# Sequential Deliberation

Alessandro Lizzeri[*] Leeat Yariv[†‡]

June 20, 2011

Abstract. We present a dynamic model of sequential information acquisition by a heterogeneous committee. At each date agents decide whether to continue deliberation, which generates costly information, or stop and take a binding vote yielding a decision. For homogeneous committees, the model is a reinterpretation of the classic Wald (1947) sequential testing of statistical hypotheses. In heterogeneous committees, the resources spent on deliberation depend on the committee's preference profile and on the rules governing the deliberation and decision-making processes. We show that voting rules at the decision stage are inconsequential when either information collection is cheap or deliberation rules are consensual. Furthermore, more heterogeneity of preferencess, more consensual deliberation rules, or more unanimous decision voting rules, lead to greater duration of deliberation and more accurate decisions. We argue that the model provides a useful way to think about jury deliberation and that it matches some stylized facts on jury decision making.

**Keywords:** Deliberation, Voting, Juries

[*]Department of Economics, NYU. E-mail: alessandro.lizzeri@nyu.edu

[†]Division of the Humanities and Social Sciences, Caltech. E-mail: lyariv@hss.caltech.edu

[‡]We thank John Morgan, Nicola Persico, Debraj Ray, and Nikita Roketskiy for comments and suggestions. We gratefully acknowledge financial support from the National Science Foundation (SES 0551014) and the Gordon and Betty Moore Foundation.

# 1. Introduction

## 1.1 Overview

Decision making in groups often involves deliberation. Juries, boards of directors, congressional and university committees, government agencies such as the FDA or the EPA, and many other committees, spend time deliberating issues before reaching a decision or issuing a recommendation. This paper presents a simple model of deliberation to study the effect of the structure of the deliberative process, and of the composition of deliberating groups, on outcomes such as the accuracy of decisions, the length of deliberation, and the degree of disagreement.

Previous literature on deliberation has focused on asymmetric information among members of the deliberating group, on how this information asymmetry can impede effective decision making, and on how different voting rules interact with this information asymmetry.[1] We abstract from private information and concentrate on a simpler aspect of collective action: how information collection responds to conflicting preferences. In a deliberating committee, there are two types of decisions to make: *deliberation decisions* and *action decisions*. Deliberation decisions are about whether to keep deliberating in order to obtain additional information. Action decisions regard the choice to be taken at the end of deliberation. Deliberation is in service of action decisions since the information that is obtained is supposed to allow more accurate action decisions.

This paper discusses a novel dimension that emerges when studying sequential deliberation in committees: the distinction between *deliberation rules* and *decision rules*. A deliberation rule governs the deliberation process and determines when information acquisition must stop. A decision rule governs the vote over issues at the end of deliberation. Deliberation rules may be different from decision rules. For instance, in many committees, the chairman of the committee has the same power as all the other members over action decisions, but has a special role to play (and more power) in deliberation decisions. While decision rules are often quite precisely described – some issues requiring a majority vote, others requiring a supermajority or unanimity – deliberation rules are often vague. Despite this vagueness, our results suggest that it is important to study the role of the deliberation process, and to understand how it functions. Much of our analysis focuses on a particularly stark way of modeling committees that have more inclusive deliberation protocols than

[1] E.g., Austen-Smith and Feddersen (2005, 2006), Coughlan (2000), Gerardi and Yariv (2007, 2008), Meirowitz (2006), and Persico (2004).

others. We model this inclusiveness as a threshold rule $R_d$ such that deliberation ends as soon as $R_d$ members of the committee vote to end deliberation. Decision rules are analogously captured by a rule $R_v$ that describes the specific qualified majority required for reaching a decision. We note that a special case of a deliberation rule in the jury setting corresponds to repeated straw polls, with a final vote taken according to the decision rule once sufficient consensus has been achieved.[2]

Our analysis discusses the effects of deliberation and decision rules on the length of deliberation, the accuracy of decisions, and the welfare of the committee and of society at large. We show that there is a sense in which deliberation rules are more "effective" (or powerful) than decision rules. For instance, we show that under certain assumptions, decision rules are irrelevant, while deliberation rules affect the length of deliberation and accuracy of ultimate decisions. Furthermore, we show that, for a range of parameters for which the decision rule does have an effect, in contrast with Feddersen and Pesendorfer (1998) and Persico (2004), unanimity leads to more informative outcomes than majority rule. We also show that, under certain symmetry assumptions, a committee would like to delegate deliberation power to a moderate chairman. Finally, we contrast sequential deliberation with static deliberation and argue that the sequential case displays a richness that is closer to the phenomena that are associated with deliberation.

The formal analysis in this paper is potentially relevant for a variety of collective decision processes such as R&D, hiring decisions, FDA drug approval, and so on. However, we focus much of our discussion on juries. This is partly because for concreteness it is useful to have a running application in mind. Furthermore, there are three additional reasons why we believe that juries are an interesting application for our model. First, in juries, the deliberation process is clear-cut and circumscribed: there is a well-defined beginning and end of deliberation, the time it takes the jury to deliberate is measurable, and one single verdict is the typical outcome of such deliberation. Second, the jury setting allows us to contrast our analysis with much of the extant body of literature on deliberation that has focused predominantly on the jury context. Third, the empirical literature has documented some patterns of deliberation in juries that we will attempt to explain with our model. We elaborate on the precise interpretation of our setup in the context of juries after we formally introduce the model.

Technically, our analysis is a natural extension of much of the analysis of individual decision

[2] The guides distributed to jurors in many U.S. courts indicate that this reflects the deliberation process suggested to juries, see Murphy and Boatright (1999a,b).

making to group contexts. Indeed, an important aspect of individual decision making is the appropriate amount of information to acquire before making a decision. An individual must weigh the cost of information against the value of making more accurate decisions. A classic approach to this question, going back to Wald (1947a,b), is that of Bayesian sequential analysis.[3] In this approach an individual acquires information sequentially, and at every stage evaluates whether he has sufficient information to make a decision: if he does, he stops and takes a decision; if he does not, he proceeds to acquire additional information. Our paper provides an analysis of how collective action affects such information acquisition. We use the term deliberation for this process.

In fact, our model is a collective action version of the analysis of sequential sampling introduced by Wald (1947a,b). In our model, a homogeneous committee deliberates in a manner that is analogous to the decision maker testing a hypothesis sequentially à la Wald: at every date, the committee evaluates its current information and decides on one of three actions: continue sampling–i.e., continue deliberating, or stop and take one of two decisions, acquittal or conviction. Wald showed that the optimal procedure involves a sequential likelihood ratio test, whereby intermediate values of the likelihood ratio require obtaining a new sample, while high (low) values of the likelihood ratio require stopping and taking one (the other) decision. We depart from Wald by introducing two possible dimensions of disagreement among committee members. The first involves disagreement exclusively on the importance of the decision (or, equivalently, on the cost of information acquisition), and hence on the length of the deliberation process. In this first version, committee members share preferences over decisions conditional on the information available, but disagree on how much information is required before making the decision. The second version involves disagreement on the appropriate decision: for example, some jurors require a higher standard of evidence in order to vote to convict. In this version, there can be disagreement at the deliberation *and* at the decision stage.

We assume that all information in the jury is public: signals are observed by all jurors, and preferences are common knowledge. This assumption represents a sharp departure from much of the extant literature on juries discussed below, where the focus is on the aggregation of private information. We view our model as a natural alternative extreme benchmark that is useful for identifying the tensions that arise in collective choice when trading off information collection costs and decision accuracy.

[3] Some precursors considered special cases, see Wald (1947a,b). See also De Groot (1970) for a simple modern exposition, and Moscarini and Smith (2001) for an extension that allows for richer sampling strategies.

## 1.2 Stylized Facts through the Lens of Sequential Deliberation

We now briefly discuss some evidence from the literature on deliberation in juries. First, we present some evidence that suggests that the structure of our model may be a realistic reduced form for thinking about deliberation. We then present evidence that is consistent with our results and that our model can help explain and interpret.

**Deliberation structure** Our model is consistent in spirit with the formal role of juries: for instance, Vidmar and Hans (2007) write "Formally their [the jurors'] task is to engage in sound fact finding from the evidence produced at trial." This formal requirement appears to be partially consistent with practice in mock juries and with surveys of actual juries: a substantial fraction of deliberation appears to be devoted to a discussion of the facts. In fact, starting from Kalven and Zeisel (1966), numerous legal scholars have argued that juries do a good job in reaching an understanding of the facts. For instance, in a study of mock juries Ellsworth (1989) writes: "In general, over the course of deliberation, jurors appear to focus more on the important facts and issues, come to a clearer understanding of them, and approach consensus on the facts." Furthermore, Hannaford et al. (2000) and Ellsworth (1989) report that some of the discussion during deliberation is about how to weigh conflicting pieces of evidence.[4]

**Importance of Deliberation** One strand of literature (appearing mostly in law and psychology scholarship) studies opinion formation by jurors. This is relevant for our model for two reasons. First, this establishes that the deliberation process is important in the formation of jurors' opinions.[5] Second, some features of the opinion formation process seem to mirror the updating process assumed in our model.

Hannaford, Hans, Mott, and Musterman (2000) studied the timing of jury opinion formation. They used a special case study of a jury reform implemented in Arizona in 1995 that allowed for discussions during civil trial (Rule 39(f) of the Arizona Rules of Civil Procedure). Their data includes survey responses of $1,385$ jurors from 172 trials in four counties (accounting for a large majority of cases in Arizona) concerning when they formed their initial opinions, whether and when they changed their minds, and when they arrived at a resolution regarding the final outcome. Over 95% of jurors

[4] We interpret discussion during deliberation about the appropriate weight of pieces of evidence as an evaluation of how the evidence should shift the posterior.

[5] This is in partial contrast with the prior received wisdom that comes from the Kalven and Zeisel (1966) landmark jury study.

reported changing their mind at least once over the course of the trial and 15% reported changing their minds more than once during trial. Importantly, over 40% of jurors reported changing their minds during the final deliberations, suggesting that deliberation is a key component of opinion formation.

Hans (2001, 2007) used surveys conducted by the National Center for State Courts (NCSC). Hans' data contains reports from close to 3,500 jurors who had participated in felony trials in four large, urban courts. Hans (2001, 2007) documents patterns of opinion change that are consistent with information collection driving a Bayesian updating process. Specifically, juries are classified into different categories depending on the outcome of a straw poll that was taken early in the deliberation process. When the initial vote in the jury strongly supports a particular outcome, that outcome is more likely to ultimately emerge. For instance, 77 of the 89 juries with strong majorities for guilt convicted the defendant, but 11 of these 89 ended up acquitting the defendant, showing that in some cases, many jurors were persuaded during deliberation to switch their vote to acquittal.[6] Figure 1 summarizes the data in Hans (2007). For each initial leaning of the jury, it describes the distribution of ultimate outcomes.

**Irrelevance of the decision rule** Baldwin and McConville (1980) studied a reform that was put in place in 1974 in England that allowed for majority verdicts in criminal trials, while prior to the reform unanimity was required. The vast majority of verdicts (311 out of 326 cases) were unanimous even after the reform, suggesting that the decision rule did not have much of an effect. Kalven and Zeisel (1966) report similar patterns for U.S. states that do not require unanimity for conviction: most verdicts are unanimous anyway.[7] Devine et al. (2001) report that in many mock jury studies there is no evidence that the decision rule has any effect on the verdict. In lab experiments, Goeree and Yariv (2011) find that, when subjects cannot talk before voting, the decision rule has an effect, whereas, when subjects can talk, the decision rule has very little effect.[8]

[6] On the other hand, 67 of the 71 juries with strong majorities for innocence in the straw poll acquitted the defendant, and 3 convicted.

These observations should be interpreted with some care, as initial polls within the jury sometimes take place after some amount of deliberations has already taken place. Thus, consensus may be overstated.

[7] A caveat to this observation is that under simple majority, when a majority of jurors agrees, any other juror's vote cannot affect the final outcome. In particular, those jurors may vote against their private assessment to satisfy social pressures at no consequence to the defendant.

[8] In the context of monetary policy, Blinder and Morgan (2008) look experimentally at two particular decision protocols (ones in which there is a designated leader, and ones in which there is not) and also find little differences in outcomes.

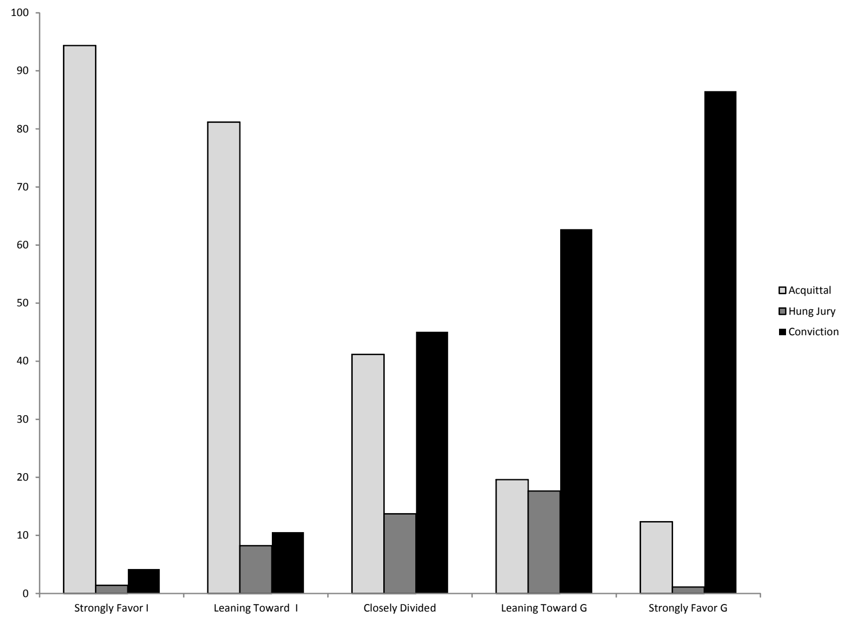


Figure 1: Distribution of Jury Outcomes (in %) According to Initial Jury Leanings (Hans, 2007)

Our model provides a possible explanation for the fact that the decision rule seems to have little or no effect. We show that, when costs of deliberation are sufficiently low, in equilibrium, deliberation always ends with unanimous decisions: whenever there is disagreement on the appropriate decision to take, members of the committee agree that it is worthwhile to continue deliberating.

**Effect on length of deliberation** Hans (2001) and Devine et al. (2001) report that the decision rule affects length and quality of deliberation. On average, using mock juries, under unanimity verdicts take as long as under majority. The quality, measured by legal experts, exhibits similar patterns – higher under unanimity than under majority.

In our model, the length of deliberation can be affected by the decision rule since pivotal members at the deliberation stage may, if not pivotal at the decision stage, prolong deliberation in order to convince the holdouts at the decision stage.

**Jury composition** Increased heterogeneity has been found to increase quality and length of deliberation (see Sommers, 2006, and Goeree and Yariv, 2011).

In our model, increased heterogeneity increases the length of deliberation since it makes the pivotal members at the deliberation stage more extreme, and therefore more definitive information is needed in order to stop deliberation. This translates immediately into longer deliberation and, in symmetric committees, more accurate decisions.

### 1.3 Literature Review

The past two decades have delivered a rich collection of work on committee decision making (see Li and Suen, 2009 for an extended survey). Our paper ties directly to several strands of studies.

In terms of jury decision making, our setup can naturally be contrasted with several papers emphasizing information aggregation within juries. Feddersen and Pesendorfer (1998) study a model in which jurors have private information about the guilt or innocence of the defendant. They show that unanimity leads to less informative outcomes than does simple majority in large juries. Persico (2004) studies a related model, but also allows for private information collection prior to voting. He characterizes the optimal voting rule and shows that unanimity leads to inferior information collection. Austen-Smith and Feddersen (2006) extend Feddersen and Pesendorfer (1998) in another way by allowing for a round of cheap-talk communication before voting. They show that unanimity leads to less communication and poorer information aggregation. Gerardi and Yariv (2007, 2008) depart from these papers by studying general communication protocols and analyzing the entire set of equilibrium outcomes. They show that the set of equilibrium outcomes is invariant to the voting rules, as long as they are non-unanimous. In fact, unanimous voting rules generate a subset of equilibrium outcomes. With respect to these papers, since one of the messages of our analysis is that more stringent decision rules entail more information collection and more accurate decisions, it implies that understanding the public component of information collection can be crucial for designing welfare enhancing institutions.

Several studies have looked at the effects of sequential collection of information within groups. Albrecht, Anderson, and Vroman (2010) and Compte and Jehiel (2010) study how group search is affected by voting. Messner and Polborn (2009) study a two-period model where voters receive information over time about the desirability of an irreversible decision. The main message of that paper is that the optimal voting rule requires a supermajority. Bognar, Meyer-ter-Vehn, and Smith (2009) also study a model of dynamic deliberation, but with very different ingredients. In their model jurors have common preferences and private information about a payoff relevant state. They assume that jurors sequentially exchange coarse messages. In their model there are many equilibria that can be ranked in terms of generated welfare. Surprisingly, longer conversations are better.[9] Strulovici (2010) analyzes a model of voting over experimentation. He shows that voting by heterogeneous voters, who are learning their preferences, leads to an inefficient level of experimentation. He then describes a voting rule that can restore efficiency. Moldovanu and Shi (2010) study a collective version of a stopping problem in which individual agents can assess the performance of alternatives on one of several dimensions. In their setting, when information is publicly transmitted, the volume of information collected, as well as the group's welfare, are shown to decrease with the variance of preferences in the group.

In comparison to all of this existing work, the main contribution of the framework proposed in this paper is that it allows for an analysis of the interplay between deliberation rules and decision rules. We identify when each plays an important role for outcomes, and how collective consequences are affected by different aspects of the environment (deliberation costs, preference heterogeneity, etc.).

From a technical perspective, the starting point of our analysis is Wald (1947a,b), who pioneered the study of sequential testing, and provided a characterization of the optimal sequential test as a sequential likelihood ratio test. We briefly describe the most directly relevant result in section 3.1.[10]

## 2. The Model

### 2.1 Setup

A jury of $n$ individuals has to determine the fate of a defendant. There are two states: $I$ (the defendant is innocent) and $G$ (the defendant is guilty), which we assume are equally likely ex-ante.[11]

All juror's want to make accurate decisions: convict the guilty and acquit the innocent. However, they differ in the importance they attribute to the two possible mistakes: convicting the innocent or acquitting the guilty. Juror $i$'s preferences are given by:

$$u_i(C,G) = u_i(A,I) = 0, \quad u_i(A,G) = -(1-q_i), \quad u_i(C,I) = -q_i,$$

[9] A related paper is that of Eso and Fong (2008), who study a dynamic cheap talk model with multiple senders, where the receiver can choose when to make her decision. They show that when the senders are all informed of the state of nature, a perfect Bayesian equilibrium exists with instantenous, full revelation, regardless of the size and direction of the senders' biases. Wilson (2011) considers exogenous costs for both sending messages and receiving them, and illustrates the dependence of effective communication on agents' quality of information and messaging costs.

[10] Moscarini and Smith (2001) consider a different extension of Wald's analysis, where they allow for simultaneous as well as sequential experimentation, and assume discounting and convex costs of sample size. Chan and Suen (2011) use the Wald setting to study the impact of heterogeneous patience among group members.

[11] Much of our analysis can be easily extended to the asymmetric case.

where $q_i$ captures the juror's concern for convicting the innocent relative to that for acquitting the guilty. In a static model without information collection, $q_i$ is also the threshold of reasonable doubt: the juror would want to convict if his posterior is higher than $q_i$, and acquit otherwise. Without loss of generality, we assume $q_1 \leq q_2 \leq ... \leq q_n$. We also assume that $q_1 > 0$ and $q_n < 1$ (otherwise at least one of the jurors is not responsive to information collection).

In determining the verdict, the jury participates in two phases: *deliberation* and *decision making*.

We assume that deliberation allows each juror to publicly acquire information about the guilt of the defendant. We formalize this collective information generation as follows. If the jurors still deliberate at time $t$, all observe the realization of the sequence of random variables $X_1, ..., X_t$, where $X_1, X_2, ....$ are independent and identically distributed conditional on the guilt or innocence of the defendant. Each random variable is drawn from an atomless distribution characterized by cumulative distribution functions $F_G(\cdot) \equiv F(\cdot|G)$ when the defendant is guilty and $F_I(\cdot) \equiv F(\cdot|I)$ when the defendant is innocent.[12]

The cost of deliberating an additional period is given by $k > 0$ per unit of time per agent. This is the cost of obtaining an additional signal (or the opportunity costs of time spent deliberating).

At each period, the jury decides whether to continue or stop deliberating using a threshold voting rule. Namely, at each period $t$, after having observed the history $X_1, ..., X_t$, each agent casts a vote whether to continue or stop information collection. Under *deliberation rule* $R_d = \lceil \frac{n}{2} \rceil, ..., n$, whenever at least $R_d$ jurors choose to stop deliberating, the deliberation phase ends.

Once deliberation comes to a halt, the decision phase takes place. The jury selects an alternative by voting. Each juror can vote to acquit, $a$, or to convict, $c$. Under the *decision rule* $R_v = \lceil \frac{n}{2} \rceil, \ldots, n$, the alternative $C$ is selected if and only if $R_v$ or more jurors vote to convict, the alternative $A$ is selected if and only if $R_v$ or more jurors vote to acquit, and the jury is hung otherwise. We assume that when the jury is hung, $A$ or $C$ are determined by the flip of a fair coin.[13]

We restrict attention to strategies that depend only on posterior beliefs $p$ (and not on the history of prior votes). Therefore, a pure strategy is a pair $(\sigma_d, \sigma_v)$, where the *deliberation strategy* is $\sigma_d : [0, 1] \to \{\text{stop, continue}\}$ and the *voting strategy* is $\sigma_v : [0, 1] \to \{a, c\}$.

[12] Allowing for atoms in the distribution would introduce some technical subtleties in the description of our results and their proofs without any qualitative differences.

[13] The exact assumption we make about the payoff consequences of hung juries is for the most part inconsequential. Our initial analysis focuses on cases where hung juries do not occur (due to low costs of deliberation). Section 8.4 discusses the case of hung juries.

### 2.2 Discussion of the model

The model is an extension of Wald (1947a,b) to study how collective action affects information collection. In the model, longer deliberation corresponds to additional signals received by the committee. Our interpretation is that this is a reasonable shortcut for thinking about how deliberation helps jurors gain an understanding of the evidence presented at trial. Of course, in a jury setting, it could be claimed that no additional information is received by the jurors during deliberation. We argue that one role of deliberation is to sift through the mass of sometimes conflicting evidence presented by two opposing parties (prosecution and defense) during the trial to figure out the relevance of different pieces of information, and the appropriate weight to attribute to these in establishing guilt or innocence of the defendant. In a way, the trial is like a lecture given by a professor, and the jury is like a study group that looks through the notes taken during class to gain some further understanding of a problem at hand.

As mentioned above, there are many alternative applications that may fit directly with the model because actual additional signals are received as deliberation continues/information is gathered. Abstractly, any scenario in which a group gathers information over time regarding two possible courses of action, with a status quo being chosen if the group cannot come to an agreement, shares features with our model. More concretely, in an R&D process, agents receive feedback about the likely success of specific avenues of research; in a drug approval process, the FDA can require additional clinical trials to be performed, and in fact, the FDA approval process is explicitly designed to require several stages of testing; in hiring practices, follow-up interviews can be requested, or additional research into a candidate can be performed; a board of directors can require additional due diligence before proceeding with a merger; and so on and so forth.

The assumption that information about both signals and preferences is public is particularly stark. It contrasts with prior work in which the focus is mostly on the aggregation of private information, and the information acquisition aspect of the collective process is disregarded. We believe that this is a useful benchmark, and that the forces highlighted in this paper would also be present under information asymmetries. However, introducing private information would naturally enrich the scope of the analysis.[14]

[14] For instance, allowing private information about preferences (the $q$'s) would be interesting. In such a setup, behavior during the deliberation phase could serve to signal agents' preferences and, technically, optimal behavior is unlikely to be stationary. We view this as a natural channel for future pursuit.

We concentrate on supermajoritarian deliberation rules, i.e., $R_d \geq n/2$. Our analysis could easily be extended to deliberation rules $R_d < n/2$. In fact, we use supermajoritarian deliberation rules simply as a way to capture succinctly the effective agenda setters during deliberation and highlight the interplay between these agents and those pivotal during the decision phase. Our analysis could allow for the specification of arbitrary agenda setters in the deliberation phase that do not come about through a vote (as is the case in settings in which, say, a committee chair determines when discussions should come to a halt).

Another restriction that we impose is that there are only two alternatives to choose from. This restriction is common to many voting models. It is not possible to obtain sharp characterizations in a model with more than two alternatives. However, we do consider a continuous action version of a simpler model (that illustrates the robustness of our basic results), which we discuss in Section 8.3.

## 3. Preliminaries

### 3.1 Homogeneous Preferences

We start by considering the benchmark of a homogenous jury containing agents with the same preference parameter: $q_1 = ... = q_n \equiv q$. In that case, the agents all face the same objective at both phases of the process. Consequently, we focus on equilibria that emulate the single person decision (by voting in unison during both the deliberation and decision phase, we discuss our equilibrium notion in Section 3.3 below). In this section we discuss the case in which $n = 1$. From Wald (1947a,b), we know the solution is unique. Formally,

**Proposition 1 (Wald, 1947)** *A unique equilibrium exists and is characterized by two thresholds:* $p^a(q) \leq \frac{1}{2} \leq p^c(q)$ *such that*[15]

- *the agent stops information collection and acquits whenever* $p_t \leq p^a$*;*
- *the agent stops information collection and convicts whenever* $p_t \geq p^c$;
- *the agent continues information collection whenever* $p_t \in (p^a, p^c)$.

[15] Recall that the prior probability that the defendant is guilty is 1/2. In particular, choosing a threshold $p^a \geq 1/2$ (or $p^c \leq 1/2$) would lead to no information collection. Our assumption that $p^a \leq 1/2 \leq p^c$ is therefore without loss of generality.

While we do not provide the proof of Proposition 1, it is useful for our analysis to illustrate the intuition behind the proposition as it is translated to our setup.[16] Fix the information cost $k$. For any posterior probability $p$, denote by $V^0(p)$ the value function associated with stopping immediately at posterior $p$.

$$V^0(p) = \max\{-q(1-p), -(1-q)p\} \tag{1}$$

Denote by $V^1(p)$ the value associated with continuing at least one more period, and $V(p)$ the overall value function for any posterior probability $p$.[17] It follows that

$$V(p) = \max\left\{V^0(p), V^1(p)\right\}. \tag{2}$$

Note that $V(0) = V(1) = 0$, and therefore, $V^1(0) = V^1(1) = -k$. Standard arguments show that $V^1(p)$ is a convex function of $p$. Indeed, consider an alternative world in which with probability $\alpha$, the probability that the defendant is guilty is given by $p_1$ and with probability $1-\alpha$, the probability that the defendant is guilty is given by $p_2$. If the (one) juror is not told which of the two probabilities had been realized, then she can guarantee the continuation value corresponding to $\alpha p_1 + (1-\alpha)p_2$. However, if she is told which of the two probabilities is realized, then with probability $\alpha$, she can guarantee the continuation value of $p_1$ and with probability $1-\alpha$ the continuation value of $p_2$. Naturally, she can ignore the information provided to her, so in the latter case she must be gaining at least as much. Convexity follows. From linearity of $-(1-q)p$ and $-q(1-p)$, and the fact that their maximal value of 0 is achieved at $p = 0, 1$, respectively, it follows that there are two posterior probabilities (that the defendant is guilty), $p^a$ and $p^c$, that define the stopping region, as in Figure 2.

When costs are high, they outweigh the benefits of information collection and stopping occurs immediately (in terms of Figure 2, when $k$ is sufficiently high, the curve corresponding to the continuation payoff lies below that corresponding to the instantaneous utility from stopping). When costs are sufficiently low, there is an interior solution. Note that convexity of the value function assures the uniqueness of such an equilibrium.[18]

[16] This discussion is adapted from De Groot (1970).

[17] The continuation value $V^1(p)$ is essentially the expectation of $V(p)$ with respect to the potential posteriors in the period that follows, minus the cost of an additional information unit $k$.

[18] In fact, from stationarity of the process, the identified thresholds would correspond to equilibria for *any* prior

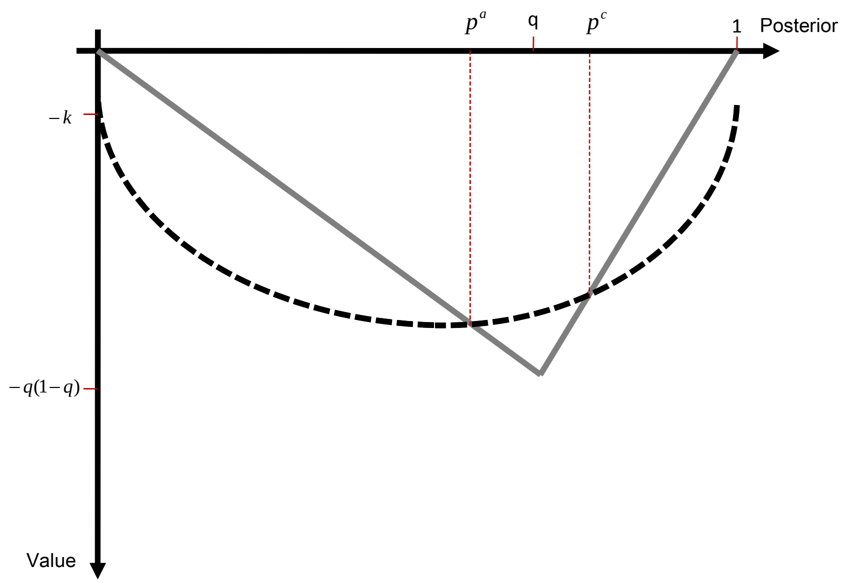


Figure 2: Homogeneous Groups – Existence and Uniqueness

The following proposition summarizes the effects of changes in the preference parameter $q$ and the cost of deliberation $k$.

**Proposition 2 (Homogeneous Juries – Comparative Statics)**

1. **Preference Parameter** $q$**.** $p^a(q;k), p^c(q;k)$ *weakly increase in* $q$*.*

2. **Cost** $k$. $p^a(q;k)$ *weakly increases in* $k$, $p^c(q;k)$ *weakly decreases in* $k$*. Consequently, the time to take a decision is weakly decreasing in* $k$*.*

Intuitively, as $q$ increases, agents care more about convicting the innocent relative to acquitting the guilty. It follows that they are willing to spend more time preventing the former relative to the latter, and that the range of posteriors for which the jury acquits becomes larger (similarly, the range of posteriors for which the jury convicts becomes smaller). When the cost $k$ increases, less information is gathered (implying that the posterior thresholds shift toward the prior) and therefore deliberation takes less time.

probability of guilt.

### 3.2 The Constrained Problem

Much of our analysis centers on heterogeneous groups. As will become transparent later, in such contexts, for different regions of the posterior, different sets of agents are potentially keen to stop or continue deliberation. In particular, the 'pivotal' agents may change with the posterior. In that respect, an agent who has the opportunity to end deliberation must evaluate the stopping problem bearing in mind that, in some other regions of the posterior, where he would ideally like to stop, other agents will force continuation. Consequently, agents face a constrained version of the original Wald problem. The following considers what will turn out to be the relevant constrained problem that is faced within heterogeneous juries.

For any two thresholds $p^a, p^c$, $p^a \leq 1/2 \leq p^c$, expected utility can be expressed as:

$$\begin{aligned} U(q;p^a,p^c) = \quad & -q\left(1-\mathbb{E}\left(p|p^c\right)\right)\Pr(p^c \text{ first} \mid p^a,p^c) - \\ & -(1-q)\,\mathbb{E}\left(p|p^a\right)\Pr(p^a \text{ first} \mid p^a,p^c) - kT(p^a,p^c), \end{aligned} \tag{3}$$

where $T(p^a,p^c)$ denotes the expected time to approach one of the posterior thresholds $p^a$ or $p^c$. The expected time $T(p^a,p^c)$ is decreasing in $p^a$ and increasing in $p^c$. The terms $\Pr(p^a \text{ first}| \; p^a,p^c)$ or $\Pr(p^c \text{ first}| \; p^a,p^c)$ correspond to the probabilities that the threshold $p^a$ or $p^c$ is reached first, respectively.[19] The expectations $\mathbb{E}\left(p|p^a\right)$ or $\mathbb{E}\left(p|p^c\right)$ denote the expected value of the posterior upon the end of deliberation conditional on passing the threshold $p^a$ or $p^c$ first, respectively.[20]

For presentational simplicity, we assume that the signal distributions $F_G$ and $F_I$ are sufficiently well-behaved so that $\Pr(p^x \text{ first}| \; p^a,p^c)$, $\mathbb{E}\left(p|p^x\right)$, and $T(p^a,p^c)$ are twice continuously differentiable with respect to $p^a$ and $p^c$, for $x=a,c$. This allows us to use calculus techniques to identify equilibrium attributes. The assumption is satisfied for many commonly used signal distributions. For instance, it holds if $F_G$ and $F_I$ are normal distributions.

We now turn to some properties of the constrained problem that prove useful subsequently.

We consider a constrained problem defined by two thresholds, $\underline{p} \leq \overline{p}$, such that the juror can only choose to stop and to acquit if $p \leq \underline{p}$, and to stop and convict if $p \geq \overline{p}$. This constrained problem is helpful when constructing best responses in the model with heterogeneous jurors. For any $\underline{p}, \overline{p}$, we define the constrained value functions as follows (dropping the arguments $q$ and $k$). As before,

[19] Note that $\Pr(p^c \text{ first}| \; p^a,p^c) = 1 - \Pr(p^a \text{ first}| \; p^a,p^c)$.

[20] Since signals are conditionally independent, $\mathbb{E}\left(p|p^a\right)$ does not depend on the value of $p^c$ and $\mathbb{E}\left(p|p^c\right)$ does not depend on the value of $p^a$. Note that if we modeled deliberation in continuous time, we would have (for many processes) $\mathbb{E}\left(p|p^a\right) = p^a$, and $\mathbb{E}\left(p|p^c\right) = p^c$.

$V^1\left(p|\underline{p},\overline{p}\right)$ is the value of continuing at least one period. The overall value function is given by:

$$V(p \mid \underline{p},\overline{p}) = \begin{cases} \max\left\{V^1\left(p|\underline{p},\overline{p}\right), -(1-q)\,p\right\} & p \leq \underline{p} \\ V^1\left(p|\underline{p},\overline{p}\right) & \underline{p} < p < \overline{p} \\ \max\left\{V^1\left(p|\underline{p},\overline{p}\right), -q\,(1-p)\right\} & p \geq \overline{p} \end{cases}. \tag{4}$$

The interpretation of this expression is the following: for $p \leq \underline{p}$, the juror chooses the best option between continuing deliberation and stopping to acquit. For $\underline{p} \leq p \leq \overline{p}$, the juror can only continue deliberation. For $p \geq \overline{p}$, the decision maker chooses the best option between continuing deliberation and stopping to convict.

**Lemma 1 (Convexity for Constrained Problem)** *The continuation value function $V^1\left(p|\underline{p},\overline{p}\right)$ of the constrained problem is convex for $p \in [0,\underline{p})$,* and for $p \in (\overline{p},1]$.

Convexity of the continuation value for the constrained problem follows through similar (standard) arguments to those used for the unconstrained problem: again, the logic is that information about $p$ is useful to the decision maker. As before, convexity implies that the solution is determined in a similar manner to that described through Figure 2 and uniqueness of the constrained solution $\left(p^a\left(\underline{p},\overline{p}\right), p^c\left(\underline{p},\overline{p}\right)\right)$ follows. The following lemma illustrates that the constrained solution is monotonic in the imposed thresholds.

**Lemma 2 (Monotonicity)** *$p^a\left(\underline{p},\overline{p}\right)$ is increasing in $\overline{p}$ and $p^c\left(\underline{p},\overline{p}\right)$ is increasing in $\underline{p}$.*

Monotonicity of the best responses is intuitive. Indeed, suppose one of the thresholds, say $\overline{p}$, increases. The consequence is that the problem facing the decision maker is more constrained, and so the continuation value decreases. By referring to Figure 2, when the continuation values decreases, the left intersection point increases, implying that $p^a\left(\underline{p},\overline{p}\right)$ increases.[21] Similar arguments hold for changes in $\underline{p}$.

An immediate consequence of Proposition 2 and Lemma 2 is the following (recalling that best responses in the constrained problem depend on the underlying preference parameter and information cost, which we do not spell out for readability's sake):

**Corollary 1 (Comparative Statics of the Constrained Problem)** *For any fixed $\left(\underline{p},\overline{p}\right)$, $p^a\left(\underline{p},\overline{p}\right)$ and $p^c\left(\underline{p},\overline{p}\right)$ weakly increase in $q$.*

[21] There is a subtlety because convexity of the value function is only guaranteed in a subinterval. However, as showed in the proof of Lemma 1, this is the relevant interval for determining the stopping thresholds.

### 3.3 Heterogeneous Preferences

We now shift our attention to juries composed of agents with potentially heterogeneous preferences. Namely, we assume $q_1 \leq q_2 \leq ... \leq q_n$ and allow for some of the inequalities to be strict. In order to isolate the effects of preference heterogeneity on outcomes, we assume that deliberation costs are homogenous and fixed at $k > 0$.[22]

Note that there are many possible sources of multiplicity in an environment such as ours: first, voting games are known to display equilibria in weakly dominated strategies in which voters vote in consensus independently of their preferences because unilateral deviations cannot affect outcomes (when voting rules are non-unanimous); second, there is a potential multiplicity linked to the infinite horizon nature of the game. Throughout the paper, we focus on equilibria satisfying two requirements:

1. Each juror's strategy is characterized by two time-independent thresholds for the posterior, $p^a \leq 1/2 \leq p^c$. That is, each agent chooses to stop deliberation whenever the timed posterior $p_t$ (that the defendant is guilty) satisfies $p_t \leq p^a$ or $p_t \geq p^c$.[23] In particular, strategies are Markov in the posterior $p_t$.

2. Sincere voting at any deliberation stage as well as the decision stage. Sincere voting at the decision stage requires that, when the deliberation stage ends with a posterior of $p$, a juror of preference parameter $q$ votes $a$, or to acquit, whenever $p < q$, and votes $c$,or to convict, whenever $p \geq q$. Sincere voting at the deliberation stage implies the following. At each period, each juror takes other jurors' equilibrium behavior as given and therefore internalizes the regions of posteriors in which she can affect deliberation choices. The juror then calculates her (constrained) continuation values from carrying on with deliberations as well as the expected payoff from stopping and votes for the alternative that provides the higher payoffs.[24]

[22] We return to the case of heterogeneous costs in Section 8.2.

[23] Recall that the prior probability that the defendant is guilty is 1/2. In particular, choosing a threshold $p^a \geq 1/2$ (or $p^c \leq 1/2$) would lead to no information collection. Our assumption that $p^a \leq 1/2 \leq p^c$ is therefore without loss of generality.

[24] Behavior according to (stationary) thresholds as well as sincerity could be thought of as the consequence of the following refinement to Markov equilibria in our setting. Consider finite horizon truncations of our game and focus on subgame perfect equilibria that survive iterated elimination of weakly dominated strategies in which last-date strategies are eliminated first. The limits of such sequences of equilibria as the horizon becomes infinitely large correspond to individual threshold strategies satisfying our sincere voting restriction.

For homogeneous committees, the above requirements boil down to selecting the Pareto-best equilibrium, which yields the same allocation as the single-person optimum. For readability purposes, and slightly abusing terminology, we will simply refer to the corresponding set as the set of 'equilibria' (or 'equilibrium' when unique).

We start by considering the case in which voting rules in the deliberation and decision stage coincide. That is, $R_d = R_v$. This allows us to focus on one set of pivotal agents, rather than consider pivotal agents at each stage of the decision-making process. It is also a case that is relevant for many applications: for example, it fits environments in which deliberation manifests as an opinion poll. Later, we inspect the impacts of discordance between the two types of rules.

From the point of view of a generic juror $i$, there are three possible cases:

(i) An $R_d$-majority of other jurors favors stopping. In this case deliberation stops, and juror $i$ has no impact on the decision.

(ii) Strictly fewer than $R_{d-1}$ other agents want to stop. In this case, deliberation must continue and juror $i$ is in a constrained region in which deliberation continues regardless of her vote.

(iii) Exactly $R_{d-1}$ other jurors want to stop. In this case juror i can affect the stopping decision.

As mentioned at the beginning of our presentation of the constrained problem in the previous subsection, because of heterogeneity in preferences, for different regions of the posterior, different agents are potentially keen to stop or continue deliberation.

The main complication, relative to the homogeneous jury setting, arises from the fact that the juror has no control over stopping in case (ii).[25] This means that the juror's optimal action in region (iii) depends on the magnitude of the region of posteriors over which case (ii) arises. From the perspective of this juror, we can see her problem as being a constrained optimization problem where she takes as given the fact that she cannot stop in a certain region.

Intuitively, when posterior probabilities of guilt are low, it is the jurors who care most about the mistake of convicting the innocent that determine the decision. Note that whenever agent $j$ prefers to continue deliberation, so does any agent $l > j$ who worries even more about innocent convictions. In particular, whenever juror $j = R_d$ chooses to continue deliberation, or vote to convict, so will all jurors $l > R_d$, and deliberation will carry on. In other words, when guilt posteriors are low, juror $R_d$ is 'pivotal'. Similarly, whenever posterior probabilities of guilt are high, it is the jurors who

[25] Region (i) will turn out to be irrelevant because, whenever this case arises, the juror is not pivotal.

worry most about guilty acquittals that determine decisions, the relevant pivotal juror being juror $n - R_d + 1$. Therefore, as in the case of the median voter theorem, we can focus on a small number of decisive people. Namely, in our model there are typically **two** decisive voters.

Lemma 3 below implies that the pivotal agent for stopping when $p_t < 1/2$ is juror $R_d$ and the pivotal agent for stopping when $p_t > 1/2$ is juror $n - R_d + 1$. In order to make the comparison with the results pertaining to homogeneous committees transparent, we focus on equilibria that are characterized by stationary thresholds.

**Lemma 3 (Reduction to Two Juror Juries)** *When $R_d = R_v$, any equilibrium thresholds corresponding to a jury composed of jurors with preference parameters $q_1 \leq q_2 \leq ... \leq q_n$ are also equilibrium thresholds of a jury composed of two jurors with preference parameters $q_{n-R_d+1}, q_{R_d}$ in which both deliberation and decision rules are unanimous.*

Best responses of the pivotal agents can be derived through optimization of the constrained decision problem with the value function from equation (4). Each of the agents takes one of the thresholds as given and optimally chooses the other one (that corresponds to the region she cares more about). We denote by $p_*$ the lower **equilibrium** threshold and by $p^*$ the upper **equilibrium** threshold. Lemma 3 implies that $p_* = p^a(p_*, p^*; q_{n-R_d+1})$, and $p^* = p^c(p_*, p^*; q_{R_d})$.[26] Note that these considerations imply that, for sufficiently low costs $k$, threshold equilibria exist.

In what follows, we move away from the assumption that $R_d = R_v$ and inspect the consequences of different deliberation and decision rules in general juries.

## 4. Arbitrary Deliberation and Voting Rules

We now consider a jury composed of $n$ jurors of arbitrary preferences $q_1 \leq q_2 \leq ... \leq q_n$ and contemplate differing constellations of voting rules. When $R_d \neq R_v$, there are two sets of relevant pivotal agents: those pertaining to the deliberation stage and those pertaining to the decision stage.

In analogy to Lemma 3, during the decision stage, whenever juror $j$ would prefer to convict if she were dictator, so would any juror $l < j$. Whenever juror $j$ would prefer to acquit if she were dictator, so would any juror $l > j$. It follows that the jurors to focus on are those pivotal during deliberation: jurors $R_d$ and $n - R_d + 1$, and those pivotal during the decision stage: jurors $R_v$ and $n - R_v + 1$.

[26] In order to stress the dependence on individual preferences, we use the natural notation $p^a(\underline{p}, \bar{p}; q)$ and $p^c(\underline{p}, \bar{p}; q)$ to denote the constrained solution for the agent of taste parameter $q$.

We first analyze environments in which deliberation costs are low. In such cases, equilibrium behavior will entail a high volume of information collection. This would suggest that in equilibrium, when information collection ends, agents would be at a consensus on what should be done. Formally,

**Lemma 4 (Low Costs – Convergence of Opinions)** Consider a jury $q_1, ..., q_n$. *For any* $R_d$ *and* $R_v$, *and any* $p_1 < p_2 \in (0, 1)$, *there is a sufficiently low deliberation cost* $\widehat{k}$ *such that, for* $k < \widehat{k}$, *equilibrium thresholds* $(p_*, p^*)$ *are such that* $p_* < p_1$ *and* $p^* > p_2$.

Lemma 4 implies that any set of jurors would agree on the decision ex-post when costs are sufficiently low even if unanimity is not a requirement for making decisions.[27]

Furthermore, Lemma 4 implies that jurors may change their stand both individually and collectively (say, as would be reflected in a poll) during deliberation. Indeed, let $p_1 = q_1$ and $p_2 = q_n$ and assume that $q_1 < 1/2$ and $q_n > 1/2$. From the lemma, for sufficiently small deliberation costs, deliberation continues even in instances in which all agents would agree to acquit ($p_* < p_t < q_1$) or all would agree to convict ($q_n < p_t < p^*$) had deliberation been forced to end. This, of course, means that if repeated polls were taken during the deliberation phase, shifts in preferred verdicts could be observed. This is, of course, a direct consequence of our modeling deliberation as learning, but it is consistent with some of the empirical research on jury deliberation processes. For example, Hans (2001, 2007) as well as Hannaford, Hans, Mott, and Musterman (2000) documented frequent opinion changes during deliberations in a variety of trials across the U.S.

The lemma also implies that for sufficiently low costs, deliberation will lead to consensus on the decision to be taken: the jury will deliberate for an amount of time that is such that the only possible posteriors are sufficiently extreme that they are either to the left of $q_1$ or to the right of $q_n$. In particular, the voting rule $R_v$ in the decision stage does not matter, and only the deliberation rule determines the length of the deliberation process.

There is another class of cases in which the voting rule $R_v$ does not affect outcomes. Fix a deliberation rule $R_d$ that is a strict super-majority. Then, any decision rule that is at most as consensual as $R_d$ ($R_v \leq R_d$) leads to the same equilibrium outcomes. The reason is fairly mechanical: whenever there is a sufficient super-majority to halt deliberation, there must be at least the same

[27] This result relies on the fact that jurors care about both types of mistakes: convicting the innocent and acquitting the guilty, so that $\tilde{q}_1, \tilde{q}_2 \neq 0, 1$. Naturally, as $\tilde{q}_1$ approaches 0 or $\tilde{q}_2$ approaches 1, the costs of assuring ex-post concensus approach 0.

super-majority at the decision stage. This is just the Wald logic: thresholds for stopping deliberation are inherently more demanding than the threshold ($q$) for taking a decision. Thus, equilibrium outcomes do not depend on $R_v$ in this region.

The following proposition summarizes our discussion of the two cases in which the decision rule has no effect on final outcomes.

**Proposition 3 (Decision Rule Irrelevance)** *For any deliberation rule $R_d$,*

1. *(Restricted Irrelevance) For any decision voting rules $R_v$, $\tilde{R}_v \leq R_d$, the set of equilibrium outcomes corresponding to $R_v$ and $\tilde{R}_v$ coincide.*

2. *(Irrelevance due to unanimous agreement: low costs) For any given preference profile, there exists a $\underline{k}$ such that, for $k \leq \underline{k}$, the voting rule at the decision stage $R_v$ is irrelevant for equilibrium outcomes. In particular, time to decision or probability of mistakes do not depend on $R_v$.*

Proposition 3 outlines two important cases in which the voting rules at the decision stage do not matter. This is consistent with the empirical and experimental evidence, discussed in the Introduction, which points to collective outcomes being insensitive to the decision rule in place when communication is available (see, e.g., Baldwin and McConville, 1980, Devine et al., 2001, and Goeree and Yariv, 2011).

There are, however, cases in which the voting rule at the decision stage is important for outcomes. Consider a case in which $R_v > R_d$ and there is a relatively high cost of deliberation $k$. Had the pivotal jurors at the deliberation stage (agents $n - R_d + 1$ and $R_d$) ignored the fact that more extreme jurors (agents $n - R_v + 1$ and $R_v$) are pivotal at the decision stage, there would potentially be disagreement in the decision stage and the jury would end up as hung. Thus, in these cases, the pivotal jurors at the deliberation stage face a trade-off: they can either prolong deliberation to convince the pivotal jurors in the decision stage, or they can halt deliberations immediately. If the costs of deliberation are not too high, some additional deliberation may therefore be beneficial. Intuitively then, in these cases one should expect that a unanimity voting rule at the decision stage will lead to longer deliberation and more accurate decisions than simple majority. We show below that this is true for juries with symmetric preferences (around 1/2). When juries are asymmetric, changing the decision rule or the deliberation rule may not lead to uniformly more accurate decisions because, for instance, more accurate acquittal decisions may come hand in hand with less accurate conviction decisions: it can be the case that, say, making the decision rule more extreme reduces the acquittal equilibrium threshold $p_*$ but also reduces the equilibrium conviction threshold $p^*$.

## 5. Symmetric Juries

For simplicity, we first go back to the case $R_d = R_v$. Lemma 3 allows us to restrict attention to two jurors within the jury with preferences: $\tilde{q}_1 = q_{n-R_d+1}$ and $\tilde{q}_2 = q_{R_d}$. Assuming $\tilde{q}_1$ and $\tilde{q}_2$ are symmetric around $\frac{1}{2}$, i.e., $\tilde{q}_1 = \frac{1}{2} - b$ and $\tilde{q}_2 = \frac{1}{2} + b$ for some $b \in \left[\frac{1}{2}, 1\right]$, simplifies equilibrium characterization significantly.

**Definition (Symmetry in Juries)** *We say the jury is quasi-symmetric with respect to $R_d = R_v$ whenever $q_{n-R_d+1}+q_{R_d} = 1$ and information is symmetric, i.e., for any $s \geq 0$, $F_G(s) = F_I(-s)$. A jury is* symmetric *whenever it is quasi-symmetric with respect to all voting rules.*

When juries are quasi-symmetric, we focus on *symmetric threshold equilibria* corresponding to the relevant deliberation rule, ones in which both posterior thresholds are symmetric around 1/2 (i.e., equally distanced from 1/2).[28] As it turns out, quasi-symmetric juries generate unique predictions, established in the following lemma.

**Lemma 5 (Quasi-symmetric Juries - Uniqueness)** *Assume the jury is quasi-symmetric with respect to $R_d$. Then, there exists a unique stationary symmetric threshold equilibrium.*

Lemma 5 implies that when the jury is symmetric and costs are sufficiently low, symmetric equilibrium thresholds are determined uniquely for any decision rule. The lemma is a direct consequence of the monotonicity implied by Lemma 2. Indeed, if there were two threshold equilibria, the ranking of the left thresholds must coincide with the ranking of the right thresholds. Therefore, it is impossible for two such equilibria to both be symmetric.

We start our analysis with juries that are quasi-symmetric with pivotal jurors as above, $\tilde{q}_1 = \frac{1}{2} - b$ and $\tilde{q}_2 = \frac{1}{2} + b$ for some $b \in \left[0, \frac{1}{2}\right]$. This allows us to identify the impact of diversity in the jury, as captured by the spread $b$, and opens the door for inspecting the effects of the voting rules, which determine how moderate or extreme the pivotal jurors are.

[28] For the next results, we will consider cases where the threshold equilibrium is non-trivial in that it entails some amount of information collection. This is the relevant case when $k$ is not too high, but it excludes the case of hung juries which will be discussed later.

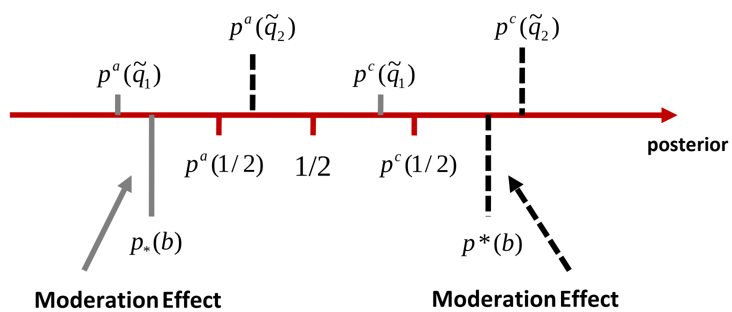


Figure 3: Equilibrium Spread and the Moderation Effect in Quasi-symmetric Juries

Denote the resulting symmetric equilibrium thresholds by $p_*(b) \leq 1/2 \leq p^*(b)$ (where we drop the cost and voting rule arguments for ease of presentation). Symmetry entails $p_*(b) + p^*(b) = 1$.

As $b$ increases, the juror with preference parameter $\tilde{q}_1$ is increasingly concerned about acquitting the guilty, while the juror with preference parameter $\tilde{q}_2$ is increasingly concerned about convicting the innocent. There are now two forces at play. The direct one is that the first agent would like to spend more time collecting information when the posterior is lower than 1/2, while the second agent would like to spend more time collecting information when the posterior is greater than 1/2. Indeed, this follows from the first part of Proposition 2, implying that (dropping the explicit dependence on $k$) $p^a(\tilde{q}_1) \leq p^a(\frac{1}{2}) \leq p^a(\tilde{q}_2) \leq 1/2 \leq p^c(\tilde{q}_1) \leq p^c(\frac{1}{2}) \leq p^c(\tilde{q}_2)$. The indirect effect comes from the strategic interaction. Consider, say, the first juror and the event in which the posterior $p_t \in (p^a(\tilde{q}_1), p^a(\tilde{q}_2))$, so that if she were by herself she would continue collecting information, while the other juror by herself would not. Importantly, the continuation value for pursuing information collection is lowered by the existence of the other juror. Indeed, juror 1 knows that when $p_t > 1/2$, juror 2 will demand longer deliberation than juror 2 would like (in the analogous range $p_t \in (p^c(\tilde{q}_1), p^c(\tilde{q}_2)))$. Thus, continuation values are lower, leading to equilibrium thresholds that are moderate relative to the most extreme individual thresholds $p^a(\tilde{q}_1), p^c(\tilde{q}_2)$, as depicted in Figure 3 and formalized in the following proposition.[29]

**Proposition 4 (Equilibrium Spread and Moderation)**

1. *(Spread)* $p_*(b)$ *is decreasing in* $b$, $p^*(b)$ *is increasing in* $b$. *In particular, the time it takes for a decision is increasing in* $b$.

2. *(Moderation Effect) For any* $q = \frac{1}{2} + b$, *where* $b \in \left(0, \frac{1}{2}\right)$, $p_*(b) = 1 - p^*(b) \leq p^a(q)$ *and* $p^*(b) \leq p^c(q)$. *Furthermore, for sufficiently small costs* $k$, *these inequalities are strict.*[30]

Part 1 of the proposition implies that increased heterogeneity in the jury (manifested in a higher $b$) will reduce the two types of mistakes, and increase the expected time to a decision. This is consistent with the empirical observations of Sommers (2006), who used a mock juries to test for the effects of racial heterogeneity on jury performance and found that heterogeneity was associated with longer deliberation and more accurate decisions. In a similar vein, Goeree and Yariv (2011) found increased preference heterogeneity to be associated with longer deliberation times in the laboratory.

The spread of symmetric pivotal agents can be manipulated through the voting rule. Indeed, the more demanding the deliberation rule (a higher $R_d = R_v$), the greater the spread. Formally, note that $q_{R_d} - q_{n-R_d+1}$ is increasing in $R_d$. Using part 1 of Proposition 4, we therefore get the following corollary.

**Corollary 2 (Accuracy and Deliberation Rules)** *In symmetric juries, deliberation length and accuracy of decisions increase with the deliberation and voting rules* $R_d = R_v$.

We now discuss the effects of decision rules beyond the cases of irrelevance highlighted in Proposition 3.

As it turns out, when the decision rule $R_v$ is more consensual than the deliberation rule $R_d$, $R_v \geq R_d$, non-trivial symmetric equilibrium thresholds are still determined uniquely. Essentially, there are two cases to consider. First, when costs of deliberation are low, the equilibrium deliberation thresholds are sufficiently wide that the super-majority requirement at the decision stage is automatically met and exceeded. In the second case, with higher costs and $R_v > R_d$, the pivotal

[29] Part 2 of the Proposition is related to Strulovici (2010), who calls a similar phenomenon in an experimentation environment a "loss of control effect". It is also related to Chan and Suen (2011), who illustrate in a collective learning environment akin to ours that an impatient member can lead a more patient committee to hasten decisions.

[30] A symmetric statement holds for $q = \frac{1}{2} - b$.

jurors at the deliberation stage would like to settle for deliberation thresholds $p_* > q_{n-R_v+1}$ and $p^* < q_{R_v}$. However, such thresholds would lead to a hung jury. In this scenario, the equilibrium deliberation thresholds are driven by the requirement to reach sufficient consensus at the decision stage so we obtain $p_* = q_{n-R_v+1}$ and $p^* = q_{R_v}$: deliberation continues until the moment the pivotal jurors at the decision stage are persuaded to join the required consensus. When costs are sufficiently high, achieving a consensus at the decision stage becomes too costly for the pivotal jurors at the deliberation stage and the unique equilibrium outcome involves no information collection and an immediate hung jury.

In the following proposition, we denote by $(p_*(R_d, R_v; k), p^*(R_d, R_v; k))$ the unique symmetric equilibrium thresholds corresponding to deliberation rule $R_d$ and decision rule $R_v \geq R_d$.

**Proposition 5 (Decision Rule Relevance: Inclusiveness Effect)** *Consider a symmetric jury. For any deliberation rule $R_d$, take two voting rules $\tilde{R}_v > R_v \geq R_d$.*

1. *There exist $\underline{k}, \overline{k}$ such that, for $\underline{k} < k < \overline{k}$, corresponding symmetric equilibrium thresholds satisfy $p_*\left(R_d, \tilde{R}_v; k\right) < p_*(R_d, R_v; k)$ and $p^*\left(R_d, \tilde{R}_v; k\right) > p^*(R_d, R_v; k)$: the larger the supermajority required for making a decision, the more information collection there is; the deliberation time and decision accuracy are greater under voting rule $\tilde{R}_v$ than under voting rule $R_v$.*

2. *There exists $\overline{k}$ such that, for $k < \overline{k}$, corresponding symmetric equilibrium thresholds satisfy $p_*\left(\tilde{R}_d, \tilde{R}_v; k\right) < p_*\left(R_d, \tilde{R}_v; k\right)$ and $p^*\left(\tilde{R}_d, \tilde{R}_v; k\right) > p^*\left(R_d, \tilde{R}_v; k\right)$: deliberation time and accuracy would be even greater under deliberation rule $\widetilde{R}_d = \tilde{R}_v$*

Part 1 of Proposition 5 is in contrast with the results in Feddersen and Pesendorfer (1998), Persico (2004), and Austen-Smith and Feddersen (2006). The intuition for this result can be understood through a special case. Consider a symmetric jury in which there is a median agent with preference parameter $q = \frac{1}{2}$. Assume also that $R_d$ corresponds to simple majority. We contemplate the effect of moving from $R_v$ = simple majority to $R_v$ = unanimity. Suppose that costs of deliberation are sufficiently high that when $R_v$ corresponds to a simple majority, it is not worthwhile for the median juror to deliberate long enough to reach consensus on the decision. Then, under unanimity, the median juror who is still pivotal in the deliberation process understands that, in order to reach a verdict, she cannot stop deliberation as early as when $R_v$ = simple majority. In order to avoid a hung jury she must convince the extreme jurors to vote with everyone else. This requires longer deliberation. When costs are not too high, it is worthwhile to deliberate just long enough to obtain these jurors' votes on the decision.

Continuing the same example, Part 2 of Proposition 5 says that moving to $R_d =$ unanimity would lead to even longer deliberation. The reason is that the most extreme jurors are now in a position to directly affect the deliberation decision, and they desire longer deliberation than the median juror.

Taken together with Proposition 3, Part 2 of Proposition 5 suggests that deliberation rules are more powerful than decision rules in driving the process of jury decision making and deliberation.

These results imply that, in essence, interior equilibria *always* depend only on two jurors. When $R_v \leq R_d$, the jury outcome is equivalent to that of a jury composed of two jurors with preferences $q_{R_d}$ and $q_{n-R_d+1}$ (and unanimous deliberation and voting rules), while when $R_v > R_d$, any jury outcome entailing non-trivial deliberation is equivalent to that of a jury composed of two jurors with preferences $q_{R_v}$ and $q_{n-R_v+1}$ (and, again, unanimous deliberation and voting rules).

In practice (and outside our model), it may be the case that a change in the decision voting rule is tied to a change in the deliberation rule, a so-called *protocol effect*. In the presence of such a protocol effect, the decision voting rule has a clear impact. Indeed, when the voting and decision rules coincide, Lemma 3 holds, so that two pivotal jurors determine the outcome. The more demanding the decision voting rule, the more extreme these two jurors are. Consequently, more stringent decision voting rules would correspond to longer deliberation and more accurate decisions.

## 6. WELFARE AND DELEGATION

Welfare effects of deliberation rules and decision rules depend on the perspective from which welfare is calculated. One point of view is to only consider the ex-ante welfare of the jury. Another is to include the benefit of accurate decisions for society at large, including individuals who do not directly bear the cost of deliberation.

First consider a homogeneous jury. From the point of view of the agents, deliberation is weakly beneficial. Indeed, the jury can always choose not to deliberate by fixing the thresholds at the prior $1/2$, $p^a = p^c = 1/2$. From an institutional point of view, when deliberating groups are homogeneous, a designer (say, the constitution writers) characterized by preference parameter $q$, who internalizes the costs (e.g., when these costs are linked to the time spent on making decisions and not engaging in other profitable activities) is best off with a committee (jury) comprised of identical agents of preference parameter $q$ as well. In fact, a committee composed of more extreme agents than the designer would entail "too much" information collection. The designer may then benefit by increasing the costs experienced by the committee members, or putting a cap on deliberation time.

From the perspective of the participating jurors, in any quasi-symmetric jury, we can assess the optimal spread of the pivotal agents. It turns out that little spread is most preferred, as captured by the following proposition.

**Proposition 6 (Optimal Delegation)** *Jurors have unanimous preferences over deliberation rules: all jurors in a symmetric jury prefer pivotal agents with as little spread as possible or* $R_d = \lceil n/2 \rceil$.

Intuitively, recall expression (3) for a juror's utility. In a symmetric jury, thresholds are symmetric, and therefore, the first two terms in (3) are a convex combination (via $q_i$) of an identical expected probability of mistake. It follows that the expected utility does not explicitly depend on $q_i$. In particular, all of the jurors gain the same level of expected utility as would a juror with preference parameter $\frac{1}{2}$ if she were to have the equilibrium thresholds imposed upon her. However, note that a juror with preference parameter $\frac{1}{2}$ would prefer no spread at all ($b = 0$ in our notation above), as then she receives her optimal thresholds. Monotonicity then implies our result.[31]

Proposition 6 is particularly stark because of symmetry. However, the effect highlighted in this proposition is more general: in a large class of asymmetric juries, the most extreme jurors will not push for unanimity at the deliberation stage because unanimity means that deliberation is long on *both* sides of the prior, making the cost of deliberation too high from an ex-ante perspective to make it worth reducing the probability of mistakes further.

It is also useful to point to a contrast between decision rules and deliberation rules at this point. Proposition 6 provides a rationale for giving moderate jurors control over the deliberation phase of the trial. This could be done by the implementation of rules requiring a small majority (optimally, a simple majority rule) or by delegating the decision to a moderate chairman (optimally, with preference parameter $q = 1/2$). Nonetheless, at least some jurors may prefer a more stringent decision rule to determine the verdict. Indeed, no juror would willingly give up her voice in the

[31] Proposition 6 hints at the possible effectiveness of deliberation taxes. Indeed, increasing the costs of deliberation would lead to shorter deliberation times, which may be preferrable to at least a fraction of the population.

decision phase and choose a decision rule that would end up excluding her. In particular, we cannot expect the same level of agreement on decision rules as the one achieved over deliberation rules.

## 7. Simultaneous Deliberation

We now discuss a case in which the decision on the amount of information to be collected takes place in one shot and contrast this case with the sequential one considered up to now. When committee members are homogeneous, this is equivalent to the classic case of choosing the optimal sample size for the test of a binary hypothesis (see De Groot 1970, Chapter 11). We retain the jury language although the jury setting is no longer the natural application for this case. Nonetheless, there are many environments in which the size of the sample is determined at the outset. For example, drug companies decide on the sample size of patients at the beginning of many drug trials, academic departments decide on the number of outside recommendation letters at the start of most promotion processes, etc.[32]

In our version with heterogeneous jurors we need to specify some additional details of the model. A deliberation decision determines the sample size $t$. A sample of size $t$ costs each juror $kt$. This is the only cost born by the committee. At time $t$, jurors observe the realization of the sequence of random variables $X_1, ..., X_t$, and they vote to acquit or convict according to a decision rule $R_v$ just as in Section 2. Deliberation can be modeled in a number of ways but, for concreteness, we assume the following process. Deliberation takes place at no cost before the sample is drawn according to deliberation rule $R_d$. An index moves over discrete time starting from 1. At index $\tau$, if jurors have not yet come to an agreement, then jurors vote on whether sample size $\tau$ is acceptable. If at least $R_d$ jurors agree that the sample size is sufficient, then the deliberation process is over and a sample of size $\tau$ is drawn. If fewer than $R_d$ jurors agree, then the index moves on to $\tau+1$. The process continues until an $R_d$ majority is satisfied. This model would be identical to our sequential deliberation model if voters had to stop deliberation without seeing the realizations of the random variables. Given our result below that deliberation is always unanimous, the exact deliberation protocol is irrelevant for symmetric juries.[33] However, the model described above is easier to work with and is a closer match

[32] It would, of course, be interesting to consider a mixed model, where decisions are partly sequential, partly simultaneous. In such a setting, at each date the commitee decides on a sample size, but the sample size can be augmented later. Moscarini and Smith (2001) provide an analysis of this problem for the single agent case.

[33] We could allow for a once and for all deliberation vote over the infinite set of alternatives corresponding to different sample sizes. We intentionally set this version of the model in the framework of binary decisions as it captures the same issues conceptually and allows us to naturally circumvent standard issues of equilibrium multiplicity that arise in voting models with more than two alternatives.

to the sequential deliberation model.[34]

Let $p_t$ be the posterior if the deliberation process has yielded a sample size $t$. Then, at date $t$, a juror of type $q$ votes to convict if $p_t \geq q$, and votes to acquit if $p_t < q$. If at least $R_v$ votes are obtained, then a decision is reached. Otherwise we have a hung jury.[35]

**Proposition 7 (Simultaneous Deliberation: Voting Rule Relevance)** *In a symmetric jury, under simultaneous deliberation, jurors have common preferences over deliberation decisions. Therefore, the deliberation rule $R_d$ is irrelevant. However, the voting rule matters: if $\widetilde{R}_v > R_v$, a jury voting under voting rule $\widetilde{R}_v$ chooses to collect more information.*

The intuition for the irrelevance of the deliberation rule is related to the intuition of Proposition 6. Given that deliberation is simultaneous, jurors evaluate the optimal amount of information to be collected ex-ante, before seeing the realization of any signals. All jurors simply trade off increased accuracy against the cost of information collection independent of their preference parameters because increased information collection reduces mistakes of both types equally.

The intuition for the effect of the voting rule is the following. Under simultaneous deliberation, for any given amount of gathered information, a larger voting rule raises the probability of a hung jury. Acquiring additional information reduces the probability of this costly event.

This result is in sharp contrast with the results we obtained for the case of sequential deliberation. This is not very surprising given the very different nature of deliberation in the two scenarios.

The contrast between the consequences of simultaneous as opposed to sequential protocols is also familiar from the literature on search and auctions. Note also that, in contrast with the case of sequential deliberation, even with symmetry, the simultaneous scenario allows for the coexistence of significant information collection and hung juries.

There are a number of interesting additional comparisons that can be made between the sequential and simultaneous scenarios. First, there is a strong general effect leading to welfare being higher under sequential deliberation: welfare is obviously unambiguously higher in the sequential case for

[34] As in the previous analysis, there are possible multiple equilibria. As before, we can think of a refinement where one considers finite truncations of the game. Equilibria surviving iterated elimination of weakly dominated strategies in the agent-form game would correspond to (timed) thresholds, and that sequence of thresholds converges to the thresholds we analyze here as the horizon of the game grows indefinitely.

[35] Recall that we maintain the assumption that in the event of a hung jury, the defendant is acquited or convicted with equal probability. The proposition would, in fact, hold as long as the ex-ante expectation of the costs of a hung jury are independent of $q$. For instance, a fixed cost of a hung jury independent of $q$ would generate similar insights.

the case of a single juror because more efficient use of information is made. Indeed, this was the original motivation behind Wald's analysis. By Proposition 6 we can also conclude that welfare from the point of view of the committee is higher in the sequential case under simple majority when juries are symmetric.

Another interesting comparison concerns the likelihood of consensual votes. For low costs, sequential deliberation leads to unanimous verdicts regardless of the decision rule (see Proposition 3). In contrast, simultaneous deliberation generates a positive probability of some disagreement for any voting rule, for any costs, because there are always positive probability sets of signals that are not very informative. For intermediate costs the comparison is less straightforward, but, for any fixed voting rule, simultaneous deliberation tends to generate more variation in consensus because in the case of sequential deliberation the vote is more likely to end at a quorum.

## 8. Discussion and Extensions

**8.1. Random Juries.** In our model, jurors' preferences are given. A natural extension would allow some randomness in the preferences of the selected jurors.[36] Suppose that, at the outset, the opposing lawyers and the judge select a jury composed of individuals with preferences drawn independently from a distribution $G$ over $[0,1]$.[37] For simplicity, assume that: 1. The population is ex-ante symmetric, so that $G(x) = 1 - G(1-x)$, and 2. Jurors, once selected, are transparent about their preferences, so that they are common knowledge.

In order to illustrate the effects of such randomness on outcomes and welfare, consider the case in which $R_d = R_v$, so that Lemma 3 holds. It follows that the relevant preferences to consider are the $R_d$'th and $n - R_d + 1$'th *order statistics* of the sampled jury preference profile. In particular, if we compare two populations, characterized by distributions $G$ and $G'$, with one more variant than the other, so that $G'$ is a mean preserving spread of $G$, the expected order statistics will be more extreme under $G'$ than under $G$.

Even though Lemma 3 holds and, for any selected jury, it is only two jurors who effectively determine outcomes, the size of the jury now plays an important role as well since it affects the

[36] In some cases, the process of voir dire in commonwealth countries as well as the U.S. effectively restricts preference profiles of juries. Nonetheless, the process cannot pick out fully the characteristics of jurors and so some randomness remains. This is true for many other collective decision processes in which the agenda is set for several generations of agents, so that rules are not tailored to a particular familiar committee.

[37] We can think of the institutional designers as the constitution writers, who put agendas in place having only a distribution of cases and juries in mind.

variance of the preferences of these two pivotal jurors. In that respect, it would be important to understand the curvature of the (constrained) best-response thresholds. This would be especially important if one considered agenda setters that experience some level of risk aversion (which, in our baseline model, plays no role). In such settings, it is the interplay between the size of the jury and the voting rule that determines the distribution of outcomes.

For sufficiently large $n$, juries will be approximately symmetric. Therefore, using the results of Proposition 4, we make two conjectures. First, more stringent voting and deliberation rules will generate more extreme pivotal jurors and therefore lead to greater expected times to decisions and smaller expected probabilities of mistakes. Similarly, fixing the voting rule and contemplating a distribution $G'$ that is a mean preserving spread of $G$, would yield more extreme pivotal jurors and analogous effects on timing and accuracy of outcomes.

**8.2. Heterogeneous Deliberation Costs.** Suppose now that jurors differ in the costs that are imposed upon them through deliberation (e.g., if costs are linked with the time away from work, variance in wages may translate to variance in deliberation costs). Formally, in order to assess the effects of cost heterogeneity, we assume that all jurors share the same preference parameter $q$, but juror $i'$s deliberation cost is given by $k_i$, where without loss of generality $k_1 \geq k_2 \geq ... \geq k_n$.

Note that the decision rule $R_v$ does not affect outcomes since for any given posterior the jurors all agree on the optimal action to be taken. The voting rule, however, does have an effect. Whenever agent $j$ wants to stop information collection, so does any agent experiencing higher costs ($l < j$). It follows that the pivotal juror during deliberation is the $R_d$'th juror. Consequently, we get the following.

**Proposition 8 (Heterogeneous Costs)** *A jury with rule $R_d$ chooses thresholds of a homogeneous committee with costs $k_{R_d}$ . Hence, deliberation length and accuracy of decisions increase with the decision rule $R_d$.*

Proposition 8 implies that a designer who does not internalize the jury's deliberation costs would be inclined to choose as demanding a deliberation rule as possible. The welfare optimal deliberation rule, however, depends on the distribution of waiting costs in the relevant population.

**8.3. Civil Juries.** We now consider an example of a related model, where a jury must take a decision from a continuum of possible choices. This can be interpreted as a model of a civil jury choosing the amount of damages to award a plaintiff. The jury is uncertain about the true level of damages $D$. The prior distribution over these true damages is given by a normal distribution $N(\mu, \frac{1}{p_0})$ with mean $\mu$ and precision $p_0$.

Suppose we allow a limited degree of heterogeneity among jurors that, as above, is given by how costly it is for them to continue collecting information (in the current setting, this will be tantamount to allowing heterogeneity in how strongly each juror desires to make the correct decision). Specifically, assume the payoffs for each juror if true damages are $D$ and the jury awards $Q$ are given by

$$U(D, Q) = -\alpha(D - Q)^2,$$

where $\alpha > 0$.

The jury deliberates as in the model presented in Section 2: at each deliberation date they observe a new signal $X_t$ at cost $k$. The sequence $X_1, X_2, ...$ is conditionally i.i.d., normal with mean $D$ and precision $p_X$ (i.e., for all $t$, $X_t \sim N(D, \frac{1}{p_X})$). The jury stops deliberating if at least $R_d$ jurors vote to stop, otherwise it continues. Note that, given the assumption about payoffs, once deliberation has ended, the jury is unanimous about the optimal decision. Thus, all disagreements arise in deliberation choices.

The optimal choice if the jury stops deliberating at $t$ is the conditional expectation of $D$ (equivalently, $X_{t+1}$) given the prior history. As is well known, in this normal quadratic setting, this conditional expectation takes a convenient form:

$$\mathbb{E}[X_{t+1} | X_1 = x_1, \ldots, X_t = x_t] = \frac{p_0\mu + p_X \sum_{s=1}^{t} x_s}{p_0 + tp_X}.$$

Thus, the payoff to a juror experiencing a cost $k$ when the jury stops at time $t$ is given by

$$-\frac{\alpha}{p_t} - kt = -\frac{\alpha}{p_0 + tp_X} - kt \tag{5}$$

Note that this payoff is independent of the realizations of $X_1, ...X_t$ so, in this setting, in contrast with our prior analysis, there is no difference between sequential and simultaneous deliberation.

From equation (5) we can immediately conclude that jurors with lower costs $k$ want to stop later and we obtain a similar result to our original one concerning the effects of deliberation rules:

**Proposition 9 (Deliberation in Civil Juries)** *The accuracy of damage awards is higher under more consensual deliberation rules.*

Note that, in equation (5), increasing the cost $k$ has similar effects to lowering the preference parameter $\alpha$. In particular, as mentioned above, cost heterogeneity plays a similar role to preference heterogeneity (when manifested through heterogeneous parameters $\alpha$ in the jury).

**8.4. Incomplete Information, Stationarity, and Hung Juries.** Throughout the paper, we have assumed that jurors' preferences are commonly known. This assumption allowed us to focus on stationary strategies and extract the main tensions between the deliberation and decision phases. Nonetheless, a natural extension to our model is to the case in which jurors have some incomplete information about the learning process at hand, either due to preferences that are not commonly known or due to the informativeness of the collective signals not being fully transparent. In such environments, the process of deliberation confounds two learning processes: regarding the guilt of the defendant, and regarding the prevailing characteristics of the jury (distribution of preferences or signal informativeness). In particular, the deliberation phase is inherently non-stationary.

While the analysis of such a model requires some novel techniques and goes beyond the scope of the current paper, we view it as especially important for explaining the patterns identified by the empirical literature regarding hung juries. Indeed, in such a model, it is conceivable that the longer deliberation goes on, the more likely it is that jurors are "high strung" or that the process is not very informative, and that agreement is likely to take a longer time than was initially estimated. Under certain additional conditions, this modification is likely to deliver that hung juries deliberate longer.[38] This would be in line with the evidence provided by Kalven and Zeisel (1966) and Hans (2003), who find that hung juries deliberate a significantly longer time than juries that deliver a verdict.[39] Furthermore, the tensions between costly information collection and decision accuracy, that are the driving force of our results, would persist in such a setting.

[38] An alternative assumption that we suspect would lead to a similar conclusion is that deliberation costs are increasing in time.

[39] In our baseline setting of complete information, a robust consequence is that even in cases of asymmetries, hung juries deliberate for a shorter time than juries that deliver a verdict. The reason is that agents can anticipate when a hung jury is likely to occur and thereby not invest in any information collection whatsoever.

## 9. Appendix

**Proof of Lemma 1**

Consider $p_1, p_2 \leq \underline{p}$. For both $p_1$ and $p_2$ the decision maker can choose to stop. In this case, $V^1\left(\alpha p_1 + (1-\alpha)\, p_2 \mid \underline{p}, \overline{p}\right) \leq \alpha V^1\left(p_1 \mid \underline{p}, \overline{p}\right) + (1-\alpha)\, V^1\left(p_2 \mid \underline{p}, \overline{p}\right)$ follows from the same argument as in the unconstrained case. Namely, consider an alternative world in which with probability $\alpha$, the posterior probability that the defendant is guilty in the period that follows is given by $p_1$ and with probability $1-\alpha$, that probability is given by $p_2$. If the agent is not told which of the two posteriors had been realized, she can guarantee the continuation value corresponding to $\alpha p_1 + (1-\alpha)\, p_2$. However, if she is told which of the two probabilities is realized, then with probability $\alpha$, she can guarantee the continuation value of $p_1$ and with probability $1-\alpha$ the continuation value of $p_2$. Since she can always ignore the information provided to her, in the latter case she must be gaining at least as much, and convexity follows. Similar arguments follow for $p_1, p_2 \geq \overline{p}$. ■

**Proof of Lemma 2**

For any $p, \underline{p}, \overline{p}$,

$$\max\left\{V^1(p \mid \underline{p}, \overline{p}), -q\,(1-p)\right\} \geq V^1(p \mid \underline{p}, \overline{p}).$$

Consider $\overline{p}_1 > \overline{p}_2$. From (4) it follows that for any $\underline{p}$, $V^1(p \mid \underline{p}, \overline{p}_1) \leq V^1(p \mid \underline{p}, \overline{p}_2)$. There are two cases to consider. In the first case, $p^a\left(\underline{p}, \overline{p}_2\right) < \underline{p}$, so that the solution is given by the intersection between the line $-(1-q)p$, which is decreasing in $p$ (see Figure 1) and the convex part of $V^1(p \mid \underline{p}, \overline{p}_2)$. If $p^a\left(\underline{p}, \overline{p}_2\right) = \underline{p}$, monotonicity follows. Suppose then that $p^a\left(\underline{p}, \overline{p}_1\right) < \underline{p}$. Since $V^1(p \mid \underline{p}, \overline{p}_1) \leq V^1(p \mid \underline{p}, \overline{p}_2)$, the intersection point with $V^1(p \mid \underline{p}, \overline{p}_1)$ must be (weakly) higher than that with $V^1(p \mid \underline{p}, \overline{p}_2)$, and the monotonicity of $p^a\left(\underline{p}, \overline{p}_2\right) \leq p^a\left(\underline{p}, \overline{p}_1\right)$ as needed. In the second case, $p^a\left(\underline{p}, \overline{p}_2\right) = \underline{p}$, and since $p^a\left(\underline{p}, \overline{p}_1\right) \leq \underline{p}$, monotonicity follows. Monotonicity of $p^c\left(\underline{p}, \overline{p}\right)$ follows analogously. ■

**Proof of Proposition 2**

**Part 1.** Suppose that $p^a \equiv p^a(q;k)$ and $p^c \equiv p^c(q;k)$ are the unique interior equilibrium thresholds for preference parameter $q$ and cost $k$. Directing our attention to $p^a$, it must be the case that first order conditions hold:

$$\frac{\partial U(q;p^a,p^c)}{\partial p^a} = -q\left(1-\mathbb{E}\left(p|p^c\right)\right)\frac{\partial \Pr(p^c \text{ first} \mid p^a,p^c)}{\partial p^a} -$$
$$-\left(1-q\right)\frac{\partial\left[\mathbb{E}\left(p|p^a\right)\Pr(p^a \text{ first} \mid p^a,p^c)\right]}{\partial p^a} - k\frac{\partial T(p^a,p^c)}{\partial p^a} = 0$$

Second order conditions must hold as well, so that $\frac{\partial^2 U(q;p^a,p^c)}{\partial (p^a)^2} < 0$ (the inequality is strict from uniqueness of the Wald solution).

Notice that, from the definition of $\Pr(p^x \text{ first}|\ p^a,p^c)$ and $\mathbb{E}\left(p|p^x\right)$ for $x = a,c$, it follows that $\frac{\partial \Pr(p^a \text{ first}|p^a,p^c)}{\partial p^a} \geq 0$, $\frac{\partial \Pr(p^c \text{ first}|p^a,p^c)}{\partial p^a} \leq 0$ and $\frac{\partial \mathbb{E}(p|p^a)}{\partial p^a} \geq 0$. Therefore, for sufficiently small $\varepsilon > 0$, $\frac{\partial U(q+\varepsilon;p^a,p^c)}{\partial p^a} \geq 0$.

From the second order condition, if follows that in order to satisfy the first order condition for $q+\varepsilon$ when we fix $p^c$, the threshold $p^a$ must increase.

From Lemma 2, we can iterate on best responses and get that the optimal threshold $p^a(q+\varepsilon;k) \geq p^a = p^a(q;k)$. Similar arguments follow for the threshold $p^c$.

**Part 2.** The cost $k$ does not affect the payoffs from stopping and taking a decision (the lines $-(1-q)p$, and $-q\left(1-p\right)$ in figure 1). However, as $k$ increases, the continuation value $V^1\left(p\right)$ decreases point-wise. Since $p^a(q;k), p^c(q;k)$, satisfy

$$V^1(p^a(q;k)) = -(1-q)p^a(q;k) \text{ and } V^1(p^c(q;k)) = -q(1-p^c(q;k)),$$

the comparative statics with respect to $k$ follows. ■

**Proof of Lemma 3**

In order to stress the dependence on preference parameters, we use $\left(p^a\left(\underline{p},\overline{p};q\right), p^c\left(\underline{p},\overline{p};q\right)\right)$ to denote the solution to the constrained problem (4) for an agent with preference parameter $q$. Consider then any candidate equilibrium defined by thresholds $p_*, p^*$ such that $\underline{p} = p^a\left(p_*,p^*;q_{n-R_d+1}\right)$, and $\overline{p} = p^c\left(p_*,p^*;q_{R_d}\right)$. By Corollary 1, the best responses $p^a\left(p_*,p^*;q\right)$ and $p^c\left(p_*,p^*;q\right)$ are increasing in $q$. This means that, when $p \geq \overline{p}$, $p \geq p^c\left(p_*,p^*;q\right)$ for at least $R_d$ agents, and, analogously, when $p \leq \underline{p}$, $p \leq p^a\left(p_*,p^*;q\right)$ for at least $R_d$ agents, so that in both cases there is a quorum for stopping deliberation whenever $p \leq \underline{p}$ or $p \geq \overline{p}$. It is also clear that whenever $p$ is in $(p_*,p^*)$ there is no such quorum. It follows that, in equilibrium, in must be the case that $p_* = p^a\left(p_*,p^*;q_{n-R_d+1}\right)$, and $p^* = p^c\left(p_*,p^*;q_{R_d}\right)$, as required. ■

**Proof of Lemma 4**

As $k$ decreases to zero, for any $q \in (0,1)$, $p^a(q;k)$ converges to zero and $p^c(q;k)$ converges to one: as information collection becomes extremely cheap, all types of jurors demand a high degree of confidence before either convicting or acquitting. Thus, for any $0 < p_1 < p_2 < 1$, there is a $k$ such that any stationary equilibrium thresholds $p_*$ and $p^*$ satisfy $p_* < p_1$ and $p^* > p_2$. ■

**Proof of Proposition 3**

**Part 1.** Consider an equilibrium $p_*(R_d, R_v)$, $p^*(R_d, R_v)$ with $R_v \leq R_d$. The pivotal jurors at the deliberation stage are $q_{n-R_d+1} \leq q_{n-R_v+1}$ and $q_{R_d} \geq q_{R_v}$. By the construction of optimal thresholds, $p_*(R_d, R_v) \leq q_{n-R_d+1}$ and $p^*(R_d, R_v) \geq q_{R_d}$. Consider any posterior $p$ such that, if reached, in equilibrium deliberation terminates: we have $p \leq p_*(R_d, R_v) \leq q_{n-R_d+1} \leq q_{n-R_v+1}$ and $p \geq p^*(R_d, R_v) \geq q_{R_d} \geq q_{R_v}$ so that whenever there is a quorum for stopping deliberation there is also a quorum for taking a decision.

**Part 2.** This is an immediate consequence of Lemma 4. ■

**Proof of Lemma 5**

Suppose there are two symmetric threshold equilibria: $(p_*, p^*)$ and $(\tilde{p}_*, \tilde{p}^*)$. Suppose $p_* < \tilde{p}_*$. From the monotonicity captured in Lemma 2, it must be the case that $p^* \leq \tilde{p}^*$. This would imply $\tilde{p}_* + \tilde{p}^* > p_* + p^* = 1$, in contradiction to the equilibrium $(\tilde{p}_*, \tilde{p}^*)$ being symmetric. ■

**Proof of Proposition 4**

**Part 1.** Assume by way of contradiction that, for $b > b'$, the symmetric equilibria corresponding to $b$ and $b'$ satisfy $p_*(b') < p_*(b)$ and $p^*(b') > p^*(b)$. By Lemma 1, the best response $p^a$ to $p^*(b)$ for the juror with preferences $\frac{1}{2} - b'$ must be such that $p^a < p_*(b') < p_*(b)$. But since $p_*(b)$ is a best response to $p^*(b)$ for the juror with preferences $\frac{1}{2} - b$, this violates monotonicity of (constrained) best responses in $q$ (Corollary 1).

**Part 2.** Note that Part 1 and monotonicity of thresholds with respect to $q$ together imply that (dropping the explicit dependence on $k$) $p_*(b) < p^a\left(\frac{1}{2} + b\right)$ and $p^*(b) > p^c\left(\frac{1}{2} - b\right)$. By Lemma 2, this implies that $p^c\left(\frac{1}{2} - b\right) > p_*(b)$. Inequalities are strict for interior equilibria. ■

**Proof of Proposition 5**

**Part 1.** Consider first voting rule $R_v = R_d$, and an associated symmetric equilibrium $p_*(R_d, R_v; k)$, $p^*(R_d, R_v; k)$. This equilibrium is unique by Lemma 5. Let $\underline{k}$ be such that $p_*(R_d, R_v; \underline{k}) = q_{n-\tilde{R}_v+1}$, $p^*(R_d, R_v; \underline{k}) = q_{\tilde{R}_v}$ a single such $\underline{k}$ is sufficient because of symmetry.[40] Thus, when the cost is $\underline{k}$, the pivotal voters under rule $R_v$ are just indifferent between voting to acquit and voting to convict, When $k$ is larger than $\underline{k}$, $p_*(R_d, R_v; k)$, $p^*(R_d, R_v; k)$ move further inwards (because, by Proposition 2, all best response thresholds involve acquiring less information) and therefore would induce a hung jury under $R_v$. Now assume that $k = \underline{k} + \varepsilon$. For $\varepsilon$ sufficiently small the unique symmetric equilibrium under rules $R_d, \tilde{R}_v$ involves extending deliberation thresholds just enough to avoid a hung jury, and for such $k$ we have, $p_*\left(R_d, \tilde{R}_v; k\right) = q_{n-\tilde{R}_v+1}$, $p^*\left(R_d, \tilde{R}_v; k\right) = q_{\tilde{R}_v}$. To see this, note that jurors $q_{n-R_d+1}, q_{R_d}$ are still pivotal at the deliberation stage. Any left threshold $p_* > q_{n-\tilde{R}_v+1}$ ($p^* < q_{\tilde{R}_v}$) would induce a hung jury, which is not optimal for $k$ sufficiently close to $\underline{k}$. This reasoning holds for any $k$ such that obtaining a verdict is better than a hung jury for the pivotal jurors at the deliberation stage. This holds as long as $k$ is not too high, i.e., lower than some $\overline{k}$. Clearly, any $p_* < q_{n-\tilde{R}_v+1}$ cannot be a best response to any $p^* \geq q_{\tilde{R}_v}$ (by Lemma 2): there is a loss in lowering the threshold even more, especially when the opposite threshold is high.

**Part 2.** Whenever $k$ is sufficiently small, there is a unique equilibrium under rules $\tilde{R}_d, \tilde{R}_v$ and under rules $R_d, \tilde{R}_v$. For such $k$, $p_*\left(\tilde{R}_d, \tilde{R}_v; k\right) \leq q_{n-\tilde{R}_d+1} = q_{n-\tilde{R}_v+1}$, and $p^*\left(\tilde{R}_d, \tilde{R}_v; k\right) \geq q_{\tilde{R}_d} = q_{\tilde{R}_v}$ since the optimal thresholds for deliberation for any type $q$ must surround $q$. From the analysis of the proof of part 1, for sufficiently small $k$, $p_*\left(R_d, \tilde{R}_v; k\right) = q_{n-\tilde{R}_v+1}$, $p^*\left(R_d, \tilde{R}_v; k\right) = q_{\tilde{R}_v}$ and the result follows. ■

**Proof of Proposition 6**

Consider the expression for juror payoffs from equation (3). In a quasi-symmetric jury, equilibrium thresholds are symmetric and so $p^*(R_d) = 1 - p_*(R_d)$, $\Pr(p^*(R_d) \text{ first} |\ p_*(R_d), p^*(R_d)) = \Pr(p_*(R_d) \text{ first} |\ p_*(R_d), p^*(R_d))$, and $1 - \mathbb{E}(p|p^*(R_d)) = \mathbb{E}(p|p_*(R_d))$. Slightly abusing notation by dropping

[40] Note that exact equality may not hold with discrete signals but the proof can be easily modified to take care of this by considering the closest feasible posteriors to the corresponding preference parameters.

the conditioning on the two thresholds, we can write (3) as:

$$\begin{aligned} U(q;R_d) &= -q\left(1-\mathbb{E}\left(p|p^*\left(R_d\right)\right)\right)\Pr(p^*\left(R_d\right)\ \text{first}) - (1-q)\,\mathbb{E}\left(p|p_*\left(R_d\right)\right)\Pr(p_*\left(R_d\right)\ \text{first}) - \\ &\quad -kT(p_*\left(R_d\right),p^*\left(R_d\right)) = \\ &= -\mathbb{E}\left(p|p_*\left(R_d\right)\right)\Pr(p_*\left(R_d\right)\ \text{first}) - kT(p_*\left(R_d\right),1-p_*\left(R_d\right)). \end{aligned}$$

This expression shows that preferences over deliberation rules are independent of $q$. To show that all jurors prefer the least inclusive deliberation rule, suppose first that two jurors with preference $q = 1/2$ existed (symmetry entails there being an even number of jurors having such a preference). For $R_d = n/2$, deliberation would precede just as it would with a dictator of preference parameter $q = 1/2$. Since all other jurors share their preferences with agents with such neutral preferences, $R_d = 1/2$ would be optimal for all jurors. Using the notation of Section 5 (with $b = 0$ for a neutral juror), we denote by $p_*(0)$ and $p^*(0)$ the resulting equilibrium thresholds. From the construction of the equilibrium and Proposition 2, if the lower threshold were constrained to be $p_{**} < p_*(0)$, an agent with preference parameter $q = 1/2$ would prefer to halt deliberation for any posterior $p > p^*(0)$,the analogue holding for the upper threshold. From Proposition 4, it follows directly that preferences would therefore be monotonically decreasing in $R_d$ for a juror of preferences $q = 1/2$ and therefore for all jurors. ■

**Proof of Proposition 7**

Recall that a hung jury leads to equal probabilities of conviction and acquittal. Given rules $R_d$ and $R_v$, payoffs to a juror with preference $q$ are then given by:

$$\begin{aligned} U\left(R_d,R_v;q\right) &= -q\left(1-\mathbb{E}\left(p_t|p_t \geq q_{R_v}\right)\right)\Pr\left(p_t \geq q_{R_v}\right) \\ &\quad -(1-q)\,\mathbb{E}\left(p_t|p_t < q_{n-R_v+1}\right)\left(\Pr\left(p_t < q_{n-R_v+1}\right)\right) \\ &\quad +\frac{1}{2}\mathbb{E}\left(-q\left(1-p_t\right)-(1-q)p_t|q_{n-R_v+1} < p_t < q_{R_v}\right)\left(\Pr\left(q_{n-R_v+1} < p_t < q_{R_v}\right)\right). \end{aligned}$$

With symmetric juries, $q_{n-R_v+1} = 1-q_{R_v}$, $\mathbb{E}\left(p_t|p_t < q_{n-R_v+1}\right) = 1-\mathbb{E}\left(p_t|p_t \geq q_{R_v}\right)$, and $\Pr\left(p_t < q_{n-R_v+1}\right) = \Pr\left(p_t \geq q_{R_v}\right)$. Furthermore, symmetry implies that

$$\frac{1}{2}\mathbb{E}\left(-q\left(1-p_t\right)-(1-q)p_t|q_{n-R_v+1} < p_t < q_{R_v}\right) = -\frac{1}{4}$$

is independent of $q$.[41] Therefore,

$$U(R_d, R_v) = \mathbb{E}(p_t | p_t < q_{R_v}) \Pr(p_t < q_{n-R_v+1}) + U(H)(\Pr(q_{n-R_v+1} < p_t < q_{R_v})).$$

This expression is independent of $q$. Therefore, jurors are unanimous in their deliberation votes, implying that the deliberation rule $R_d$ is irrelevant. However, the decision rule $R_v$ does matter: a larger $R_v$ raises the probability of a hung jury. This feeds back into the optimal sample size (for the unanimous jurors). Thus, a larger $R_v$ implies more information collection. ■

[41] Indeed, note that

$$\Pr(p_t < \frac{1}{2} | q_{n-R_v+1} < p_t < q_{R_v}) = \Pr(p_t > \frac{1}{2} | q_{n-R_v+1} < p_t < q) = \frac{1}{2}.$$

Furthermore, for any $p_t$, $q_{n-R_v+1} < p_t < q_{R_v}$, it must be that $q_{n-R_v+1} < 1 - p_t < q_{R_v}$ and

$$-q(1-p_t) - (1-q)p_t - q(1-(1-p_t)) - (1-q)(1-p_t) = 1.$$

## References


[1] Albrecht J., A. Anderson, and S. Vroman (2010), "Search by Committee," *Journal of Economic Theory*, Volume 145(4), pages 1386-1407.

[2] Austen-Smith, D. and T. Feddersen (2005), "Deliberation and Voting Rules," In *Social Choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks*, edited by David Austen-Smith and John Duggan, Springer-Verlag.

[3] Austen-Smith, D. and T. Feddersen (2006), "Deliberation, Preference Uncertainty, and Voting Rules," *American Political Science Review*, Volume 100(2), pages 209-217.

[4] Baldwin J. and M. McConville (1980), "Juries, Foremen, and Verdicts," *British Journal of Criminology*, Volume 20(1), pages 34-44.

[5] Bognar, K., M. Meyer-ter-Vehen, and L. Smith (2009), "A Conversational War of Attrition," mimeo.

[6] Blinder, A. S. and J. Morgan (2008), "Leadership in Groups: A Monetary Policy Experiment," *International Journal of Central Banking*, Volume 4(4), pages 117-150.

[7] Chan, J. and W. Suen, "Does Simple Majority Rule Produce Hasty Decisions?," mimeo.

[8] Compte O., and P. Jehiel (2010) "Bargaining and Majority Rules: A Collective Search Perspective," *Journal of Political Economy*, Volume 118(2), pages 189-221.

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

[10] De Groot, M. H. (1970), *Optimal Statistical Decisions*, McGraw-Hill.

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

[12] Ellsworth, P.C. (1989) "Are Twelve Heads Better than One?" *Law and Contemporary Problems*, Volume 52(4), pages 205-224.

[13] Eso, P. and Y. Fong (2008), "Wait and See: A Theory of Communication Over Time," mimeo.

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

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

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

[17] Goeree J. and L. Yariv (2011), "An Experimental Study of Collective Deliberation," *Econometrica*, Volume 89(3), pages 893-921.

[18] Harrison, J. M. (1985), *Brownian Motion and Stochastic Flow Systems*, John Wiley and Sons, New York.

[19] Hannaford P., V. Hans, N. Mott, and T. Musterman (2000), "The Timing of Opinion Formation by Jurors in Civil Cases: an Empirical Examination," *Tennessee Law Review*, Volume 67, pages 627-652.

[20] Hans V. (2001), "The Power of Twelve: The Impact of Jury Size and Unanimity on Civil Jury Decision Making," *Delaware Law Review*, Volume 4(1), pages 1-31.

[21] Hans V. (2007), "Deliberation and Dissent: *12 Angry Men* versus the Empirical Reality of Juries," *Chicago-Kent Law Review*, Volume 82, pages 579-589.

[22] Hans, V., P. Hannaford-Agor, N. Mott, and T. Musterman (2003), "The Hung Jury: *The American Jury*'s Insights and Contemporary Understanding," *Criminal Law Bulletin*, Volume 39, pages 33-50.

[23] Kalven H. and H. Zeisel (1966), *The American Jury*, Little Brown, Boston.

[24] Li, H. and W. Suen (2009), "Decision-making in Committees," *Canadian Journal of Economics*, Volume 42(2), pages 359-392.

[25] Meirowitz, A. (2006), "Designing Institutions to Aggregate Private Beliefs and Values," *Quarterly Journal of Political Science*, Volume 1(4), pages 373-392.

[26] Messner, M. and M. Polborn (2009), "The Option to Wait in Collective Decisions," mimeo.

[27] Moldovanu, B. and X. Shi (2010), "Search Committees," mimeo.

[28] Moscarini, G. and L. Smith (2001), "The Optimal Level of Experimentation," *Econometrica*, Volume 69(6), pages 1629-1644.

[29] Murphy, B. and R. G. Boatright (1999a), *Behind Closed Doors: A Resource Manual to Improve Jury Deliberation*.

[30] Murphy, B. and R. G. Boatright (1999b), "How Judges Can Help Deliberating Juries: Using the Guide for Jury Deliberations," *Court Review*, 38-45.

[31] Persico N. (2004), "Committee Design with Endogenous Information," *Review of Economic Studies*, Volume 71(1), pages 165-194.

[32] Sommers S. (2006), "On Racial Diversity and Group Decision Making: Identifying Multiple Effects of Racial Composition on Jury Deliberations," *Journal of Personality and Social Psychology*, Volume 90, pages 597-612.

[33] Strulovici B. (2010), "Learning While Voting: Determinants of Collective Experimentation," *Econometrica*, Volume 78(3), pages 933-971.

[34] Vidmar and Hans (2007) *American Juries: the Verdict* Prometheus Books.

[35] Wald, A. (1947a), "Foundations of a General Theory of Sequential Decision Functions," *Econometrica*, Volume 15(4), pages 279-313.

[36] Wald, A. (1947b), *Sequential Analysis*, John Wiley and Sons, New York.

[37] Wilson, A. J. (2011), "Costly Communication and Group Decision Making," mimeo.