jurors are assigned the role of strong-red partisan with probability 1/6, a role in which the red outcome is preferred regardless of the realized jar. Subjects are informed of the ex-ante distribution of preferences, and their own realized preferences in each round (but not the full realization of preferences in their group). The top panel of Table 1 displays the payoffs (in cents) used in the different treatments.

To summarize, the experiments employ a $3 \times 3 \times 2$ design based on variations in voting rules, jurors' preferences, and the availability of communication amongst the subjects. Each experimental session implemented one particular voting rule and one particular preference distribution. Within sessions, we conducted 15 periods without communication followed by 15 periods with communication (with one practice round preceding each). Three of the sessions were repeated with the chat periods preceding the no-chat periods to check for order effects. These "reverse order" sessions led to qualitatively identical insights as our baseline treatments. In our analysis below, we therefore pool the data from both types of sessions.[9]

[9] Separate analysis of the sessions in which rounds with communication preceded the rounds without communication is available from the authors upon request.

| Preferences | Homogeneous | | Heterogeneous | | Partisan | |
|---|---|---|---|---|---|---|
| | Neutral [1] | | Weak Red Partisan [1/2] | | Neutral [5/6] | |
| | True Jar Red | True Jar Blue | True Jar Red | True Jar Blue | True Jar Red | True Jar Blue |
| Jury Choice Red | 100 | 10 | 150 | 10 | 100 | 10 |
| Jury Choice Blue | 10 | 100 | 10 | 50 | 10 | 100 |
| | | | Weak Blue Partisan [1/2] | | Strong Red Partisan [1/6] | |
| | | | True Jar Red | True Jar Blue | True Jar Red | True Jar Blue |
| Jury Choice Red | | | 50 | 10 | 150 | 50 |
| Jury Choice Blue | | | 10 | 150 | 10 | 25 |

| Subjects | r=5 | r=7 | r=9 | r=5 | r=7 | r=9 | r=5 | r=7 | r=9 |
|---|---|---|---|---|---|---|---|---|---|
| no-chat, chat | 36 | 27°+36*+45* | 27 | 18+45*^^ | 27+45* | 36+45*^ | 27 | 36 | 36 |
| chat, no chat | N/A | N/A | 18 | 27 | 18 | N/A | N/A | N/A | N/A |

* Starred sessions were ones in which only the no-chat treatment was run.
° Chat treatment was run for 18 rounds (in addition to the practice round).
^ This session was run for only 9 rounds (in addition to the practice round).
^^ Chat treatment was run for only 9 rounds (in addition to the practice round).

Table 1: Experimental Design

The experiments were conducted at the California Social Sciences Experimental Laboratory (CASSEL) at UCLA. The bottom panel of Table 1 describes the number of subjects participating in each of the treatments (where summands correspond to separate sessions). Overall, 549 subjects participated. The average payoff per subject from the no chat segment of each session was \$9.53, while the corresponding average payoff in the chat segment was \$13.11. In addition, each subject received a \$5 show-up fee.

## 3. Theoretical Predictions

Our experimental design matches the basic jury setup introduced by Feddersen and Pesendorfer (1998). Formally, consider a group of $n = 2k+1$ individuals (subjects, jurors, etc.) who collectively choose one out of two alternatives, $\{red, blue\}$ (as suggested above, this can be understood as a metaphor for a choice between convicting or acquitting a defendant) using a threshold voting rule parameterized by $r = 1, \ldots, n$. That is, $red$ (convict) is chosen if and only if at least $r$ agents vote in favor of it. In our experimental treatments, $n = 9$ and $r = 5, 7, 9$. At the outset, a state of nature is chosen randomly from $\{R, B\}$ (experimentally, red or blue jar; metaphorically, guilty or innocent defendant), and individuals's private preference types are randomized from $T = \{$Neutral, Weak Red Partisan, Weak Blue Partisan, Strong Red Partisan$\}$ according to the prior probability $p = (p_N, p_{WR}, p_{WB}, p_{SR})$.

Utility mappings for each type are determined naturally according to Table 1 above. After preference types had been determined, each agent observes a conditionally independent signal $s \in \{red, blue\}$ of accuracy $q$. That is,

$$\Pr(s = red \mid R) = \Pr(s = blue \mid B) = q,$$

where $q = 0.7$ in all our experimental treatments.

After observing all of their private information (comprised of preference type and signal), when communication is not available, agents vote simultaneously, the group choice is determined according to $r$, and agents' earnings are determined accordingly. In our experimental design, each treatment corresponds to a different prior $p$. In particular, in the homogeneous treatment, $p_N = 1$, in the heterogeneous treatment, $p_{WR} = p_{WB} = \frac{1}{2}$, and in the partisan treatment, $p_N = \frac{5}{6}$ and $p_{SP} = \frac{1}{6}$. A strategy is then a mapping $\sigma : T \times \{red, blue\} \to [0, 1]$, which associates a probability of choosing red (or convict) for each realization of private preference type and revealed signal. We concentrate on *symmetric responsive* equilibria in which agents of the same extended type (comprising preference type and private signal) use the same strategy, and not all extended types use the same strategy. Using the techniques of Feddersen and Pesendorfer (1998), we identify the equilibrium strategies generated by the assortment of our experimental sessions.

Consider first the homogeneous treatments. When $p_N = 1$ and $r = k + 1$, the unique symmetric equilibrium entails agents following their signals, selecting *red* (*blue*) when observing *red* (*blue*), as in Austen-Smith and Banks (1996). Intuitively, if all agents follow their signals, then a pivotal agent knows that precisely $k$ agents observed the signal *red* and $k$ agents observed the signal *blue*. These signals cancel one another, and the agent best responds by following her own signal.

For $r > k+1$, this sincere behavior is no longer part of an equilibrium. Indeed, if all vote sincerely, then pivotality implies that there are at least two more *red* signals in the group,

| | | r=5 | r=7 | | | | r=9 | | |
|---|---|---|---|---|---|---|---|---|---|
| Homogeneous | red votes with red signals | 1 | 1 | | | | 1 | | |
| | red votes with blue signals | 0 | 0.31 | | | | 0.77 | | |
| | Pr(red\|blue) [="Pr(C\|I)"] | 0.099 | 0.108 | | | | 0.206 | | |
| | Pr(blue\|red) [="Pr(A\|G)"] | 0.099 | 0.280 | | | | 0.474 | | |
| | | | Eq 1 | Eq 2 | Eq 3 | Eq 4 | Eq1 | Eq2 | Eq3 |
| Heterogeneous | **Weak Red Partisans** | | | | | | | | |
| | red votes with red signals | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| | red votes with blue signals | 1 | 1 | 1 | 0.32 | 1 | 1 | 1 | 1 |
| | **Weak Blue Partisans** | | | | | | | | |
| | red votes with red signals | 0 | 0.41 | 0.64 | 0 | 0 | 0.14 | 0.02 | 0 |
| | red votes with blue signals | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | Pr(red\|blue) [="Pr(C\|I)"] | 0.5 | 0.166 | 0.224 | 0.002 | 0.5 | 0.003 | 0.002 | 0.5 |
| | Pr(blue\|red) [="Pr(A\|G)"] | 0.5 | 0.678 | 0.472 | 0.972 | 0.5 | 0.995 | 0.998 | 0.5 |
| Partisan | **Strong Red Partisans** | | | | | | | | |
| | red votes with red signals | 1 | 1 | | | | 1 | | |
| | red votes with blue signals | 1 | 1 | | | | 1 | | |
| | **Neutrals** | | | | | | | | |
| | red votes with red signals | 0.97 | 1 | | | | 1 | | |
| | red votes with blue signals | 0 | 0.18 | | | | 0.72 | | |
| | Pr(red\|blue) [="Pr(C\|I)"] | 0.286 | 0.113 | | | | 0.201 | | |
| | Pr(blue\|red) [="Pr(A\|G)"] | 0.064 | 0.275 | | | | 0.48 | | |

Table 2: Theoretical Predictions

implying a best response of $red$ regardless of one's signal. As it turns out, for $r > k+1$, the unique responsive equilibrium entails agents with a $red$ signal voting $red$ and those with a $blue$ signal mixing between a $red$ and $blue$ vote. Let the equilibrium probability of choosing $red$ when observing a blue signal be $\alpha$. Then, after simplifying terms we get,

$$\Pr(red \mid pivotal) = \Pr(red \mid r-1 \; red \text{ votes}, \, n-r \; blue \text{ votes}) =$$

$$= \frac{[q+(1-q)\alpha]^{r-1}[(1-q)\,(1-\alpha)]^{n-r}}{[q+(1-q)\alpha]^{r-1}[(1-q)\,(1-\alpha)]^{n-r}+[1-q+q\alpha]^{r-1}[q\,(1-\alpha)]^{n-r}},$$

which, for indifference, must equal $q$. The solution of this equality for different values of $q, n$, and $r$ identifies the corresponding equilibria, as they appear in the top panel of Table 2 for $q = 0.7$, $n = 9$, and $r = 7, 9$.

The analysis of the heterogeneous and partisan treatments is similar in spirit and therefore omitted. Table 2 summarizes all equilibrium predictions germane to our no-communication experimental sessions, as well as the probabilities of the different errors, associated with choosing $R$ (red, or convict) when the state is actually $B$ (blue, or innocent) or alternatively choosing $B$ (blue, or acquit) when the state is actually $R$ (red, or guilty).[10] The former is often referred to in the jury literature as the probability of convicting the innocent, which is thus denoted $\Pr(C \mid I)$, while the latter is referred to as the probability of acquitting the guilty and denoted $\Pr(A \mid G)$.

## 4. Strategic Voting

**4.1. Aggregate Analysis.** We start by considering the extent to which subjects behaved strategically. Table 3 summarizes the relevant results for all sessions. Numbers in round parentheses correspond to theoretical predictions.[11] As will be seen in Section 6, in the treatments allowing for communication, subjects revealed their private signals at very high rates across treatments. We therefore report the aggregate choices in those sessions as a pair of percentages $x\%/y\%$, where $x\%(y\%)$ is the appropriate percentage of choices when, given the agent's preferences and the entire signal profile, the optimal decision was red (blue). Thus, a best response to truthful revelation would constitute of the pair 100%/0%.[12] Strong partisans had a dominant action entailing a vote for red and we therefore report their aggregate choices only.[13] Last, for the homogeneous case, there is an appealing equilibrium (in terms of Pareto optimality or efficiency) in which all reveal their signals and vote for the commonly preferred alternative. The errors that would have resulted in the experiment with such behavior are reported in the square parentheses of the top panel.

[10] The multiplicity of equilibria in the heterogeneous case when $r = 7$ or $r = 9$ is inherent for symmetric settings in which there are weak red and weak blue partisans. In particular, this multiplicity could not be avoided by specifying different *symmetric* rewards for correct matches between group choice and actual states for both types of partisans.

[11] Since there are multiple equilibria for the heterogeneous treatment, we do not include any theoretical predictions for the corresponding sessions. The theoretical error predictions are based on the equilibrium strategies and *realized* signal profiles in the experimental sessions.

[12] For instance, in the heterogeneous treatment, red types require only 4 out of 9 signals to be red for red to be the optimal choice. So, for example, under simple majority ($r = 5$), in 93% of the time in which there were at least 4 red signals and a red type received a red signal, she voted red. Similarly, blue types require 6 out of 9 red signals to prefer red over blue and numbers are calculated accordingly.

[13] Partisan subjects did not always use their dominant action. This can be explained by either a desire to conform or match the winner (see Goeree and Yariv, 2007) combined with probability matching (Siegel and Goldstein, 1959), or some form of altruism (particularly in the case of the two supermajoritarian rules), as in Feddersen, Gailmard, and Sandroni (2009). We return to their behavior in some of the individual-level analysis below.

| | | Without Communication | | | With Communication | | |
|---|---|---|---|---|---|---|---|
| | | r=5 | r=7 | r=9 | r=5 | r=7 | r=9 |
| Homogeneous | **Number of Individual Decisions** | 540 | 1620 | 675 | 540 | 486 | 675 |
| | **Number of Group Decisions** | 60 | 180 | 75 | 60 | 54 | 75 |
| | **Red Votes with Red Signal** | 91%(100%) | 89%(100%) | 90%(100%) | 99%/36% | 98%/10% | 98%/10% |
| | **Red Votes with Blue Signal** | 7%(0%) | 24%(31%) | 39%(77%) | 69%/5% | 91%/5% | 95%/4% |
| | **Wrong Jury Outcomes** | 10%(8%) | 35%(22%) | 48%(40%) | 10%[7%] | 7%[5%] | 8%[8%] |
| | **True Jar Blue** | 7%(10%) | 5%(10%) | 0%(21%) | 16%[9%] | 7%[3%] | 5%[8%] |
| | **True Jar Red** | 13%(6%) | 60%(30%) | 97%(54%) | 4%[4%] | 7%[6%] | 11%[8%] |
| Heterogeneous | **Number of Individual Decisions** | 1080 | 1350 | 945 | 675 | 675 | 540 |
| | **Number of Group Decisions** | 120 | 150 | 105 | 75 | 75 | 60 |
| | **Red Types** | | | | | | |
| | Red Votes with Red Signal | 86% | 88% | 91% | 93%/25% | 82%/33% | 84%/6% |
| | Red Votes with Blue Signal | 37% | 44% | 49% | 78%/3% | 56%/6% | 72%/11% |
| | **Blue Types** | | | | | | |
| | Red Votes with Red Signal | 64% | 59% | 62% | 91%/44% | 84%/28% | 84%/38% |
| | Red Votes with Blue Signal | 16% | 15% | 19% | 91%/13% | 73%/12% | 73%/24% |
| | **Wrong Jury Outcomes** | 23% | 41% | 60% | 7% | 13% | 30% |
| | **True Jar Blue** | 24% | 4% | 0% | 11% | 3% | 10% |
| | **True Jar Red** | 23% | 83% | 100% | 3% | 22% | 52% |
| Partisan | **Number of Individual Decisions** | 405 | 540 | 540 | 405 | 540 | 540 |
| | **Number of Group Decisions** | 45 | 60 | 60 | 45 | 60 | 60 |
| | **Neutral Types** | | | | | | |
| | Red Votes with Red Signal | 91%(97%) | 90%(100%) | 71%(100%) | 100%/16% | 98%/41% | 80%/33% |
| | Red Votes with Blue Signal | 18%(0%) | 21%(18%) | 28%(72%) | 95%/2% | 82%/17% | 38%/8% |
| | **Partisan Types** | | | | | | |
| | Red Votes with Red Signal | 81%(100%) | 90%(100%) | 68%(100%) | 83%(100%) | 90%(100%) | 82%(100%) |
| | Red Votes with Blue Signal | 57%(100%) | 45%(100%) | 50%(100%) | 48%(100%) | 47%(100%) | 42%(100%) |
| | **Wrong Jury Outcomes** | 27%(23%) | 25%(16%) | 43%(34%) | 9% | 13% | 15% |
| | **True Jar Blue** | 36%(35%) | 3%(9%) | 0%(20%) | 12% | 21% | 4% |
| | **True Jar Red** | 12%(3%) | 48%(23%) | 100%(52%) | 5% | 8% | 22% |

Table 3: Strategic Voting Across Treatments

There are several insights one gains by inspecting Table 3. First of all, in the homogeneous and partisan no-communication treatments, behavior generally follows the comparative statics (if not the precise numbers) predicted by theory. In particular, voting against one's blue signal under rules $r = 7$ and $r = 9$ is significantly different than 0 for any conventional levels of confidence. Furthermore, voting against a blue signal increases in a significant way with the voting rules (again, for any conventional levels of confidence).[14] Nonetheless, in all of our treatments, subjects took at least 20% longer to make a decision when ultimately voting against their signal, suggesting that voting against one's signal may involve a more complex cognitive process.[15]

[14] Results for homogeneous preferences can readily be compared to those obtained by Guarnaschelli, McKelvey, and Palfrey (2000), for groups of size 3 and 6, and majoritarian and unanimous voting rules. Our observations are consistent with those reported there.

[15] Voting with the signal took an average of 41.4, 55.1, and 30.5 seconds within the homogeneous, heterogeneous, and partisan treatments, respectively. Voting against one's signal took an average of $51.3, 72.2,$and

The qualitative deviations from the theoretical predictions pertain to the probability of convicting an innocent defendant (i.e., the probability that the group outcome is red when the blue jar is being used).[16] In the homogeneous and partisan no-communication treatments, this probability declines with the size $r$ of the super-majority needed for conviction (a choice of red). This comparative static, which is not predicted by theory, has been observed before in the experiments of Guarnaschelli, McKelvey, and Palfrey (2000), who focused on simple majority and unanimity. Furthermore, under unanimous voting rules ($r = 9$), convictions (red choices) are hardly observed, and so wrong convictions ("$\Pr(C \mid I)$") are rare. Indeed, without the ability to communicate, it is hard to achieve a unanimous profile of votes. This is important from a policy perspective, as the levels of $\Pr(C \mid I)$ are often the object of minimization when assessing institutions. In the lab, absent deliberation, unanimous rules generate very low innocent convictions (see also Guarnaschelli, McKelvey, and Palfrey, 2000).

Looking at the communication treatments, Table 3 illustrates that subjects respond to the entire profile of signals available in their group, though appear to place too much weight on their own signals (conditional on full revelation). This ties to the reduced overall probabilities of wrong outcomes when communication is available. Note, however, that under unanimity, the probabilities of wrong outcomes when the jar is blue (wrongful convictions) are significantly higher with communication than without at any conventional confidence level. Indeed, as will be shown below, subjects can more easily create a majority, super-majority, or even a unanimous vote for red when deliberation precedes choice.

Throughout the paper, we report results from all sessions. It is important to note that when looking at sessions in which the order of the communication and no communication treatment was reversed, we see very little difference in strategic behavior[17], and wrong jury

36.7 seconds within the respective homogeneous, heterogeneous, and partisan treatments. All differences were significant at any reasonable level.

[16] The theoretical values concerning wrong decisions (bottom three rows in each panel) capture the probabilities that would have been generated had subjects used the theoretical equilibrium strategies for the *experimental* signal realizations.

[17] For the sessions with homogeneous preferences and $r = 9$, in which reversed sessions were run and theoretical predictions are unique, looking at votes for red with red signal and with blue signal, we get p-values corresponding to differences in the baseline sessions of 0.82 and 0.62, respectively.

outcomes occur at similar, though slightly lower frequencies.

**4.2. Individual Behavior.** To uncover the determinants of strategic voting and to test for learning, we estimate a discrete choice model on each individual's decision to vote red as a function of several explanatory variables. In addition to dummy variables corresponding to voting rules 7 and 9, we consider several additional dummy variables: *Red Sample* takes the value 1 when the subject's signal is red, *Red Type* takes the value of 1 when the subject is a weak red partisan in the heterogeneous treatments, and when they are a strong partisan in the partisan treatments, *Past Wrong Blue Dec(ision)* takes the value of 1 when blue was the outcome in the previous round and ended up not coinciding with the realized state and thereby allows us to identify reinforcement forces, and *Late* allows us to account for learning by taking the value of 1 when the decision is taken in the last 5 periods of the session. In addition, *Number of Red Signals* captures the number of red signals in the group, and we consider several natural interaction terms. Table 4 contains the marginal effects corresponding to our estimations (where errors are clustered by subject).

Several insights come out of these estimations. First, and in line with our aggregate analysis, subjects put significant weight on their private information captured by our *Red Sample* variable. They do so in a significantly more prominent manner in the treatments without communication. As we will see below, subjects frequently reveal their private information in the communication treatments. Therefore, the *Number of Red Signals* variable is a proxy for the public information available in the communication treatments. Table 4 illustrates the significant impact of the group's information whenever communication is possible (in fact, in the homogeneous and heterogeneous treatments, two additional red signals within the group influence behavior approximately as much as a private red signal, while in the partisan treatment an additional red signal in the group outweighs the effect of a private red signal). Second, voting rules have some effect on behavior and response to private signals, but the effect is limited and appears most dominant in the homogeneous preference treatments. Third, types have some effect on behavior, particularly in treatments with

| Preferences | Homogeneous | | Heterogenous | | Partisan | |
|---|---|---|---|---|---|---|
| Communication | No | Yes | No | Yes | No | Yes |
| Red Sample | 0.814*** | 0.504*** | 0.514*** | 0.248*** | 0.711*** | 0.226*** |
| | (0.069) | (0.072) | (0.047) | (0.071) | (0.070) | (0.082) |
| Past Wrong Blue Decision | -0.044 | -0.088 | -0.005 | -0.0002 | -0.0328 | 0.209*** |
| | (0.043) | (0.122) | (0.036) | (0.070) | (0.058) | (0.060) |
| Rule 7 | 0.271*** | -0.426* | 0.001 | 0.104 | 0.010 | 0.644*** |
| | (0.103) | (0.219) | (0.062) | (0.168) | (0.105) | (0.099) |
| Rule 9 | 0.385*** | -0.611** | 0.052 | 0.228 | 0.119 | 0.234 |
| | (0.066) | (0.296) | (0.062) | (0.183) | (0.085) | (0.218) |
| Number of Red Signals | | 0.220*** | | 0.221*** | | 0.360*** |
| | | (0.027) | | (0.030) | | (0.043) |
| Red Sample * Rule 7 | -0.311* | -0.422*** | -0.029 | -0.051 | 0.025 | 0.055 |
| | (0.184) | (0.076) | (0.081) | (0.106) | (0.157) | (0.093) |
| Red Sample * Rule 9 | -0.449*** | -0.485*** | -0.050 | -0.162** | -0.349*** | -0.061 |
| | (0.131) | (0.074) | (0.079) | (0.079) | (0.120) | (0.099) |
| Late | 0.0003 | -0.644*** | 0.026 | -0.292** | 0.014 | -0.320* |
| | (0.057) | (0.187) | (0.049) | (0.120) | (0.056) | (0.185) |
| Late * Red Sample | 0.049 | 0.076 | -0.083* | -0.184** | -0.075 | -0.072 |
| | (0.045) | (0.121) | (0.045) | (0.073) | (0.065) | (0.071) |
| Late * Rule 7 | 0.070 | -0.003 | -0.048 | 0.078 | -0.004 | 0.155** |
| | (0.067) | (0.093) | (0.057) | (0.052) | (0.087) | (0.072) |
| Late * Rule 9 | 0.012 | 0.094 | -0.072 | 0.219*** | -0.044 | 0.206*** |
| | (0.059) | (0.093) | (0.062) | (0.056) | (0.080) | (0.078) |
| Late * Past Wrong Blue Dec | -0.122 | 0.158 | 0.003 | 0.033 | -0.052 | -0.041 |
| | (0.082) | (0.167) | (0.059) | (0.134) | (0.108) | (0.154) |
| Late * Number of Red Signals | | 0.152** | | 0.052** | | 0.066* |
| | | (0.064) | | (0.024) | | (0.037) |
| Number of Red Signals * Rule 7 | | 0.136** | | -0.047 | | -0.186*** |
| | | (0.054) | | (0.040) | | (0.050) |
| Number of Red Signals * Rule 9 | | 0.219*** | | -0.054 | | -0.067 |
| | | (0.085) | | (0.042) | | (0.058) |
| Red Type | | | 0.257*** | -0.106 | 0.309*** | 0.451*** |
| | | | (0.048) | (0.105) | (0.070) | (0.036) |
| Red Type * Past Wrong Blue Dec | | | 0.016 | 0.080 | -0.007 | -0.182 |
| | | | (0.039) | (0.121) | (0.122) | (0.202) |
| Red Type * Rule 7 | | | 0.084 | 0.010 | -0.032 | -0.449*** |
| | | | (0.055) | (0.075) | (0.106) | (0.125) |
| Red Type * Rule 9 | | | 0.088 | 0.019 | -0.059 | -0.463*** |
| | | | (0.075) | (0.082) | (0.095) | (0.126) |
| Red Type * Red Sample | | | 0.031 | 0.007 | -0.353*** | -0.063 |
| | | | (0.038) | (0.063) | (0.069) | (0.114) |
| Red Type * Number of Red Signals | | | | 0.055** | | -0.050* |
| | | | | (0.023) | | (0.027) |
| Pseudo-Rsquared | 0.376 | 0.710 | 0.218 | 0.465 | 0.28 | 0.620 |
| Observations | 2835 | 1701 | 3375 | 1890 | 1485 | 1485 |

Robust standard errors in parentheses
* significant at 10% level; ** significant at 5% level; *** significant at 1% level

Table 4: Probit Estimations Explaining Red Individual Decisions

strong red partisans. In these treatments, partisan subjects, for whom a red vote is a weakly dominant action, vote red at a significantly greater frequency (notably under the non-unanimous voting rules). Last, learning seemed to play a limited role. Indeed, behavior in later periods is for the most part not significantly different than early behavior when communication is unavailable. With communication, subjects did tend to choose the red action less frequently at later periods. Nonetheless, the reaction to the environment (as captured by the interaction terms) did not change significantly across the experimental periods.

In relation to our theoretical predictions, note that in the treatments without communication, individual equilibrium choices depend on the voting rule, the private sample, and the private preference type. This conforms with what we observe using our regression analysis, implying again a qualitative match of our subjects' behavior with the theoretical predictions when communication was unavailable.

In what follows, we analyze how this individual behavior aggregates into group decisions, which will allow us to assess outcomes of the institutions we consider.

## 5. Voting Outcomes

A natural object when comparing institutions is the resulting outcome, i.e., the mapping from the characteristics of the group (preferences, information, etc.) to final decisions (e.g., probabilities of conviction in a jury). Theoretically, without communication, the different voting rules generate different outcomes for any of the preference distributions (see Table 2 above). On the other hand, the availability of free-form communication yields an equivalence of the set of outcomes generated by intermediate voting rules (and to a subset of outcomes under unanimity).

Comparison of outcomes is particularly important when making policy decisions. It is the natural basis upon which to choose one institution over the other, as it captures information about the likelihood of specific decisions (say, conviction or acquittal) for particular profiles of agents (e.g., jurors' political stands) and available information (such as testimonies).

We start with the homogeneous treatments, which are the easiest to analyze in that characteristics of the group can be fully summarized by the number of red signals in the group. In these treatments, symmetry assures that outcomes are encapsulated formally by the correspondence between the number of red signals in the group and the eventual probability of collectively choosing the red jar. Table 5 contains the experimental outcomes with and without communication.

| Number of Red Signals | Without Communication | | | With Communication | | |
|---|---|---|---|---|---|---|
| | r=5 | r=7 | r=9 | r=5 | r=7 | r=9 |
| 0 | - (0) | 0% (2) | 0% (2) | - (0) | - (0) | 0% (1) |
| 1 | 0% (3) | 0% (11) | 0% (8) | 0% (4) | 0% (5) | 0% (12) |
| 2 | 0% (12) | 0% (30) | 0% (10) | 0% (9) | 0% (4) | 0% (9) |
| 3 | 0% (9) | 0% (21) | 0% (11) | 0% (10) | 0% (8) | 0% (8) |
| 4 | 25% (4) | 0% (19) | 0% (8) | 29% (7) | 10% (10) | 0% (7) |
| 5 | 56% (9) | 24% (25) | 0% (9) | 100% (4) | 50% (4) | 60% (5) |
| 6 | 100% (8) | 29% (31) | 0% (12) | 100% (9) | 100% (9) | 100% (17) |
| 7 | 100% (7) | 54% (24) | 0% (9) | 100% (9) | 100% (10) | 100% (11) |
| 8 | 100% (7) | 82% (11) | 0% (5) | 100% (7) | 100% (3) | 100% (4) |
| 9 | 100% (1) | 100% (6) | 100% (1) | 100% (1) | 100% (1) | 100% (1) |

Parentheses contain the corresponding number of observations

Table 5: Frequency of Red Choices/Convictions when Preferences are Homogeneous

Table 5 illustrates the stark differences between outcomes that institutions can impose when communication is not available. For simple majority ($r = 5$) the empirical outcome approximates the statistically efficient outcome (prescribing a guess of red with 100% probability whenever 5 or more signals within the group are red, and a guess of blue, i.e., a guess of red with 0% probability, otherwise) rather well. However, under unanimity, subjects are unable to reach a consensus of red votes and the resulting outcome yields significantly less efficient outcomes.

The availability of communication overturns these results. Once communication is available, empirical outcomes are both nearly efficient as well strikingly similar across the differentvoting rules. Outcomes coincide across all voting rules when there are less than 4 or more than 5 red signals. When there are 4 or 5 signals, rule $r = 5$ generates different outcomes than the other rules $r = 7$ and $r = 9$, which generate outcomes that are not significantly different from one another (with a p-value of 0.518 corresponding to the null that the two rules do not generate different outcomes).[18]

In fact, a (non-parametric) Fisher exact probability test on group decisions rejects outcomes being identical across voting rules without communication when the number of red signals is $5-8$ at conventional significance levels. When communication is available, no pair-

[18] While communication may seem simple to conduct when agents share preferences, a large segment of the theoretical literature analyzing institutions has focused on this particular case. The results suggest the importance of accounting for communication in such circumstances.

| Preferences | HOMOGENEOUS | | | HETEROGENEOUS | | | PARTISANS | | |
|---|---|---|---|---|---|---|---|---|---|
| Voting Rule | r=5 | r=7 | r=9 | r=5 | r=7 | r=9 | r=5 | r=7 | r=9 |
| Majority Red Signals | 100% | 92.6% | 94.7% | 97.6% | 84.2% | 51.7% | 100% | 97.2% | 84.4% |
| | [100%, 100%] | [89.3%, 95.9%] | [92.4%, 97.1%] | [96.0%, 99.1%] | [80.3%, 88.1%] | [45.7%, 57.8%] | [100%, 100%] | [95.4%, 99.0%] | [80.2%, 88.6%] |
| Majority Blue Signals | 6.7% | 3.7% | 0% | 5.9% | 2.7% | 6.5% | 4% | 12.5% | 10.7% |
| | [3.7%, 9.6%] | [1.3%, 6.1%] | [0%, 0%] | [3.2%, 8.5%] | [1.0%, 4.4%] | [3.6%, 9.3%] | [1.4%, 6.6%] | [8.1%, 16.9%] | [6.9%, 14.5%] |

Square parentheses contain corresponding 95% confidence intervals

Table 6: Percentage of Red Choices/Convictions with Communication

wise comparison, for any number of red signals or two voting rules, generated a difference significant with 10% confidence.[19,20]

When preferences are heterogeneous, the analysis is complicated by the fact that it matters who holds either kind of signal. For example, a weak red partisan observing a red signal may affect decisions differently than a weak blue partisan observing a red signal.

The effect of communication on outcomes is illustrated in Table 6, which shows the percentage of red choices (convictions) when the majority of signals in the group are red or blue for the different treatments, together with their 95% confidence intervals (approximating a normal distribution). Table 6 highlights the observation that groups are highly responsive to the majority of signals within the group. For non-unanimous rules, whenever the majority of signals is red, the probability the group outcome is red exceeds 84%, regardless of the preference distribution and voting rule. Whenever the majority of signals is blue, the probability the group outcome is red is lower than 13% for all preference distributions and voting rules (including unanimous ones). In particular, the outcomes corresponding to different

[19] While the numbers reflecting rates of red choices as a function of number of red signals do not, strictly speaking, represent a cumulative distribution, they are monotonically increasing from 0 to 1. If one were to then use the (non-parametric) Kolmogorov-Smirnov test, similar results would emerge when the null is taken to be that two voting rules are identical. The values corresponding to any two rules when communication is unavailable are lower than 0.0001. When communication is available, the comparison of rules 5 and 7 leads to a value of 0.466, of rules 5 and 9 to a value of 0.255, and of rules 7 and 9 to a value of 1.

[20] We note that similar conclusions can be drawn using regression analysis. Indeed, suppose a group's decision (a dummy achieving the value of 1 when the group decision is red) is explained by the voting rule at place (accounted for by two of the voting rules, say, $r = 7$ and $r = 9$, or $r = 5$ and $r = 9$) when controlling for the number of red signals being 4 or 5 (and their interactions with the voting rules). The corresponding probit regression yields all of the coefficients regarding voting rules as not significantly different from 0.

rules appear rather similar.[21,22]

To conclude, without communication different voting rules yield significantly different group outcomes. The availability of communication reduces the effects of voting rules on outcomes. Specifically, non-unanimous voting rules generate similar outcomes in all of our experimental circumstances. Unanimous rules make it harder for groups to achieve the red outcome (conviction) and therefore appear different at times when the majority of signals in the group is red. Even this difference vanishes when preferences are homogeneous.

In terms of efficiency, individuals' response to group information is echoed in the generated outcomes that are significantly more efficient in the presence of communication. From a policy perspective, this suggests that deliberation may be an important instrument for design, and when introduced, voting rules in and of themselves may be far less so.

In the next section we analyze the communication protocols that emerged and gain more understanding regarding how group outcomes are determined in the presence of communication.

## 6. Communication Protocols

**6.1. Aggregate Protocol Characteristics.** We start by reporting general properties of the communication protocols. Table 7 summarizes the percentage of agents reporting truthfully their signals, misreporting their private signals (in the "Lies" rubric), or not

[21] In fact, looking at the 95% confidence intervals, we gain very similar insights. With the exception of unanimous voting with heterogeneous preferences, the lower bound of the 95% confidence interval corresponding to a majority of red signals in the group exceeds 80% across all treatments. Similarly, the upper bound of the 95% confidence interval corresponding to a majority of blue signals in the group lies below 17% across all treatments (for the homogeneous and heterogeneous treatments, it is below 10%).

[22] Kolmogorov-Smirnov tests generate similar messages. Without communication, a Kolmogorov-Smirnov test to compare group decisions across voting rules leads to a rejection of the null hypothesis that outcomes are the same across voting rules when communication is not available (at any conventional level of significance). Kolmogorov-Smirnov tests do not reject the coincidence of outcomes across voting rules $r = 5$ and $r = 7$ when conditioning on the more prevalent signal within the group. Outcomes from voting rule $r = 9$ are significantly different than those corresponding to rules $r = 5$ and $r = 7$ when the majority of red signals is red in the heterogeneous treatment (at 5% level) and the partisan treatment (at 10% level). For those treatments, unanimity generates significantly less red outcomes (convictions) when the information suggests red (guilt) is more likely. In all other cases of Table 6, voting rule $r = 9$ generates statistically similar outcomes to those produced under rules $r = 5, 7$.

| | | Messages | | | | | |
|---|---|---|---|---|---|---|---|
| | | Truthful | Lies | Nothing | Public | Average Number of Signal Messages | Average Number of Type Messages |
| Homogeneous | r = 5 | 90% | 10% | 0% | 100% | 8.67 | - |
| | r = 7 | 98% | 2% | 0% | 96% | 8.49 | - |
| | r = 9 | 98% | 2% | 0% | 100% | 13.88 | - |
| Heterogeneous | r = 5 | 88% | 12% | 0% | 100% | 13.79 | 0.21 |
| | r = 7 | 88% | 10% | 2% | 100% | 15.50 | 0.77 |
| | r = 9 | 89% | 10% | 1% | 100% | 8.67 | 1.20 |
| Partisan | r=5 | 93% | 6% | 1% | 100% | 7.92 | 0.04 |
| | r=7 | 89% | 9% | 2% | 100% | 7.91 | 0.31 |
| | r=9 | 92% | 8% | 0% | 100% | 8.09 | 0.16 |

Table 7: Aggregate Message Profiles

revealing anything regarding their private information. Furthermore, we account for the percentage of messages (truthful or not) that were sent publicly to the entire group.[23]

As can be seen, across treatments, a striking percentage of subjects reveal their signals truthfully and almost all subjects send messages to their entire group.

These results contrast those regarding voting without communication. While subjects are perfectly capable of behaving strategically when casting a vote, they are not very strategic when sending messages. Indeed, given that subjects react to group signals in a substantial way (see, e.g., Table 4), partisan subjects in the heterogeneous or partisan treatments would have an incentive to misrepresent signals that go against their leaning.[24]

Table 7 also reports the average number of messages conveying signal realizations and the average number of messages conveying individual types (that are relevant for the heterogeneous and partisan treatments). The former is significantly greater than the latter. In fact, type revelation occurs very rarely. For example, in the partisan treatments, the average number of types revealed is significantly lower than 0.5 with any conventional significance levels.

[23] The coding was done for the sessions in which no communication preceded the communication treatments. All coding was done by two independent research assistants that were not privy to our research questions.

[24] This is consistent with "excessive" truthful reporting observed in other experimental setups, such as the Crawford and Sobel (1982) setting, see Cai and Wang (2006).

It is worth noting that in the homogeneous treatments, unanimous chat sessions were (insignificantly) faster than majoritarian ones. The average round length under unanimity (majority) was $39 \pm 9$ ($55 \pm 9$) seconds.[25] In the heterogeneous treatments, however, communication was significantly longer under unanimity ($96 \pm 13$ seconds) than under simple majority ($26 \pm 11$ seconds) or 2/3 super-majority ($36 \pm 13$ seconds).

**6.2. Sequencing.** In order to gain insights regarding the endogenous formation of communication protocols, we identified messages that contained information about private signals, and messages that had to do with suggestions regarding how the group or particular individuals should act.[26]

Figure 1 depicts the sequencing of messages as follows. We normalized the length of all conversations within a treatment to 20 periods. For each period, we calculated the percentage of messages sent that contained signals, or suggestions, as described above. Each rubric of the Figure corresponds to a different treatment and contains two graphs – the left one depicting the evolution of signal messages, the right one illustrating the evolution of suggestion messages.[27]

Roughly speaking, conversations are consistently composed of two phases. First, subjects exchange information. Later, they converse on how to act upon the collective information. This depiction is true across the different preference settings and the different voting rules.

This split into phases allows us to identify so-called "leaders", subjects who consistently make suggestions for group and individual ultimate decisions. As it turns out, leaders do not always appear. Some sessions had unique individuals that sent numerous messages (namely, the homogeneous treatment with simple majority or the partisan treatment with unanimity). In other treatments, no clear leaders appeared.

[25] This relates to Blinder and Morgan (2005), who conducted an experiment in which groups were required to solve two problems - a statistical urn problem and a monetary policy puzzle. The groups could converse before casting their votes. They found no significant difference in the decision lag when group decisions were made by majority rule relative to when they were made under a unanimity requirement. See Cooper and Kagel (2005) for another related study.

[26] Again, these were coded by an independent research assistant.

[27] Since preference types were rarely revealed as described above, we do not include them in Figure 1.

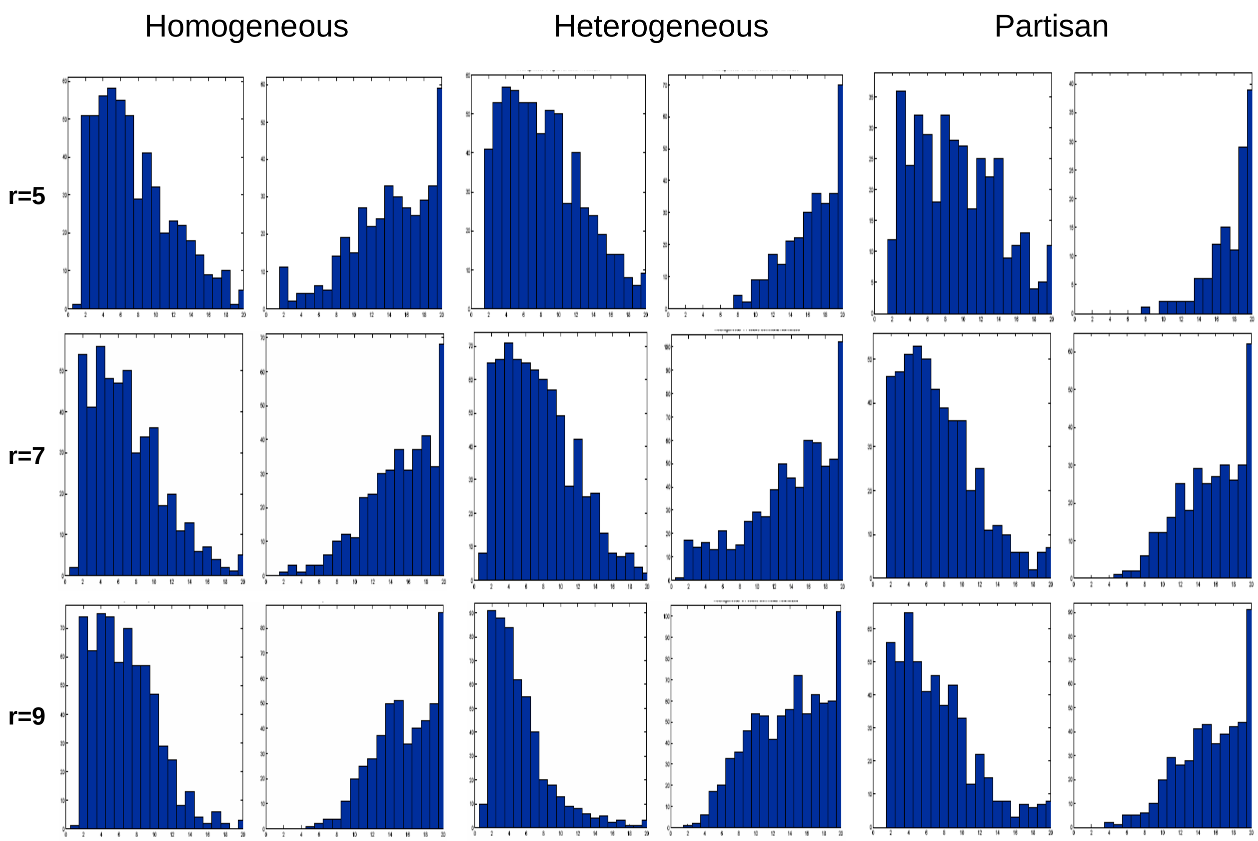


Figure 1: Sequencing within Communication Protocols (x-axis denotes normalized period and y-axis denotes percentage of signal or suggestion messages on left or right panels respectively)

We suspect that the emergence of leaders, while certainly a possibility when communication is available, is group specific.[28]

**6.3. Communication Volume and Outcomes.** We now inspect the relation between the volume of communication and the accuracy of decisions. Table 8 describes the average number of signals, the average number of overall messages (termed chat length), and the percentage of messages pertaining to observed signals in all treatments, for group decisions

[28]For the jury context, the sessions in which leaders emerged may be particularly germane. Indeed, in many U.S. courts, a jury foreperson is nominated, either by the jury itself or by the judge. The jury's foreperson effectively acts as leader - having control over some of the deliberation process as well as serving as the jury's delegate in all communications with the judge in charge (see, e.g., Abbott and Batt, 1999)..

that matched the actual state (so-called correct) and group decisions that did not match the actual state (so-called incorrect).

| | | Homogeneous | | | Heterogeneous | | | Partisan | | | Wilcoxon |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | r=5 | r=7 | r=9 | r=5 | r=7 | r=9 | r=5 | r=7 | r=9 | |
| Signals | Correct | 9.31 | 8.96 | 15.19 | 14.35 | 16.42 | 8.57 | 8.39 | 8.46 | 8.55 | W=77, |
| | Incorrect | 8.67 | 9.00 | 9.33 | 17.60 | 17.14 | 10.83 | 9.00 | 8.25 | 9.11 | p<0.48 |
| Chat Length | Correct | 20.96 | 19.36 | 32.50 | 23.93 | 40.50 | 30.00 | 11.90 | 17.27 | 18.67 | W=57, |
| | Incorrect | 30.67 | 44.25 | 34.00 | 38.00 | 47.29 | 38.06 | 30.50 | 28.63 | 31.44 | p<0.01 |
| % Signals | Correct | 0.44 | 0.46 | 0.47 | 0.60 | 0.41 | 0.29 | 0.70 | 0.49 | 0.46 | W=54, |
| | Incorrect | 0.28 | 0.20 | 0.27 | 0.46 | 0.36 | 0.28 | 0.30 | 0.29 | 0.29 | p<0.004 |

Table 8: Volume of Chats and Decision Accuracy

As can be seen from the table, while the number of signals transmitted is not significantly correlated with the groups' accuracy, the length of conversation as well as the percentage of signals transmitted within the conversation are significantly correlated with decision accuracy. Indeed, correct decisions are associated with shorter communication phases and, consequently, greater fractions of the conversations being dedicated to the transmission of information.

## 7. Group Behavior and Supermajorities

One reasoning for the equivalence of voting rules when free-form communication is available is that agents can simply circumvent the voting rule by deciding which alternative they would like to implement during deliberations and then voting unanimously for that alternative. A slight subtlety arises for unanimous voting rules for which unanimous choices in the voting stage are not robust to unilateral deviations (hence, the equivalence pertains only to intermediate voting rules, and the unanimous voting rules generate a *subset* of outcomes).

Figure 2 depicts the cumulative distribution function corresponding to all possible supermajorities (5-9) for all treatments. Note that for all of our treatments, the cumulative distribution functions corresponding to the treatments without communication (solid lines) are stochastically dominated by those corresponding to treatments with communication (dashed

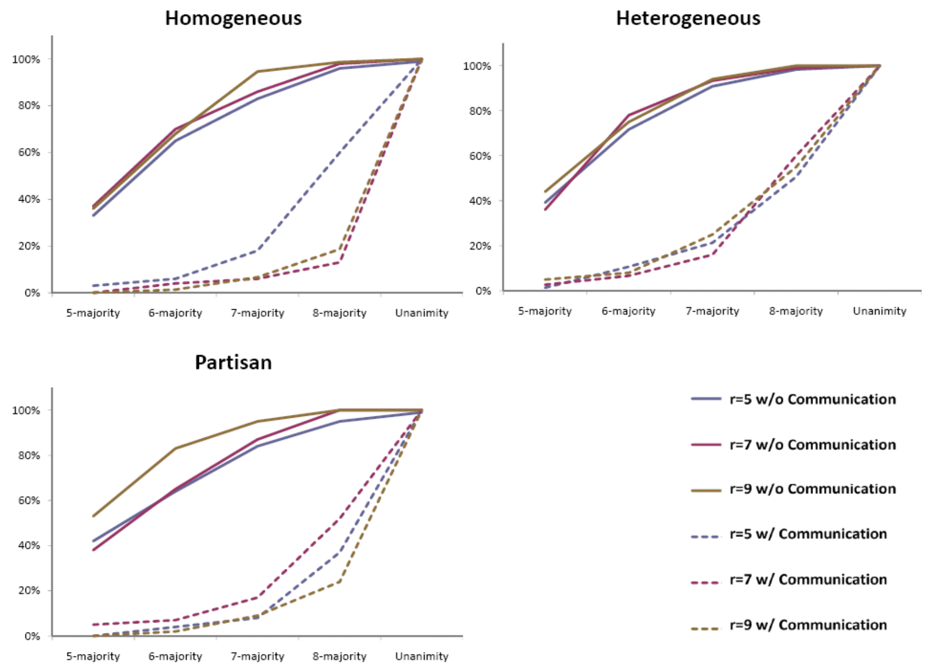


Figure 2: Cumulative Distribution Functions for Size of Supermajorities Acting in Consensus

lines). Furthermore, the cumulative distribution functions relating to the no-communication treatments are concave, while those relating to the communication treatments are convex. This captures the fact that when communication is not available, most outcomes are achieved with small supermajorities (in fact, the modal outcome is achieved with a 5 or 6 supermajority), while with communication most outcomes are achieved with large supermajorities (indeed, the modal outcomes are achieved with 8 or 9 supermajorities).

Table 7 illustrated a high percentage of subjects revealing truthfully their signals. Furthermore, Table 6 demonstrated the match between group decisions and the majority of reports in the communication stage. These numbers exceed 85% in all treatments with intermediate voting rules. These combined with the evidence captured in Figure 2 are suggestive of a heuristic process underlying the groups' decision-making algorithm. Namely, subjects share their private information and then unanimously (or almost unanimously) select the alternative supported by the majority of the signals.

## 8. Conclusions

We reported observations from an array of experiments assessing the joint impacts of heterogeneous preferences, voting rules, and the availability of communication on group (jury) outcomes. Several important insights emerge from our analysis. First, in the absence of communication, individuals behave strategically much in the spirit of theoretical jury models and consequently different voting rules yield different outcomes. Second, deliberation makes voting rules less crucial for outcomes, particularly non-unanimous ones. This is especially true when preferences of individuals are aligned. Last, communication protocols have consistent characteristics: messages are public and truthful, they are a powerful determinant of the collective choice, and are broadly divided into two phases – first, information is shared and next, a discussion ensues as to how to aggregate that information into a group decision.

The observed similarity in outcomes for non-unanimous experimental juries is consistent with the high variance of non-unanimous voting rules specified in U.S. civil jurisdiction, where non-unanimous decision rules range anywhere from simple majority to 7/8 majority. Beyond the jury context, the results are valuable for any collective decision making in which individuals communicate prior to taking decisions, be it faculty making hiring decisions, managerial teams making investment decisions, political entities deciding on policies, and so on and so forth.

The insights of the paper suggest the importance of using communication as an instrument in institutional design in conjunction with voting rules. Indeed, imposing restrictions on deliberation protocols may be an important avenue for generating desirable collective outcomes. Put differently, while much of the focus of the literature on collective decision making is on agents who are pivotal during the voting stage, understanding the agents who are effectively pivotal in the communication stage could be equally important. In fact, in practice, in many environments, agenda setting plays an important role in the design of collective decisions. In a way, an agenda can be thought of as a pre-determined communication protocol, which, as the experimental results advise, may be crucial for generating sought-after outcomes.

In fact, even without restricting protocols, the consistent sequencing of endogenous protocols we observe opens the door to new questions regarding institutional design. So far, the theoretical literature on deliberative voting has assumed that communication is either very short (entailing one round of communication, as in Austen-Smith and Feddersen, 2005) or is free-form (as in Gerardi and Yariv, 2007), much like in the experiments.[29]

Theoretical results suggests that when communication protocols are unrestricted (e.g., Gerardi and Yariv, 2007), intermediate voting rules are equivalent in terms of the *set* of sequential equilibrium outcomes they generate. Under unanimity, only a subset of the outcomes that can result with intermediate voting rules can be implemented. These results illustrate the potential effects of communication on collective outcomes, but offer little guidance on the precise product of the collective process. Our experimental results suggest stronger impacts of communication: the *selected* outcomes are the same across institutions.

We suspect that this is due to the particular format the observed (endogenous) communication protocols take. In that respect, our study suggests the importance of comparing different institutions with protocols that are in between the two polar specifications commonly studied: one-shot and fully unrestricted.

[29] One exception is Lizzeri and Yariv (2010), who study protocols resembling the two-staged ones observed in our experiments. In their setup, agents first need to decide when to halt costly communication that generates public information. Agents then collectively choose an action. The paper identifies environments in which different decision rules generate identical predictions.

## References

[1] Abbott, W. F. and Batt, J. (1999), *A Handbook of Jury Research*, American Law Institute – American Bar Association.

[2] Andreoni, J. and Rao, J. (2009), "Just Ask: One- and Two-way Communication in Dictator Games," mimeo.

[3] Austen-Smith, D. and Banks, J. (1996), "Information Aggregation, Rationality, and The Condorcet Jury Theorem," *American Political Science Review*, Volume 90(1), pages 34-45.

[4] Austen-Smith, D. and Feddersen, T. (2005), "Deliberation and Voting Rules," in Austen-Smith, D. and Duggan, J. (eds) *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks*, Berlin: Springer

[5] Austen-Smith, D. and Feddersen, T. (2006), "Deliberation, Preference Uncertainty and Voting Rules," *American Political Science Review*, Volume 100, pages 209-218.

[6] Battaglini, M., Morton, R., and Palfrey, T. R. (2010), "The Swing Voter's Curse in the Laboratory," *Review of Economic Studies*, Volume 77(1), pages 61-89.

[7] Blinder, A. S., and Morgan, J. (2005), "Are Two Heads Better than One? An Experimental Analysis of Group versus Individual Decision-making," *Journal of Money, Credit and Banking*, Volume 37(5), pages 789-811.

[8] Bottom, W., Ladha, K., and Miller, G. (2002), "Propagation of individual bias through group judgment: Error in the treatment of asymmetrically informative signals," *Journal of Risk and Uncertainty*, Volume 25(2), pages 147-163.

[9] Cai, H. and Wang, J. (2006), "Overcommunication in Strategic Information Transmission Games," *Games and Economic Behavior*, Volume 56(1), pages 7-36.

[10] Charness, G. and Dufwenberg, M. (2006), "Promises and Partnership," *Econometrica*, Volume 74, pages 1579-1601.

[11] Condorcet, Marquis de. (1785), *Essai sur l'application de l'analyse a la probabilite des decisions rendues a la probabilite des voix*. Paris: De l'imprimerie royale, translated in 1976 to "Essay on the Application of Mathematics to the Theory of Decision-Making." in Condorcet: Selected Writings, ed. Keith M. Baker. Indianapolis, IN: Bobbs-Merrill.

[12] Cooper, D. and Kagel, J. (2005) "Are Two Heads Better Than One? Team versus Individual Play in Signaling Games," *American Economic Review*, Volume 95(3), pages 477-509.

[13] Coughlan, P. (2000), "In Defense of Unanimous Jury Verdicts: Mistrials, Communication, and Strategic Voting," *American Political Science Review*, Volume 94(2), pages 375-393.

[14] Crawford, V. P. and Sobel, J. (1982), "Strategic Information Transmission," *Econometrica*, Volume 50(6), pages 1431-1451.

[15] Devine, D. J., Clayton, L. D., and Dunford, B. B. (2001), "Jury Decision Making: 45 Years of Empirical Research on Deliberating Groups," *Psychology, Public Policy, and Law*, Volume 7(3), pages 622-727.

[16] Dickson, E., Hafer, C., and Landa D. (2008), "Cognition and Strategy: A Deliberation Experiment," *The Journal of Politics*, Volume 70(4), pages 974-989.

[17] Elster, J. (1998), *Deliberative Democracy*, Cambridge: Cambridge University Press.

[18] Feddersen, T. J., Gailmard, S., and Sandroni, A. (2009), "Moral Bias in Large Elections: Theory and Experimental Evidence," *American Political Science Review*, Volume 103(2), pages 175-192.

[19] Feddersen, T. J. and Pesendorfer, W. (1996), "The Swing Voter's Curse," *American Economic Review*, Volume 86(3), pages 408-424.

[20] Feddersen, T. J. and Pesendorfer, W. (1997), "Voting Behavior and Information Aggregation in Elections with Private Information," *Econometrica*, Volume 65(5), pages 1029-58.

[21] Feddersen, T. J. and Pesendorfer, W. (1998), "Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts under Strategic Voting," *American Political Science Review*, Volume 92(1), pages 23-35.

[22] Gerardi, D. and Yariv, L. (2007), "Deliberative Voting," *Journal of Economic Theory*, Volume 134, pages 317-338.

[23] Gerardi, D. and Yariv, L. (2008), "Information Acquisition in Committees," *Games and Economic Behavior*, Volume 62, pages 436-459.

[24] Goeree, J. K. and Yariv, L. (2007), "Conformity in the Lab," mimeo.

[25] Guarnaschelli, S., McKelvey, R. C., and Palfrey, T. R. (2000), "An Experimental Study of Jury Decision Rules," *American Political Science Review*, Volume 94(2), pages 407-423.

[26] Kalven, H. and Zeisel, H. (1966), *The American Jury*, Chicago: University of Chicago Press.

[27] Ladha, K., Miller, G., and Oppenheimer, J. (1999), "Information Aggregation by Majority Rule: Theory and Experiments," mimeo.

[28] Lizzeri, A. and Yariv, L. (2010), "Sequential Deliberation," mimeo.

[29] McCubbins, M. and Rodriguez, D. B. (2006), "When does Deliberating Improve Decisionmaking?," *Journal of Contemporary Legal Studies*, Volume 15(9), pages 9-50.

[30] Meirowitz, A. (2006), "Designing Institutions to Aggregate Preferences and Information," *Quarterly Journal of Political Science*, Volume 1, pages 373-392.

[31] Myerson, R. (1998), "Extended Poisson Games and the Condorcet Jury Theorem," *Games and Economic Behavior*, Volume 25, pages 111-131.

[32] Palfrey, T. (2006), "Laboratory Experiments in Political Economy," in Weingast, B. and Wittman, D. (eds), *Handbook of Political Economy*, Oxford University Press: Oxford.

[33] Priest, G. L. and B. Klein (1984), "The Selection of Disputes for Litigation," *The Journal of Legal Studies*, Volume 13(1), pages 1-55.

[34] Siegel, S. and Goldstein, D. A. (1959), "Decision Making Behavior in a Two-choice Uncertain Outcome Situation," *Journal of Experimental Psychology*, Volume 57, pages 37-42.