

A Bayesian Model of Voting in Juries

John Duggan, and Cesar Martinelli

Working Paper No. 18
November 1998

W. ALLEN
WALLIS
Institute of
POLITICAL
ECONOMY

UNIVERSITY OF
ROCHESTER

A Bayesian Model of Voting in Juries

John Duggan

Department of Political Science
and Department of Economics
University of Rochester
Rochester, NY 14627

César Martinelli

Centro de Investigación Económica
Instituto Tecnológico Autónomo de México
Santa Teresa 930, México, D.F. 10700

November 9, 1998

Abstract

We take a game-theoretic approach to the analysis of juries by modelling voting as a game of incomplete information. Rather than the usual assumption of two possible signals (one indicating guilt, the other innocence), we allow jurors to perceive a full spectrum of signals. We offer three main results. First, given any voting rule requiring a fixed fraction of votes to convict, we characterize the unique symmetric equilibrium of the game. Second, we obtain a condition under which unanimity rule exhibits a bias toward convicting the innocent. Third, we prove a “jury theorem” for the continuous signal model: as the size of the jury increases, the probability of making a mistaken judgment goes to zero for every voting rule, *except* unanimity rule; for unanimity rule, the probability of a mistake is bounded strictly above zero.

1 Introduction

Consider a group of decision-makers who must choose one of two alternatives. Voters agree on the overall objective, but, on the basis of differential information, they may disagree on which alternative best achieves that goal. Some examples are:

- A jury deciding whether to convict or acquit a defendant. Jurors agree about the desirability of acquitting an innocent and convicting a guilty defendant, but they have different opinions about whether the defendant is innocent or guilty.
- The board of directors of a company deciding whether to approve a new investment project. All members of the board agree about the objective of maximizing profits, but they disagree in their estimates of the profitability of the project.
- An academic department deciding whether to hire a job candidate. The common objective is to hire a productive researcher, but there is disagreement about the academic potential of the candidate.

Although our results apply equally well to all situations exhibiting this structure, we follow an old literature on information aggregation in elections by focusing on the jury example. The literature traces back to Condorcet's (1785) jury theorem, which asserts that, under majority voting, large electorates should reach correct decisions with very high probability (cf. Miller (1986), Grofman and Feld (1988), Young (1988), Ladha (1992)). It is traditionally assumed that each voter simply behaves "naively," i.e., as if the voter were deciding the outcome alone, but Austen-Smith and Banks (1996)

observe that, given naive behavior on the part of jurors, some may have an incentive to vote “strategically.” In other words, naive behavior does not generally constitute an equilibrium. We follow subsequent papers in analyzing voting in juries as a Bayesian game in which the jurors’ opinions of guilt or innocence, i.e., their “signals,” are private information.

Several versions of the jury theorem under strategic voting have been offered. Feddersen and Pesendorfer (1997) assume discrete distributions of signals and a continuum of states. Myerson (1997) introduces uncertainty about the size of the electorate and considers a countable set of players’ types (signals), where the number of voters receiving any given signal is drawn from a Poisson distribution, the mean of which depends on guilt or innocence. McClellan (1998) proves that there is at least one equilibrium that maximizes the *ex ante* payoffs of jurors over the class of symmetric strategy profiles. Feddersen and Pesendorfer (1998) analyze a simple model in which there are two possible signals, one indicating guilt and the other innocence. Given any voting rule requiring a fixed fraction of votes to convict, they are able to explicitly solve for the unique symmetric, responsive Bayesian equilibrium of the voting game. They show that a jury theorem holds for all voting rules other than unanimity: as the size of the jury increases, the probability of making a mistaken judgment goes to zero for all voting rules, except unanimity; in that case, the probability of a mistake is bounded strictly above zero. Feddersen and Pesendorfer (1998) also give an example comparing different voting rules for a fixed jury size: there, the probability of convicting an innocent defendant under unanimity rule is greater than the probability under majority or any other supermajority rule. McKelvey and Palfrey (1998) offer experimental results on the binary signal model roughly consistent with

the equilibrium predictions.

We depart from the previous literature on juries by assuming that the signals representing the jurors' opinions of guilt or innocence are drawn from continuous, rather than discrete (usually binary), distributions. This is meant to capture the fact that a juror's opinion of the evidence against the defendant, the case made by the prosecutor, etc., may reflect a very rich spectrum of possibilities — possibilities that cannot be summarized by a dichotomous signal merely indicating guilt or innocence. We impose few restrictions on the distributions of signals, and we are actually able to obtain a continuous analogue of the “binary signal” model as a special case. Unlike Feddersen and Pesendorfer (1997), we confine our attention to the case in which the objectives of the jurors are perfectly aligned. Within this framework, we analyze voting in juries as a game of incomplete information.

We offer three main results. First, we establish the existence of a symmetric, responsive equilibrium characterized by a cutoff signal: jurors who get signals indicating a higher likelihood of guilt vote for convicting the defendant while those who get signals indicating a lower likelihood vote for acquittal. The equilibrium is unique within that class. Moreover, under a strict monotone likelihood ratio condition, all equilibria are cutoff equilibria; as a consequence, our uniqueness result extends to the class of all symmetric, responsive equilibria, even allowing for mixed strategies. With McClellan's (1998) result, this implies that the equilibrium is *optimal*: it maximizes the jurors' *ex ante* payoffs over the set of symmetric strategy profiles. An undesirable artifact of the binary signal model of Feddersen and Pesendorfer (1998), and of the continuous version we consider, is that, typically, jurors who vote to acquit must use mixed strategies; therefore, they are indifferent

between voting to acquit and voting to convict. When continuous distributions are allowed for, that is no longer typically the case. Indeed, under the strict monotone likelihood ratio condition, all equilibria are essentially *strict*: only a juror who receives the cutoff signal (a zero probability event) is indifferent to which vote he casts.

Second, we give a sufficient condition for unanimity rule to exhibit a bias toward convicting the innocent, independently of the size of the jury. This condition is met in our continuous version of the binary signal model, with the following implication: unanimity rule leads to a higher probability of convicting an innocent defendant than majority or any other supermajority rule. Thus, we find that Feddersen and Pesendorfer's (1998) fixed jury size example generalizes within the binary signal model. It generalizes beyond the binary signal model as well: we give an example with exponentially distributed signals in which the sufficient condition also holds.

Third, we obtain a “jury theorem” for the general continuous signal model: for all voting rules other than unanimity, the probability of making a mistaken judgment goes to zero for any voting rule as the size of the jury increases; for unanimity rule, the probability of making a mistake is bounded strictly above zero. Thus, the asymptotic inferiority of unanimity rule and the asymptotic efficiency of all other voting rules are fully general — not merely artifacts of the binary signal model.

We leave several important issues for further work. One is the extension of our results to situations in which the jurors' preferences are not perfectly aligned. Other issues are allowing for correlation among the jurors' signals and for the possibility of more than two “states,” the latter not relevant to the jury model perhaps, but certainly to others. Finally, there is the possibility

of limited communication among jurors. As stressed by Coughlan (1997), a single nonbinding “straw vote” is enough to allow jurors to share all their information in the binary signal model, thus eliminating the strategic aspects of voting in a common preference environment. In general, a finite number of “straw votes” is enough to allow jurors to share all their information if the distribution of signals is discrete. But, when the opinions of jurors can reflect subtle nuances of trials, a continuous distribution of signals seems better suited to model the difficulties associated with limited communication.

2 Preliminaries

We consider $n \geq 2$ jurors who must decide whether to convict or acquit a defendant. The defendant is either innocent, I , or guilty, G , with probabilities $P(I)$ and $P(G)$. Each juror i receives a real-valued signal s_i distributed according to $F(\cdot|G)$ or $F(\cdot|I)$, depending on whether the defendant is guilty or innocent. After receiving their signals, which are private information, the jurors simultaneously vote to convict or acquit. Once the votes are tallied, the defendant’s fate is determined by an anonymous, monotonic decision rule, i.e., there is some integer, k , such that the defendant is convicted, C , if k or more jurors vote to convict and acquitted, A , otherwise.

We assume the jurors have a common preference to convict the guilty and acquit the innocent. We assume that these outcomes are equally desirable and normalize the jurors’ payoffs in those cases to $u(C|G) = u(A|I) = 0$. In the cases of convicting the innocent or acquitting the guilty, the jurors receive negative payoffs $u(C|I)$ and $u(A|G)$. In effect, the *ex ante* cost of conviction is $u(C|I)P(I)$, and the cost of acquittal is $u(A|G)P(G)$. We use

$$\rho = \frac{u(A|G) P(G)}{u(C|I) P(I)}$$

to denote the relative cost of acquittal.

A strategy for juror i is a measurable mapping $\sigma_i: \mathfrak{R} \rightarrow [0, 1]$, where $\sigma_i(s_i)$ is the probability that the juror votes to convict. The probability that i votes to convict, conditional on the defendant being innocent, is

$$\int \sigma_i(s) \mu_I(ds),$$

where μ_I is the probability measure induced by $F(\cdot|I)$. The probability that i votes to convict conditional on guilt is identical, except that μ_G , the probability measure induced by $F(\cdot|G)$, is used. Probabilities of acquittal are written similarly, but integrating $1 - \sigma_i$ rather than σ_i .

A profile of strategies is denoted $\sigma = (\sigma_1, \dots, \sigma_n)$. Given σ , the probability that the defendant is convicted conditional on being innocent, denoted $P_\sigma(C|I)$, is

$$\sum_{\substack{M \subseteq N \\ |M| \geq k}} \prod_{j \in M} \left(\int \sigma_j(s) \mu_I(ds) \right) \prod_{j \notin M} \left(\int [1 - \sigma_j(s)] \mu_I(ds) \right).$$

The probability that the defendant is guilty conditional on being acquitted, denoted $P_\sigma(A|G)$, is

$$\sum_{\substack{M \subseteq N \\ |M| < k}} \prod_{j \in M} \left(\int \sigma_j(s) \mu_G(ds) \right) \prod_{j \notin M} \left(\int [1 - \sigma_j(s)] \mu_G(ds) \right).$$

The *ex ante* payoff of a juror is

$$u(C|I)P_\sigma(C|I)P(I) + u(A|G)P_\sigma(A|G)P(G).$$

Let σ_{-i} represent the strategies of jurors other than i . The probabilities that i is pivotal (that is, that his vote decides the fate of the defendant)

conditional on innocence and guilt, $P_{\sigma_{-i}}(piv|I)$ and $P_{\sigma_{-i}}(piv|G)$, are defined as

$$\sum_{\substack{M \subseteq N \\ |M|=k-1 \\ i \notin M}} \prod_{j \in M} \left(\int \sigma_j(s) \mu_I(ds) \right) \cdot \prod_{\substack{j \notin M \\ j \neq i}} \left(\int [1 - \sigma_j(s)] \mu_I(ds) \right)$$

and

$$\sum_{\substack{M \subseteq N \\ |M|=k-1 \\ i \notin M}} \prod_{j \in M} \left(\int \sigma_j(s) \mu_G(ds) \right) \cdot \prod_{\substack{j \notin M \\ j \neq i}} \left(\int [1 - \sigma_j(s)] \mu_G(ds) \right),$$

respectively.

Using the above definitions, we obtain an expression for a juror's payoff in terms of his probability of being pivotal. It translates to our framework the insight from the literature on strategic voting that a voter should condition his vote on being pivotal, as this is the only event where his vote might affect his payoff.

Proposition 1 *Given σ_{-i} , the ex ante payoff to juror i of σ_i is an affine transformation of*

$$\begin{aligned} & \left(\int \sigma_i(s) \mu_I(ds) \right) u(C|I) P_{\sigma_{-i}}(piv|I) P(I) \\ & - \left(\int \sigma_i(s) \mu_G(ds) \right) u(A|G) P_{\sigma_{-i}}(piv|G) P(G). \end{aligned}$$

The proof of this and other auxiliary results is found in the appendix.

An *equilibrium* is a profile σ such that, for every juror i and every σ'_i ,

$$\begin{aligned} & u(C|I) P_{\sigma}(C|I) P(I) + u(A|G) P_{\sigma}(A|G) P(G) \\ & \geq u(C|I) P_{\sigma'_i, \sigma_{-i}}(C|I) P(I) + u(A|G) P_{\sigma'_i, \sigma_{-i}}(A|G) P(G). \end{aligned}$$

A *responsive equilibrium* is an equilibrium σ such that each σ_i is *responsive*:

$$0 < \int \sigma_i(s) \mu_G(ds) < 1 \text{ and } 0 < \int \sigma_i(s) \mu_I(ds) < 1.$$

There are always unresponsive equilibria: if $k < n$, it is an equilibrium for the jurors to convict regardless of their signals; if $k > 1$, it is an equilibrium to acquit regardless of signal. A *cutoff equilibrium* is a pure strategy equilibrium σ such that each σ_i is a *cutoff strategy*: there is some $\bar{s}_i \in [-\infty, \infty]$ such that

$$\sigma_i(s) = \begin{cases} 1 & \text{if } s > \bar{s}_i \\ 0 & \text{if } s < \bar{s}_i \end{cases}$$

for all $s \in \mathfrak{R}$. The cutoff strategy associated with a given cutoff is unique up to the behavior of the juror upon receiving the cutoff signal. Given assumption (A1), below, this is a zero-probability event, and we will not distinguish between cutoff strategies that differ only at the cutoff.

In what follows we maintain several assumptions on $F(\cdot|I)$ and $F(\cdot|G)$. We will see that these assumptions enable us to restrict our attention to cutoff equilibria.

(A1) *The distribution functions are absolutely continuous with respect to Lebesgue measure and have piecewise continuous densities $f(\cdot|I)$ and $f(\cdot|G)$.*

This assumption implies that the distribution functions are differentiable at all but a finite number of points. We will use S^d to denote the subset of signals in S on which $F(\cdot|I)$ and $F(\cdot|G)$ are both differentiable.

(A2) *The densities have common support, $S = (\underline{S}, \bar{S})$, where $\underline{S}, \bar{S} \in [-\infty, \infty]$: $f(s|I) > 0$ and $f(s|G) > 0$ for all $s \in S$.*

The latter implies that μ_I and μ_G have the same sets of measure zero. The terms “ μ_I -a.e.” and “ μ_G -a.e.” are thus synonymous, so we can use “a.e.” without ambiguity.

(A3) *The likelihood ratio, $f(s|I)/f(s|G)$, is weakly decreasing on S .*

This assumption is standard and amounts to assuming that higher signals are stronger (or at least not weaker) indications of guilt. Sometimes we will want a stronger condition to hold locally: we will say that the likelihood ratio is *strictly decreasing at $x \in S$* if for all $x', x'' \in S$ with $x' < x < x''$,

$$\frac{f(x'|I)}{f(x'|G)} > \frac{f(x|I)}{f(x|G)} > \frac{f(x''|I)}{f(x''|G)}.$$

As we will see, a juror who behaves “naively” (i.e., as if his vote alone determines the outcome) after receiving signal s would prefer to convict if $f(s|I)/f(s|G) > \rho$ and would prefer to acquit if $f(s|I)/f(s|G) < \rho$.

$$(A4) \quad \lim_{s \downarrow \underline{s}} \frac{f(s|I)}{f(s|G)} > \rho > \lim_{s \uparrow \bar{s}} \frac{f(s|I)}{f(s|G)}.$$

Thus, (A4) implies that there must be a signal low enough to induce a naive juror to acquit, and a signal high enough to induce him to convict.

Lemma 0, stated in the appendix, establishes some implications of (A1)–(A4). Among those that are well-known, $F(\cdot|G)$ exhibits (strict) first order stochastic dominance over $F(\cdot|I)$, and the ratios

$$\frac{1 - F(s|I)}{1 - F(s|G)} \quad \text{and} \quad \frac{F(s|I)}{F(s|G)}$$

are weakly decreasing.

3 Existence and Uniqueness

Consider any profile σ of responsive strategies and any juror i . Since the strategies are responsive, $P_{\sigma_{-i}}(piv|G)$ and $P_{\sigma_{-i}}(piv|I)$ are positive. Hence, under our assumptions the expression

$$J(\sigma_{-i}, s) = \frac{P_{\sigma_{-i}}(piv|I)}{P_{\sigma_{-i}}(piv|G)} \frac{f(s|I)}{f(s|G)} - \rho$$

is well-defined on S . Moreover, for fixed σ_{-i} , it is weakly decreasing in its second argument. Note that $J(\sigma_{-i}, s) > 0$ if and only if

$$u(C|I)P_{\sigma_{-i}}(piv|I)P(I)f(s|I) < u(A|G)P_{\sigma_{-i}}(piv|G)P(G)f(s|G).$$

That is, $J(\sigma_{-i}, s) > 0$ if and only if a juror's expected gain from voting to convict, conditional on receiving signal s and on the strategies of others, is less than the expected gain from voting to acquit. Hence, as shown in the following lemma, jurors will be inclined to acquit when J is positive and to convict when it is negative. In contrast, a naive juror would behave as if the terms $P_{\sigma_{-i}}(piv|I)$ and $P_{\sigma_{-i}}(piv|G)$ were equal to one, and he would vote to acquit if

$$u(C|I)P(I)f(s|I) < u(A|G)P(G)f(s|G)$$

and to convict if the inequality were reversed.

Lemma 1 *Given responsive strategies σ_{-i} for jurors other than i , a strategy σ_i is a best response for i if and only if it satisfies the following a.e.:*

$$\sigma_i(s) = \begin{cases} 1 & \text{if } J(\sigma_{-i}, s) < 0 \\ 0 & \text{if } J(\sigma_{-i}, s) > 0. \end{cases}$$

If the likelihood ratio is strictly decreasing at $\inf\{s \in S \mid J(\sigma_{-i}, s) \square 0\}$, σ_i is a best response for i if and only if it is equivalent a.e. to the following cutoff strategy $\tilde{\sigma}_i$:

$$\tilde{\sigma}_i(s) = \begin{cases} 1 & \text{if } J(\sigma_{-i}, s) \square 0 \\ 0 & \text{else.} \end{cases}$$

Since J is weakly decreasing in its second argument, an implication of the first part of the preceding lemma is that jurors always have best response cutoff strategies. Hence, if a profile of strategies is an equilibrium when

jurors are restricted to cutoff strategies, it will be an equilibrium of the unrestricted game. From the second part of the lemma, if the likelihood ratio is strictly decreasing, then *all* best responses for a juror are equivalent to cutoff strategies, regardless of the strategies of others, and *all* equilibria are equivalent to cutoff equilibria.

When all jurors other than i use the same cutoff strategy, given by cutoff \bar{s} , we will write $J(\bar{s}, s)$ for $J(\sigma_{-i}, s)$. That is,

$$J(\bar{s}, s) = \left(\frac{1 - F(\bar{s}|I)}{1 - F(\bar{s}|G)} \right)^{k-1} \left(\frac{F(\bar{s}|I)}{F(\bar{s}|G)} \right)^{n-k} \frac{f(s|I)}{f(s|G)} - \rho.$$

Henceforth, we will only consider cutoff strategies, and we will view J as a mapping defined on $S \times S$.

We have already noted that J is weakly decreasing in its second argument. The following lemma further characterizes J for the case in which the jurors use the same cutoff strategy.

Lemma 2 *J is continuous and weakly decreasing in its first argument. In addition,*

$$\lim_{s \downarrow \underline{s}} J(s, s) > 0 \quad \text{and} \quad \lim_{s \uparrow \bar{s}} J(s, s) < 0.$$

Finally, $J(s, s) = 0$ has at most one solution.

The following theorem establishes existence of a cutoff equilibrium under our assumptions (A1)–(A4). The equilibrium is unique within the class of symmetric, responsive cutoff equilibria. If the likelihood ratio is strictly decreasing, there are no other symmetric, responsive equilibria.

Theorem 1 *There exists a symmetric, responsive cutoff equilibrium with cutoff given by $s^* = \inf\{s \in S \mid J(s, s) \square 0\}$. It is unique within the class*

of symmetric, responsive cutoff equilibria. If the likelihood ratio is strictly decreasing at s^* , then this equilibrium is unique a.e. within the class of all symmetric, responsive equilibria.

Proof: By Lemma 1, s^* defines a symmetric, responsive cutoff equilibrium if and only if $s^* \in S$, $J(s^*, s) \geq 0$ for all $s < s^*$, and $J(s^*, s) \leq 0$ for all $s > s^*$.

Take any $s < s^*$ and suppose $J(s^*, s) < 0$. Since J is continuous in its first argument, by Lemma 2, there is some $\epsilon > 0$ such that $s^* - \epsilon > s$ and $J(s^* - \epsilon, s) < 0$. Since J is weakly decreasing in its second argument, $J(s^* - \epsilon, s^* - \epsilon) < 0$, contradicting our definition of s^* . Therefore, $s < s^*$ implies $J(s^*, s) \geq 0$. Now take any $s > s^*$ and suppose $J(s^*, s) > 0$. Since J is continuous in its first argument, there is some $\epsilon > 0$ such that $s^* + \epsilon < s$ and $J(s^* + \epsilon, s) > 0$. Since J is weakly decreasing in its second argument, $J(s^* + \epsilon, s^* + \epsilon) > 0$. But then, since J is also decreasing in its first argument, by Lemma 2, $J(\hat{s}, \hat{s}) > 0$ for all $\hat{s} < s^* + \epsilon$, contradicting the definition of s^* . Therefore, $s > s^*$ implies $J(s^*, s) \leq 0$, giving us the first part of the theorem.

To prove the second part of the theorem, consider any s' such that $J(s', s) \geq 0$ for all $s < s'$ and $J(s', s) \leq 0$ for all $s > s'$. If $s' < s^*$, take $\epsilon > 0$ such that $s^* - \epsilon > s'$. Then $J(s', s^* - \epsilon) \leq 0$. Since J is decreasing in its first argument, $J(s^* - \epsilon, s^* - \epsilon) \leq 0$, contradicting our definition of s^* . If $s' > s^*$, take $\epsilon > 0$ such that $s^* + \epsilon < s'$. Then $J(s', s^* + \epsilon) \geq 0$. Since J is decreasing in its first argument, $J(s^* + \epsilon, s^* + \epsilon) \geq 0$. Since J is decreasing in its second argument, by Lemma 2, as well as its first, $J(s, s) \geq 0$ for all $s < s^* + \epsilon$. Then, since $J(s, s) = 0$ has at most one solution by Lemma 2, $J(s, s) > 0$ for all $s < s^* + \epsilon$, contradicting our definition of s^* .

To prove the third part of the theorem, assume that the likelihood ratio

is strictly decreasing at s^* . In any symmetric, responsive equilibrium, the jurors use the same best response strategy. By Lemma 1, this strategy is equivalent a.e. to a cutoff strategy, and we just proved uniqueness within the class of symmetric, responsive cutoff equilibria. ■

4 Examples

We give two examples to illustrate the general model developed above. The first is a continuous analogue of the binary signal model of Feddersen and Pesendorfer (1998), in which our pure strategy cutoff equilibria can be interpreted as purifications of the mixed strategy equilibria of their model. More generally, any discrete signal model could be matched with a continuous analogue in a similar way. We will see that a condition on the ratio of hazard rates used in the next section holds for majority and supermajority rules. In the second example, the signals of the jurors are exponentially distributed. This functional form is particularly tractable: the ratio of hazard rates is constant in this model, easily delivering the condition of the next section; and the probabilities of convicting an innocent and of acquitting a guilty defendant are positive and independent of the number of jurors under unanimity rule, anticipating our result on the asymptotic inferiority of unanimity rule.

To define the *binary signal model* in our framework, let $S = (0, 2)$,

$$f(s|I) = \begin{cases} p & \text{if } 0 < s \leq 1 \\ 1 - p & \text{if } 1 < s < 2, \end{cases}$$

and

$$f(s|G) = \begin{cases} 1 - p & \text{if } 0 < s \leq 1 \\ p & \text{if } 1 < s < 2. \end{cases}$$

In order to satisfy (A3) and (A4), we impose $1/2 < p < 1$ and

$$\frac{p}{1-p} > \rho > \frac{1-p}{p}.$$

It follows that

$$J(s, s) = \begin{cases} \left(\frac{1-sp}{1-s+sp} \right)^{k-1} \left(\frac{p}{1-p} \right)^{n-k+1} - \rho & \text{if } 0 < s \leq 1 \\ \left(\frac{1-p}{p} \right)^k \left(\frac{s-1-ps+2p}{1+ps-2p} \right)^{n-k} - \rho & \text{if } 1 < s < 2. \end{cases}$$

Note that J is weakly decreasing (strictly so if $1 < k < n$), and that it is discontinuous at $s = 1$. See Figure 1.

[Figure 1 about here.]

Recall that s^* is defined in the statement of Theorem 1 as $\inf\{s \in S \mid J(s, s) \leq 0\}$. Since

$$\begin{aligned} J(1, 1) &= \left(\frac{p}{1-p} \right)^{n-2k+2} - \rho \\ \lim_{s \downarrow 1} J(s, s) &= \left(\frac{p}{1-p} \right)^{n-2k} - \rho, \end{aligned}$$

we see that $J(1, 1) > 0$ if and only if (1) $k < \frac{n}{2} + 1$ or (2) $k = \frac{n}{2} + 1$ and $\rho < 1$; and $\lim_{s \downarrow 1} J(s, s) \leq 0$ if and only if (3) $k > \frac{n}{2}$ or (4) $k = \frac{n}{2}$ and $\rho \geq 1$. We conclude that $s^* < 1$ if neither (1) nor (2) hold; $s^* > 1$ if neither (3) nor (4) hold; and $s^* = 1$ in the remaining cases. Once the location of s^* is determined, the unique equilibrium is either given by $s^* = 1$ or by the solution to $J(s^*, s^*) = 0$.

In particular, we note that if a supermajority is required to convict, $s^* < 1$ follows; if a majority is required and $\rho \geq 1$, again $s^* < 1$. That is, in

equilibrium, jurors who receive some signals below one (indicating innocence) will vote to acquit, while those who receive other signals below one will vote to convict. Note also that $s^* = 1$ holds if and only if either $k = \frac{n}{2} + 1$ and $\rho < 1$ or else $k = \frac{n}{2}$ and $\rho \geq 1$. Thus, it is an equilibrium for jurors to “vote with their signals” only under majority rule (or close to it) and then only for restricted ρ 's.

In the model of Feddersen and Pesendorfer (1998), jurors get one of only two possible signals: a signal that innocence is likely, which occurs with probability p if the defendant is innocent and with probability $1 - p$ if the defendant is guilty, and a signal that guilt is likely, which occurs with the same probabilities reversed. Our example replaces the innocence signal with a continuum of signals from 0 to 1, and the guilt signal with a continuum of signals from 1 to 2. Whereas jurors who receive the innocence signal in their model vote to acquit with some probability, say a , and to convict with some probability $1 - a$ (assuming supermajority rule), we partition $(0, 1)$ into two intervals $(0, a)$ and $(a, 1)$; jurors who receive signals in $(0, a)$ vote to acquit, and those who receive signals in $(a, 1)$ vote to convict. Thus, the cutoff equilibrium in our version of the binary signal model is a purification of the mixed strategy equilibrium in their model.

Theorem 1 guarantees existence of a symmetric, responsive cutoff equilibrium and uniqueness within that class; but because the likelihood ratio is not strictly decreasing in our version of the binary signal model, the theorem does not guarantee uniqueness within the larger class of all symmetric, responsive equilibria. Indeed, because jurors who receive signals between 0 and 1 are indifferent between voting to convict and voting to acquit, there is a continuum of symmetric, responsive non-cutoff equilibria: we could spec-

ify any subset of $(0, 1)$ with Lebesgue measure a and have jurors receiving signals therein vote to acquit, jurors receiving other signals vote to convict. Note, however, that the likelihood ratio *is* strictly decreasing at one; thus, if $s^* = 1$, then there are no symmetric, responsive non-cutoff equilibria.

Before leaving the binary signal model, we calculate the hazard rates

$$\frac{f(s|I)}{1 - F(s|I)} = \begin{cases} \frac{p}{1-sp} & \text{if } 0 < s \leq 1 \\ \frac{1}{2-s} & \text{if } 1 < s < 2 \end{cases}$$

and

$$\frac{f(s|G)}{1 - F(s|G)} = \begin{cases} \frac{1-p}{1-s+sp} & \text{if } 0 < s \leq 1 \\ \frac{1}{2-s} & \text{if } 1 < s < 2. \end{cases}$$

Thus, the ratio of the hazard rate when innocent to the hazard rate when guilty increases from $p/(1-p)$ at $s = 0$ to $(p/(1-p))^2$ at $s = 1$, and then drops to one for $1 < s < 2$.

To define the *exponential model*, let $S = (0, \infty)$, $f(s|I) = \lambda e^{-\lambda s}$, and $f(s|G) = \gamma e^{-\gamma s}$. In order to satisfy (A3) and (A4), we need $\lambda > \gamma$ and $\lambda/\gamma > \rho$. It follows that

$$J(s, s) = \left(\frac{e^{-\lambda s}}{e^{-\gamma s}} \right)^{k-1} \left(\frac{1 - e^{-\lambda s}}{1 - e^{-\gamma s}} \right)^{n-k} \frac{\lambda e^{-\lambda s}}{\gamma e^{-\gamma s}} - \rho,$$

which is strictly decreasing and continuous on S . See Figure 2.

[Figure 2 about here.]

The unique equilibrium is found by solving $J(s^*, s^*) = 0$; for the special case of unanimity rule, we readily obtain

$$s^* = \frac{1}{(\lambda - \gamma)n} \ln \left(\frac{\lambda}{\gamma\rho} \right).$$

The hazard rates when innocent and guilty are λ and γ , respectively, so the ratio of hazard rates is constant. (Other examples with constant ratios of

hazard rates can be easily obtained from certain parameterizations of the Pareto and Weibull distributions.) Under unanimity rule, the probability of convicting the defendant, conditional on innocence, is

$$(1 - F(s^*|I))^n = \left(\frac{\rho\gamma}{\lambda}\right)^{\frac{\lambda}{\lambda-\gamma}},$$

and the probability of acquitting the defendant, conditional on guilt, is

$$1 - (1 - F(s^*|G))^n = 1 - \left(\frac{\rho\gamma}{\lambda}\right)^{\frac{\gamma}{\lambda-\gamma}}.$$

Note that both probabilities are strictly positive and independent of the size of the jury. As a consequence, the probability of making a mistake does not go to zero as the size of the jury increases, a result we prove later in more generality.

5 The Bias of Unanimity Rule

In this section, we investigate jury decision-making under different possible voting rules. The main result of the section gives a sufficient condition for unanimity rule to exhibit a bias in favor of convicting innocent defendants independently of the size of the jury. We write s_k for the cutoff corresponding to the unique symmetric, responsive cutoff equilibrium when the number of votes needed to convict is k . Thus, the cutoff corresponding to unanimity rule is given by s_n , and the cutoff corresponding to simple majority rule (with an odd number of voters) is given by $s_{(n+1)/2}$. We write J_k to make explicit the dependence of J on the decision rule. We write $P_k(C|I)$ for the probability of conviction conditional on innocence, and $P_k(A|G)$ for the probability of acquittal conditional on guilt, under the unique symmetric, responsive cutoff equilibrium, when the number of votes needed to convict is k . Define

$$H(s|I) = \frac{f(s|I)}{1 - F(s|I)} \quad \text{and} \quad H(s|G) = \frac{f(s|G)}{1 - F(s|G)},$$

the hazard rates when the defendant is innocent and when he is guilty.

Theorem 2 *If*

$$\lim_{s \uparrow s_n} \frac{H(s|I)}{H(s|G)} \quad \square \quad \lim_{s \downarrow s_k} \frac{H(s|I)}{H(s|G)}$$

then

$$\frac{P_n(C|I)}{1 - P_n(A|G)} > \frac{P_k(C|I)}{1 - P_k(A|G)},$$

for all $k = 1, 2, \dots, n - 1$.

Proof: Note that

$$\frac{P_n(C|I)}{1 - P_n(A|G)} = \frac{(1 - F(s_n|I))^n}{(1 - F(s_n|G))^n}$$

and, for $k = 1, 2, \dots, n - 1$,

$$\begin{aligned} \frac{P_k(C|I)}{1 - P_k(A|G)} &= \frac{\sum_{m=k}^n \binom{n}{m} (1 - F(s_k|I))^m (F(s_k|I))^{n-m}}{\sum_{m=k}^n \binom{n}{m} (1 - F(s_k|G))^m (F(s_k|G))^{n-m}} \\ &= \frac{\binom{n}{k} (1 - F(s_k|I))^k (F(s_k|I))^{n-k} + \sum_{m=k+1}^n \binom{n}{m} \left(\frac{1 - F(s_k|I)}{F(s_k|I)} \right)^{m-k}}{\binom{n}{k} (1 - F(s_k|G))^k (F(s_k|G))^{n-k} + \sum_{m=k+1}^n \binom{n}{m} \left(\frac{1 - F(s_k|G)}{F(s_k|G)} \right)^{m-k}}. \end{aligned}$$

By (4) of Lemma 0, on strict first order stochastic dominance,

$$\sum_{m=k+1}^n \binom{n}{m} \left(\frac{1 - F(s_k|I)}{F(s_k|I)} \right)^{m-k} < \sum_{m=k+1}^n \binom{n}{m} \left(\frac{1 - F(s_k|G)}{F(s_k|G)} \right)^{m-k}.$$

Hence, for $k = 1, 2, \dots, n - 1$,

$$\frac{P_k(C|I)}{1 - P_k(A|G)} < \frac{(1 - F(s_k|I))^k (F(s_k|I))^{n-k}}{(1 - F(s_k|G))^k (F(s_k|G))^{n-k}},$$

so it remains only to be shown that

$$\left(\frac{1 - F(s_n|I)}{1 - F(s_n|G)}\right)^n \geq \left(\frac{1 - F(s_k|I)}{1 - F(s_k|G)}\right)^k \left(\frac{F(s_k|I)}{F(s_k|G)}\right)^{n-k}$$

for $k = 1, 2, \dots, n - 1$.

Take any $\epsilon > 0$ such that $s_n - \epsilon \in S$ and $s_k + \epsilon \in S$. By definition of s_n and s_k and because J is decreasing in both arguments, $J_n(s_n - \epsilon, s_n - \epsilon) > 0 \geq J_k(s_k + \epsilon, s_k + \epsilon)$. Note that we can write

$$J_k(s, s) = \left(\frac{1 - F(s|I)}{1 - F(s|G)}\right)^k \left(\frac{F(s|I)}{F(s|G)}\right)^{n-k} \frac{H(s|I)}{H(s|G)} - \rho$$

for all $s \in S$, giving us

$$\begin{aligned} \left(\frac{1 - F(s_n - \epsilon|I)}{1 - F(s_n - \epsilon|G)}\right)^n \frac{H(s_n - \epsilon|I)}{H(s_n - \epsilon|G)} \\ > \left(\frac{1 - F(s_k + \epsilon|I)}{1 - F(s_k + \epsilon|G)}\right)^k \left(\frac{F(s_k + \epsilon|I)}{F(s_k + \epsilon|G)}\right)^{n-k} \frac{H(s_k|I)}{H(s_k|G)}. \end{aligned}$$

Taking limits and using continuity of $F(\cdot|I)$ and $F(\cdot|G)$, we have

$$\begin{aligned} \left(\frac{1 - F(s_n|I)}{1 - F(s_n|G)}\right)^n \lim_{s \uparrow s_n} \frac{H(s|I)}{H(s|G)} \\ \geq \left(\frac{1 - F(s_k|I)}{1 - F(s_k|G)}\right)^k \left(\frac{F(s_k|I)}{F(s_k|G)}\right)^{n-k} \lim_{s \downarrow s_k} \frac{H(s|I)}{H(s|G)}, \end{aligned}$$

and the assumption of the theorem delivers the desired inequality. \blacksquare

Theorem 2 has a clear interpretation when $P(G) = P(I)$, in which case it follows from the theorem that

$$\frac{P_n(C|I)P(I)}{P_n(C|G)P(G) + P_n(C|I)P(I)} > \frac{P_k(C|I)P(I)}{P_k(C|G)P(G) + P_k(C|I)P(I)}.$$

That is, the probability that the defendant is innocent, conditional on conviction, is higher under unanimity rule than when k votes are required to convict.

To give a more transparent sufficient condition for the result of Theorem 2, we use the following lemma, which establishes that s_k is weakly decreasing in k . That is, as the number of votes required to convict the defendant increases, jurors become more willing to vote for convicting.

Lemma 3 *If $n \geq k' > k \geq 1$, then $s_{k'} \square s_k$.*

We can now state a corollary of Theorem 2.

Corollary 1 *If $H(s|I)/H(s|G)$ is weakly increasing on an open interval including $[s_n, s_k]$, then*

$$\frac{P_n(C|I)}{1 - P_n(A|G)} > \frac{P_k(C|I)}{1 - P_k(A|G)},$$

for all $k = 1, 2, \dots, n - 1$.

The proof, given Theorem 2 and Lemma 3, is straightforward. Applied to the binary signal model, because the ratio of hazard rates is strictly increasing from 0 to 1, we know that the conclusion of the corollary holds if $s_k < 1$. In particular, it holds if $k > \frac{n}{2} + 1$ or if both $k = \frac{n}{2} + 1$ and $\rho \geq 1$. Applied to the exponential model, since the ratio of hazard rates is constant, unanimity rule exhibits a bias toward convicting the innocent compared to any other voting rule.

6 A Jury Theorem

In this section, we investigate the asymptotic properties of jury decision rules as jury size increases. We show that, as the jury size becomes large, the probability of making a wrong judgment goes to zero for every voting rule *except* unanimity. Rather than specify the number of votes needed to

convict, we will here define a rule by the fraction, say α , of votes needed. Thus, unanimity rule will be represented by $\alpha = 1$. Given n , the decision rule requiring k votes to convict would be represented by $\alpha = k/n$. For ease of exposition, we only consider combinations of α and n such that αn is an integer. We write $P_{\alpha,n}(C|I)$ for the probability of conviction conditional on innocence and $P_{\alpha,n}(A|G)$ for the probability of acquittal conditional on guilt, under the unique symmetric, responsive cutoff equilibrium, when the α rule is used and the number of jurors is n .

Theorem 3 *If the ratio of hazard rates is bounded above by some $\overline{H} < \infty$ and below by some $\underline{H} > 0$, then*

$$\liminf_{n \rightarrow \infty} P_{1,n}(C|I) + P_{1,n}(A|G) > 0,$$

and, for all $0 < \alpha < 1$,

$$\limsup_{n \rightarrow \infty} P_{\alpha,n}(C|I) + P_{\alpha,n}(A|G) = 0.$$

Proof: We use the notation $J_{\alpha,n}$ to make explicit the dependence of J on the decision rule and number of jurors. Note that we can write

$$\begin{aligned} J_{\alpha,n}(\overline{s}, s) &= \left(\frac{1 - F(\overline{s}|I)}{1 - F(\overline{s}|G)} \right)^{\alpha n - 1} \left(\frac{F(\overline{s}|I)}{F(\overline{s}|G)} \right)^{n - \alpha n} \frac{f(s|I)}{f(s|G)} - \rho \\ &= [L_{\alpha}(\overline{s})]^n \left(\frac{f(s|I)}{f(s|G)} \right) \left(\frac{1 - F(\overline{s}|G)}{1 - F(\overline{s}|I)} \right) - \rho, \end{aligned}$$

where we define

$$L_{\alpha}(s) = \left(\frac{1 - F(s|I)}{1 - F(s|G)} \right)^{\alpha} \left(\frac{F(s|I)}{F(s|G)} \right)^{1 - \alpha}$$

for all $s \in S$. Note also that $J_{\alpha,n}(s, s)$ can be written

$$J_{\alpha,n}(s, s) = [L_{\alpha}(s)]^n \frac{H(s|I)}{H(s|G)} - \rho.$$

From Theorem 1, for each n there is a unique symmetric, responsive cutoff equilibrium characterized by the cutoff $s_n = \inf\{s \in S \mid J_{\alpha,n}(s, s) \square 0\}$. We claim that $L_\alpha(s_n) \rightarrow 1$. If not, we can extract a subsequence with limit less than one or with limit greater than one. Without loss of generality, we suppose this is true of $\{s_n\}$. In the first case, take m high enough that $L_\alpha(s_m)^m \overline{H} < \rho$. Using continuity of L_α , we can take $s < s_m$ close enough to s_m so that $L_\alpha(s)^m \overline{H} < \rho$. But then

$$J_{\alpha,m}(s, s) < \rho - \rho = 0,$$

contradicting the definition of s_m . In the second case, take m high enough that $L_\alpha(s_m)^m \underline{H} > \rho$. Using continuity of L_α , we can take $s > s_m$ close enough to s_m so that $L_\alpha(s)^m \underline{H} > \rho$. But then

$$J_{\alpha,m}(s, s) > \rho - \rho = 0,$$

contradicting the definition of s_m .

To prove the first part of the theorem, note that, from $L_1(s_n) \rightarrow 1$ and strict first order stochastic dominance ((4) of Lemma 0), $s_n \rightarrow \underline{S}$. By assumption (A1), the likelihood ratio has at most a finite number of discontinuity points, so there exists l such that, for all $m > l$, s_m is a continuity point of $J_{1,m}$. Hence $J_{1,m}(s_m, s_m) = 0$ for all such s_m , by definition. Thus,

$$\frac{P_{1,m}(C|I)}{1 - P_{1,m}(A|G)} = \left(\frac{1 - F(s_m|I)}{1 - F(s_m|G)} \right)^m = \rho \frac{H(s_m|G)}{H(s_m|I)}$$

for all $m > l$. Since $s_n \rightarrow \underline{S}$,

$$\lim_{n \rightarrow \infty} \frac{P_{1,n}(C|I)}{1 - P_{1,n}(A|G)} = \rho \lim_{s_n \rightarrow \underline{S}} \frac{H(s_n|G)}{H(s_n|I)} \geq \frac{\rho}{\underline{H}},$$

and no subsequence of $\{P_{1,m}(C|I) + P_{1,m}(A|G)\}$ can go to zero, proving the first part of the theorem.

To prove the second part of the theorem, take $0 < \alpha < 1$. We first claim that $L_\alpha(s) = 1$ implies $1 - F(s|G) > \alpha > 1 - F(s|I)$. We use the facts that $x^\alpha(1-x)^{1-\alpha}$ is single-peaked at $x = \alpha$ and, by strict first order stochastic dominance, $1 - F(s|I) < 1 - F(s|G)$ for all $s \in S$. If $\alpha \square 1 - F(s|I)$ then $\alpha \square 1 - F(s|I) < 1 - F(s|G)$ and, by single-peakedness,

$$(1 - F(s|I))^\alpha (F(s|I))^{1-\alpha} > (1 - F(s|G))^\alpha (F(s|G))^{1-\alpha},$$

or equivalently $L_\alpha(s) > 1$, a contradiction. Similarly, if $1 - F(s|G) \square \alpha$, then $1 - F(s|I) < 1 - F(s|G) \square \alpha$ and, by single-peakedness, $L_\alpha(s) < 1$, a contradiction establishing the claim. Since L_α is decreasing, continuous, and

$$\begin{aligned} \lim_{s \rightarrow \underline{S}} L_\alpha(s) &= \lim_{s \rightarrow \underline{S}} \left(\frac{f(s|I)}{f(s|G)} \right)^{1-\alpha} > 1 \\ \lim_{s \rightarrow \bar{S}} L_\alpha(s) &= \lim_{s \rightarrow \bar{S}} \left(\frac{f(s|I)}{f(s|G)} \right)^\alpha < 1 \end{aligned}$$

(using L'Hôpital's rule and (3) of Lemma 0), the set $L_\alpha^{-1}(1)$ is a non-empty closed interval, $[s', s'']$, with $\underline{S} < s', s'' < \bar{S}$. By continuity of the distribution functions, we can take $\delta > 0$ such that $1 - F(s|G) > \alpha > 1 - F(s|I)$ for all $s \in [s' - \delta, s'' + \delta]$. Since $L_\alpha(s_n) \rightarrow 1$, there exists l such that, for all $m > l$, $s_m \in [s' - \delta, s'' + \delta]$.

The last part of the proof is a straightforward application of the law of large numbers. To prove $P_{\alpha,n}(A|G) \rightarrow 0$, define the probability space $S^\infty = S \times S \times \dots$ with the product measure generated by μ_G , and define the sequence X_1, X_2, \dots of i.i.d. random variables satisfying

$$X_i = \begin{cases} 1 & \text{if } s_i \geq s'' + \delta \\ 0 & \text{else,} \end{cases}$$

where s_i is the i th component of $(s_1, s_2, \dots) \in S^\infty$. By the strong law of large numbers, $\frac{1}{n} \sum_{i=1}^n X_i$ converges almost surely to $1 - F(s'' + \delta|G)$ as n

goes to infinity. In particular, it converges in probability:

$$P\left(1 - F(s'' + \delta|G) - \frac{1}{n} \sum_{i=1}^n X_i > \epsilon\right) \rightarrow 0$$

for all $\epsilon > 0$. Define the sequence Y_1, Y_2, \dots of random variables as

$$Y_n = \frac{1}{n} \#\{i \square n \mid s_i \geq s_n\},$$

and note that, for $m > l$, $Y_n \geq \frac{1}{n} \sum_{i=1}^n X_i$. Hence,

$$P(1 - F(s'' + \delta|G) - Y_n > \epsilon) \rightarrow 0,$$

or equivalently,

$$P(Y_n < 1 - F(s'' + \delta|G) - \epsilon) \rightarrow 0$$

for all $\epsilon > 0$. Since $1 - F(s'' + \delta|G) > \alpha$, we can set $\epsilon = 1 - F(s'' + \delta|G) - \alpha$, yielding $P(Y_n < \alpha) \rightarrow 0$. That is, the probability that the fraction of jurors voting to convict a guilty defendant is smaller than α goes to zero as the size of the jury goes to infinity. Therefore, $P_{\alpha,n}(A|G) \rightarrow 0$. The proof that $P_{\alpha,n}(C|I) \rightarrow 0$ is analogous. ■

Note that the alternative condition

$$\lim_{s \downarrow \underline{s}} \frac{f(s|I)}{f(s|G)} < \infty \quad \text{and} \quad \lim_{s \uparrow \bar{s}} \frac{f(s|I)}{f(s|G)} > 0$$

implies the existence of bounds $\overline{H} < \infty$ and $\underline{H} > 0$, since $(1 - F(s|I))/(1 - F(s|G))$ is bounded below by one and bounded above by $\lim_{s \downarrow \underline{s}} f(s|I)/f(s|G)$. This sufficient condition for the theorem has a clear interpretation: a juror cannot be arbitrarily sure that the defendant is guilty or arbitrarily sure that he is innocent. The alternative condition is stronger than the one used in the theorem: it is fulfilled in the binary signal model but not in the exponential model, where the likelihood ratio goes to zero as s goes to infinity.

Appendix

Proposition 1 *Given σ_{-i} , the ex ante payoff to juror i of σ_i is an affine transformation of*

$$\begin{aligned} & \left(\int \sigma_i(s) \mu_I(ds) \right) u(C|I) P_{\sigma_{-i}}(piv|I) P(I) \\ & - \left(\int \sigma_i(s) \mu_G(ds) \right) u(A|G) P_{\sigma_{-i}}(piv|G) P(G). \end{aligned}$$

Proof: Let N denote the set of jurors. Note that

$$\begin{aligned} P_\sigma(C|I) &= \left(\int \sigma_i(s) \mu_I(ds) \right) \sum_{\substack{M \subseteq N \\ |M| \geq k-1 \\ i \notin M}} \Pi_{j \in M} \left(\int \sigma_j(s) \mu_I(ds) \right) \cdot \\ & \quad \left[\Pi_{\substack{j \notin M \\ j \neq i}} \left(\int [1 - \sigma_j(s)] \mu_I(ds) \right) \right] \\ & + \left(\int [1 - \sigma_i(s)] \mu_I(ds) \right) \sum_{\substack{M \subseteq N \\ |M| \geq k \\ i \notin M}} \Pi_{j \in M} \left(\int \sigma_j(s) \mu_I(ds) \right) \cdot \\ & \quad \left[\Pi_{\substack{j \notin M \\ j \neq i}} \left(\int [1 - \sigma_j(s)] \mu_I(ds) \right) \right], \end{aligned}$$

and

$$\begin{aligned} P_\sigma(A|G) &= \left(\int \sigma_i(s) \mu_G(ds) \right) \sum_{\substack{M \subseteq N \\ |M| < k-1 \\ i \notin M}} \Pi_{j \in M} \left(\int \sigma_j(s) \mu_G(ds) \right) \cdot \\ & \quad \left[\Pi_{\substack{j \notin M \\ j \neq i}} \left(\int [1 - \sigma_j(s)] \mu_G(ds) \right) \right] \\ & + \left(\int [1 - \sigma_i(s)] \mu_G(ds) \right) \sum_{\substack{M \subseteq N \\ |M| < k \\ i \notin M}} \Pi_{j \in M} \left(\int \sigma_j(s) \mu_G(ds) \right) \cdot \\ & \quad \left[\Pi_{\substack{j \notin M \\ j \neq i}} \left(\int [1 - \sigma_j(s)] \mu_G(ds) \right) \right]. \end{aligned}$$

Inserting these expressions into $u(C|I)P_\sigma(C|I)P(I) + u(A|G)P_\sigma(A|G)P(G)$ and simplifying, we get

$$\begin{aligned}
& \left(\int \sigma_i(s) \mu_I(ds) \right) u(C|I) \sum_{\substack{M \subseteq N \\ |M|=k-1 \\ i \notin M}} \prod_{j \in M} \left(\int \sigma_j(s) \mu_I(ds) \right) \cdot \\
& \quad \prod_{\substack{j \notin M \\ j \neq i}} \left(\int [1 - \sigma_j(s)] \mu_I(ds) \right) \Big] P(I) \\
& - \left(\int \sigma_i(s) \mu_G(ds) \right) u(A|G) \sum_{\substack{M \subseteq N \\ |M|=k-1 \\ i \notin M}} \prod_{j \in M} \left(\int \sigma_j(s) \mu_G(ds) \right) \cdot \\
& \quad \prod_{\substack{j \notin M \\ j \neq i}} \left(\int [1 - \sigma_j(s)] \mu_G(ds) \right) \Big] P(G) + \text{constant},
\end{aligned}$$

where the last term is independent of σ_i . ■

Lemma 0

$$(1) \quad \frac{1 - F(s|I)}{1 - F(s|G)} \text{ and } \frac{F(s|I)}{F(s|G)} \text{ are weakly decreasing.}$$

$$(2) \quad \frac{1 - F(s|I)}{1 - F(s|G)} \square \frac{f(s|I)}{f(s|G)} \square \frac{F(s|I)}{F(s|G)} \text{ for all } s \in S.$$

$$(3) \quad \lim_{s \downarrow \underline{s}} \frac{f(s|I)}{f(s|G)} > 1 \text{ and } \lim_{s \uparrow \bar{s}} \frac{f(s|I)}{f(s|G)} < 1.$$

$$(4) \quad F(s|I) > F(s|G) \text{ for all } s \in S.$$

$$(5) \quad \text{If } \frac{f(\bar{s}|I)}{f(\bar{s}|G)} > \lim_{s \uparrow \bar{s}} \frac{f(s|I)}{f(s|G)} \text{ then } \frac{1 - F(s|I)}{1 - F(s|G)} > \frac{1 - F(\bar{s}|I)}{1 - F(\bar{s}|G)}$$

for all $s < \bar{s}$ and all $\bar{s} \in S$.

$$(6) \quad \text{If } \frac{f(\bar{s}|I)}{f(\bar{s}|G)} < \lim_{s \downarrow \underline{s}} \frac{f(s|I)}{f(s|G)} \text{ then } \frac{F(s|I)}{F(s|G)} < \frac{F(\bar{s}|I)}{F(\bar{s}|G)}$$

for all $s > \bar{s}$ and all $\bar{s} \in S$.

Proof: Results (1) and (2) follow from (A1)–(A3) and are well-known. Result (3) follows easily from (A3) and (A4). Result (4), stated with weak inequality (that is, first order stochastic dominance) is a well-known implication of (A3). Strict inequality follows from result (3) above. If (5) fails then, by (1) above,

$$\frac{1 - F(\bar{s}|I)}{1 - F(\bar{s}|G)} = \frac{1 - F(s|I)}{1 - F(s|G)} \text{ for all } s \in [\hat{s}, \bar{s}]$$

for some $\hat{s} < \bar{s}$. Consequently,

$$D \left(\frac{1 - F(s|I)}{1 - F(s|G)} \right) = 0,$$

or equivalently

$$(7) \quad \frac{1 - F(s|I)}{1 - F(s|G)} = \frac{f(s|I)}{f(s|G)},$$

for all $s \in [\hat{s}, \bar{s}] \cap S^d$. Note that

$$\frac{f(\bar{s}|I)}{f(\bar{s}|G)} > \lim_{s \uparrow \bar{s}} \frac{f(s|I)}{f(s|G)}$$

and assumption (A4) imply that there is a $s' \geq \bar{s}$ such that

$$\frac{f(\bar{s}|I)}{f(\bar{s}|G)} > \frac{f(x|I)}{f(x|G)}$$

for all $x \geq s'$. Taking any $s \in (\hat{s}, \bar{s})$ and using (7),

$$\begin{aligned} 1 - F(s|I) &= \int_s^{\bar{s}} \frac{f(s|I)}{f(s|G)} f(x|G) dx \\ &= \int_s^{s'} \frac{f(s|I)}{f(s|G)} f(x|G) dx + \int_{s'}^{\bar{s}} \frac{f(s|I)}{f(s|G)} f(x|G) dx \\ &> \int_s^{s'} \frac{f(x|I)}{f(x|G)} f(x|G) dx + \int_{s'}^{\bar{s}} \frac{f(x|I)}{f(x|G)} f(x|G) dx \\ &= 1 - F(s|I), \end{aligned}$$

where the inequality follows from assumption (A3) and our choice of s' . But this is a contradiction, establishing (5). The proof of (6) is analogous. \blacksquare

Lemma 1 *Given responsive strategies σ_{-i} for jurors other than i , a strategy σ_i is a best response for i if and only if it satisfies the following a.e.:*

$$(8) \quad \sigma_i(s) = \begin{cases} 1 & \text{if } J(\sigma_{-i}, s) < 0 \\ 0 & \text{if } J(\sigma_{-i}, s) > 0. \end{cases}$$

If the likelihood ratio is strictly decreasing at $\inf\{s \in S \mid J(\sigma_{-i}, s) \square 0\}$, σ_i is a best response for i if and only if it is equivalent a.e. to the following cutoff strategy $\tilde{\sigma}_i$:

$$\tilde{\sigma}_i(s) = \begin{cases} 1 & \text{if } J(\sigma_{-i}, s) \square 0 \\ 0 & \text{else.} \end{cases}$$

Proof: Suppose σ_i satisfies (8). Take any strategy σ'_i , and define the sets

$$\begin{aligned} V &= \{s \in S \mid J(\sigma_{-i}, s) < 0 \text{ and } \sigma'_i(s) < 1\} \\ W &= \{s \in S \mid J(\sigma_{-i}, s) > 0 \text{ and } \sigma'_i(s) > 0\}. \end{aligned}$$

Note that $\sigma_i(s) = 1$ for all $s \in V$ and $\sigma_i(s) = 0$ for all $s \in W$. Thus, using Proposition 1, the payoff from σ_i to juror i exceeds the payoff from σ'_i by

$$\begin{aligned} &\int_V (1 - \sigma'_i(s)) [u(C|I)P_{\sigma_{-i}}(piv|I)P(I)f(s|I) \\ &\quad - u(A|G)P_{\sigma_{-i}}(piv|G)P(G)f(s|G)] ds \\ &- \int_W \sigma'_i(s) [u(C|I)P(I)P_{\sigma_{-i}}(piv|I)f(s|I) \\ &\quad - u(A|G)P(G)P_{\sigma_{-i}}(piv|G)f(s|G)] ds. \end{aligned}$$

By construction, $s \in V$ implies $J(\sigma_{-i}, s) < 0$, which implies that the integrand of the first integral is positive; $s \in W$ implies $J(\sigma_{-i}, s) > 0$, which implies that the integrand of the second integral is negative. Since σ'_i violates (8) if and only if $V \cup W$ has positive measure, any strategy satisfying (8) is a best response and any strategy violating (8) is not.

If the likelihood ratio is strictly decreasing at $\inf\{s \in S \mid J(\sigma_{-i}, s) \square 0\}$, $J(\sigma_{-i}, s) = 0$ has at most one solution, and hence (8) implies σ_i is equivalent a.e. to $\tilde{\sigma}_i$, a cutoff strategy. ■

Lemma 2 *J is continuous and weakly decreasing in its first argument. In addition,*

$$\lim_{s \downarrow \underline{s}} J(s, s) > 0 \quad \text{and} \quad \lim_{s \uparrow \bar{s}} J(s, s) < 0.$$

Finally, $J(s, s) = 0$ has at most one solution.

Proof: The first part follows directly from (A1), (A3), and (1) of Lemma 0.

Note that

$$\begin{aligned}\lim_{s \downarrow \underline{S}} J(s, s) &= \lim_{s \downarrow \underline{S}} \left(\frac{F(s|I)}{F(s|G)} \right)^{n-k} \lim_{s \downarrow \underline{S}} \frac{f(s|I)}{f(s|G)} - \rho \\ &= \left(\lim_{s \downarrow \underline{S}} \frac{f(s|I)}{f(s|G)} \right)^{n-k+1} - \rho \\ &> 0,\end{aligned}$$

and

$$\begin{aligned}\lim_{s \uparrow \bar{S}} J(s, s) &= \lim_{s \uparrow \bar{S}} \left(\frac{1 - F(s|I)}{1 - F(s|G)} \right)^{k-1} \lim_{s \uparrow \bar{S}} \frac{f(s|I)}{f(s|G)} - \rho \\ &= \left(\lim_{s \uparrow \bar{S}} \frac{f(s|I)}{f(s|G)} \right)^k - \rho \\ &< 0,\end{aligned}$$

where we make use of L'Hôpital's rule, (A4), and (3) of Lemma 0. This proves the second part of the lemma.

For the third part, take any signal \bar{s} such that $J(\bar{s}, \bar{s}) = 0$. We claim that, for all $s' > \bar{s}$, $J(s', s') < J(\bar{s}, \bar{s}) = 0$. By the definition of J and (1) of Lemma 0, if

$$\frac{f(s'|I)}{f(s'|G)} < \frac{f(\bar{s}|I)}{f(\bar{s}|G)},$$

we are done. Thus, by assumption (A3), we suppose the two likelihood ratios are equal. Note that, by assumption (A4), either

$$\frac{f(\bar{s}|I)}{f(\bar{s}|G)} < \lim_{s \downarrow \underline{S}} \frac{f(s|I)}{f(s|G)}$$

or

$$\frac{f(s'|I)}{f(s'|G)} = \frac{f(\bar{s}|I)}{f(\bar{s}|G)} > \lim_{s \uparrow \bar{S}} \frac{f(s|I)}{f(s|G)}.$$

If the first inequality holds, then

$$\frac{F(s'|I)}{F(s'|G)} < \frac{F(\bar{s}|I)}{F(\bar{s}|G)}$$

by (6) of Lemma 0. If the second holds, then

$$\frac{1 - F(s'|I)}{1 - F(s'|G)} < \frac{1 - F(\bar{s}|I)}{1 - F(\bar{s}|G)}.$$

by (5) of Lemma 0.

We look at three cases. If $1 < k < n$, then by the definition of J and the preceding discussion, we are done. If $k = n$, then

$$J(s, s) = \left(\frac{1 - F(s|I)}{1 - F(s|G)} \right)^{n-1} \frac{f(s|I)}{f(s|G)} - \rho.$$

If

$$\frac{f(\bar{s}|I)}{f(\bar{s}|G)} > \lim_{s \uparrow \bar{s}} \frac{f(s|I)}{f(s|G)},$$

then by (5) of Lemma 0 we are done. Otherwise, we have

$$f(s|I) = f(s|G) \cdot \lim_{s \uparrow \bar{s}} \frac{f(s|I)}{f(s|G)}$$

for all $s \geq \bar{s}$. Then, after integrating and rearranging terms,

$$\frac{1 - F(\bar{s}|I)}{1 - F(\bar{s}|G)} = \lim_{s \uparrow \bar{s}} \frac{f(s|I)}{f(s|G)}.$$

Using $J(\bar{s}, \bar{s}) = 0$, we get

$$\rho = \left(\lim_{s \uparrow \bar{s}} \frac{f(s|I)}{f(s|G)} \right)^n.$$

But

$$\lim_{s \uparrow \bar{s}} \frac{f(s|I)}{f(s|G)} < 1,$$

by (3) of Lemma 0, so

$$\lim_{s \uparrow \bar{s}} \frac{f(s|I)}{f(s|G)} > \rho,$$

contradicting (A4). The case $k = 1$ is analogous. This establishes the claim that $s' > \bar{s}$ implies $J(s', s') < 0$. Therefore, $J(s, s) = 0$ has at most one solution. ■

Lemma 3 *If $n \geq k' > k \geq 1$, then $s_{k'} \square s_k$.*

Proof: Recall that

$$J_k(s, s) = \left(\frac{1 - F(s | I)}{1 - F(s | G)} \right)^{k-1} \left(\frac{F(s | I)}{F(s | G)} \right)^{n-k} \frac{f(s | I)}{f(s | G)} - \rho.$$

By (2) of Lemma 0, given arbitrary $s \in S$,

$$\frac{1 - F(s | I)}{1 - F(s | G)} \square \frac{F(s | I)}{F(s | G)},$$

which implies $J_{k'}(s, s) \square J_k(s, s)$ for $k' > k$. This implies

$$\{s \in S \mid J_k(s, s) \square 0\} \subseteq \{s \in S \mid J_{k'}(s, s) \square 0\},$$

from which we conclude $s_{k'} \square s_k$. ■

References

- [1] Austen-Smith, D. and J. Banks (1996) Information Aggregation, Rationality and the Condorcet Jury Theorem, *American Political Science Review* 90: 34-45.
- [2] Condorcet, Marquis de (1785) *Essais sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris.

- [3] Coughlan, P. (1997) In Defense of Unanimous Jury Verdicts: Mistrials, Communication, and Sincerity. Social Science Working Paper 1012, California Institute of Technology.
- [4] Feddersen, T. and W. Pesendorfer (1997) Voting Behavior and Information Aggregation in Elections with Private Information, *Econometrica* 65: 1029-1058.
- [5] Feddersen, T. and W. Pesendorfer (1998) Convicting the Innocent: The Inferiority of Unanimous Jury Verdicts, *American Political Science Review* 92: 23-35.
- [6] Grofman, B. and S. Feld (1988) Rousseau's General Will: A Condorcetian Perspective, *American Political Science Review* 82: 567-576.
- [7] Ladha, K. (1992) The Condorcet Jury Theorem, Free Speech, and Correlated Votes, *American Journal of Political Science* 36: 617-634.
- [8] McKelvey, R. and T. Palfrey (1998) An Experimental Study of Jury Decisions. Mimeo., California Institute of Technology.
- [9] McLennan, A. (1998) Consequences of the Condorcet Jury Theorem for Beneficial Information Aggregation by Rational Players, *American Political Science Review* 92: 413-418.
- [10] Miller, N. (1986) Information, Electorates, and Democracy: Some Extensions and Interpretations of the Condorcet Jury Theorem, *Information Pooling and Group Decision Making*, ed. B. Grofman and G. Owen. Greenwich, CT: JAI Press.

- [11] Myerson, R. (1997) Extended Poisson Games and the Condorcet Jury Theorem. Discussion Paper 1103, Center for Mathematical Studies in Economics and Management Science, Northwestern University.
- [12] Young, P. (1988) Condorcet's Theory of Voting, *American Political Science Review* 82: 1231-1244.