Economics Working Paper Series
Faculty of Business, Economics and Law, AUT

Unanimity under Ambiguity*

Simona Fabrizi, Steffen Lippert, Addison Pan & Matthew Ryan

2024/01

* This paper supersedes the working paper released in 2021 found here.

Unanimity under Ambiguity*

Simona Fabrizi[†] Steffen Lippert[‡] Addison Pan[§] Matthew Ryan[¶]

March 4, 2024

Abstract

This paper considers a binary decision to be made by a committee – canonically, a jury – through a voting procedure. Each juror must vote on whether a defendant is guilty or not guilty. The voting rule aggregates the votes to determine whether the defendant is convicted or acquitted. We focus on the unanimity rule (convict if and only if all vote guilty), and we consider jurors who share ambiguous prior beliefs as in Ellis (2016). Our contribution is twofold. First, we identify all symmetric equilibria of these voting games. Second, we show that ambiguity may drastically undermine McLennan's (1998) results on decision quality: unlike in the absence of ambiguity, the ex ante optimal symmetric strategy profile need not be an equilibrium; indeed, there are games for which it is possible to reduce both types of error starting from any (non-trivial) equilibrium.

Keywords: ambiguous priors, voting problems, decision quality.

JEL codes: C02, D71, D81.

*Financial support from the Royal Society of New Zealand through Marsden Fund grant UOA1617 is gratefully acknowledged. We would also like to thank Hülya Eraslan and Ronald Stauber for helpful comments on earlier drafts, as well as audiences at PET 2022 (Marseille), SAET 2022 (Canberra), AETW 2023 (Canberra), the University of Auckland, Victoria University of Wellington, and the Online Social Choice and Welfare Seminar Series.

[†]Department of Economics and Centre for Mathematical Social Science, University of Auckland, New Zealand. E-mail: `s.fabrizi@auckland.ac.nz`.

[‡]Department of Economics and Centre for Mathematical Social Science, University of Auckland, New Zealand. E-mail: `s.lippert@auckland.ac.nz`.

[§]International Office and Centre for Mathematical Social Science, University of Auckland, New Zealand. E-mail: `addison.pan@auckland.ac.nz`.

[¶]`Corresponding author`. School of Economics and Finance, Auckland University of Technology, New Zealand, and Centre for Mathematical Social Science, University of Auckland, New Zealand. E-mail: `matthew.ryan@aut.ac.nz`.

1 Introduction

This paper considers a binary decision to be made by a committee – canonically, a jury – through a voting procedure. Each juror must vote on whether a defendant is guilty or not guilty. The voting rule aggregates the votes to determine whether the defendant is convicted or acquitted.

We consider jurors who share *ambiguous* prior beliefs as in Ellis (2016). However, unlike Ellis, who considers the majority rule, we focus on the unanimity rule (convict if and only if all vote guilty); we allow an asymmetric utility function (lower utility from convicting the innocent than from acquitting the guilty); and we identify *all* symmetric equilibria of these voting games. A companion paper (Fabrizi *et al.*, 2022) examines the unanimity rule when ambiguity affects the signals received by jurors rather than the prior. That paper showed that signal ambiguity has negligible impact on the structure of equilibria, and generally improves decision quality by lowering the probability of the more serious error. By contrast, prior ambiguity substantially complicates the structure of equilibria and can seriously degrade decision quality: there may exist symmetric profiles that reduce *both* types of error relative to the (non-trivial) equilibrium. For cases where the (non-trivial) equilibrium is unique, the model also exhibits interesting comparative statics, which we are currently testing in the lab.

The remainder of this Introduction explains these results in more detail and puts them into the context of the existing literature.

Recall the classical model that formalises the famous analysis of Condorcet. The state (guilt or innocence) is unobserved but jurors share a common prior and jurors receive private signals before voting. A juror may receive a "guilty" signal or an "innocent" signal, with signals being independently and identically distributed conditional on the state. There is no communication amongst jurors between receiving their private information and casting their votes. If the outcome is determined by majority rule (i.e., convict if a majority vote guilty and acquit if a majority vote not guilty), and if all voters vote according to their signals (i.e., vote guilty if they receive a guilty signal and not guilty if they receive an innocent signal – the *informative* voting strategy), then a correct decision becomes certain as the jury size goes to infinity. This is Condorcet's famous result (Young, 1988).

Condorcet *assumed* informative voting behaviour. Modern authors have re-examined Condorcet's result through a game-theoretic lens. In these game-theoretic models all jurors share a common Bernoulli utility function – they are "common interest" games. Austen-Smith and Banks (1996) establish conditions under which informative voting is an equilibrium when the common Bernoulli utility function is the indicator function for a correct decision (convict if guilty, acquit if innocent). McLennan (1998) shows that the *ex ante* optimal strategy profile – and also the *ex ante* optimal *symmetric* profile – is an equilibrium irrespective of the specification of the common Bernoulli utility function. In other words, even if the informative strategy profile is not an equilibrium, there will exist another (symmetric) profile that *is* an equilibrium and is even better from an *ex ante* expected utility standpoint. These results give powerful normative justification for majority voting.

More recently, Ellis (2016) has shown that ambiguity may drastically undermine these normative foundations. In his model, jurors share a *set* of prior probability distributions over the states, and each prior is updated (in the usual manner) based on new information to determine a *set* of posteriors. The jurors in Ellis' model make decisions according to the *maxmin expected utility (MEU)* rule. Ellis does not identify all the (symmetric) equilibria of this modified game, but he does establish some important general results. When juror posterior sets include probabilities of innocence both above and below $\frac{1}{2}$ conditional on either signal, majority rule may be no better than a coin toss: there is an equilibrium in which jurors cast each vote with equal probability, irrespective of the signals received or the number of jurors (Ellis, 2016, Proposition 1). On the other hand, if the signal-conditional posterior intervals on the innocent state are *disjoint*,[1] then the essence of Condorcet's result is preserved (Ellis, 2016, Theorem 2).

Austen-Smith and Banks (1996) and the literature which followed has also considered voting rules other than simple majority. The unanimity rule is especially salient in jury contexts. When analysing this rule it is natural to relax the symmetry of the Bernoulli utility function. Correct decisions are still treated symmetrically but "Type I error" (convicting the innocent) is assigned a lower Bernoulli utility than "Type II error" (acquitting the guilty). Feddersen and Pesendorfer (1998) provide the definitive analysis of the unanimity rule, as well as the intermediate super-majority rules. They show that, paradoxically, the unanimity rule is unique in exhibiting an anti-Condorcet property: the asymptotic probability of convicting the innocent is bounded away from zero.[2] However, it is important to note that McLennan's (1998) result applies to these games. The *ex ante* optimal (symmetic) strategy profile is an equilibrium. Any deficiency in decision quality is inherent in the voting rule itself, not the individual rationality constraints imposed by equilibrium – there is no conflict between individual and collective rationality in behaviour.

Ryan (2021) shows that Feddersen and Pesendorfer's asymptotic result also holds in the unanimity version of Ellis' (2016) model, at least when restricting attention to symmetric strategies. Like Ellis (2016), Ryan (2021) establishes his result without needing to identify all (symmetric) equilibria of the voting game, though he does show that equilibria may take more exotic forms than in the absence of ambiguity.

The present paper fills this gap in the analysis of Ryan (2021). We identify all symmetic equilibria of unanimity voting games with an ambiguous prior. This allows us to evaluate decision quality for fixed jury size. A complicated picture emerges but two clear lessons. First, as in Ellis' (2016) analysis of the majority rule, overlapping posterior in-

[1] In fact, it suffices that the intersection of these posterior intervals has an empty interior.

[2] Specifically, along any convergent sequence of *non-trivial* equilibria. Under the unanimity rule there is always a trivial equilibrium in which jurors vote not guilty irrespective of the signal received, so the defendant is always acquitted.

tervals may fundamentally alter equilibrium behaviour and drastically undermine decision quality. Second, if posterior intervals are disjoint, equilibrium structure mostly resembles the case of no ambiguity, with one exception: the generic possibility of an equilibrium in which jurors cast an innocent vote with positive probability following either signal; but equilibrium behaviour may deviate from the *ex ante* optimum. Ambiguity therefore drives a wedge between individual and collective rationality, despite common interest. To some extent this reflects the well-known potential for conflict between *ex ante* and *ad interim* preferences under MEU (i.e., dynamic inconsistency; see, for instance, Siniscalchi, 2011); ambiguity drives a wedge between mixed and behaviour strategies.[3] However, more striking is the finding that when posteriors overlap, it may be possible to reduce *both* Type I *and* Type II error starting from *any* (non-trivial, symmetric) equilibrium profile. The latter effect cannot be explained by dynamic inconsistency alone; equilibrium constraints directly impede decision quality.

2 The model

2.1 Voting problems

We adopt the model of Ryan (2021), which in turn is a hybrid of Feddersen and Pesendorfer (1998) and Ellis (2016). The model is described in detail in Section 2 of Ryan (2021) so we only provide a brief summary here.

There is a set $I = \{1, 2, ..., N+1\}$ of jurors, with generic member i, which makes a decision $d \in D = \{A, B\}$ by secret ballot. We interpret A as the decision to "acquit" the defendant; hence B corresponds to entering a conviction.[4] We use the same notation for decisions and votes: each juror may vote A for acquittal (the "not guilty" vote) or B for conviction (the "guilty" vote). The outcome is determined by the unanimity rule: the defendant is acquitted – decision $d = A$ is made – unless *all* jurors vote for conviction, in which case decision $d = B$ is made.

The defendant may be innocent or guilty, represented by the state $s \in S = \{a, b\}$, where $s = a$ is the state of innocence and $s = b$ the state of guilt. (Think of b as the state in which the defendant is "bad".) Jurors share common ambiguous prior information about s. The prior probability of $s = a$ is objectively known to lie in the interval $\left[\underline{p}, \overline{p}\right] \subseteq (0, 1)$ but nothing more than this. Prior to casting their vote, juror i receives a private signal $t_i \in T = \{1, 2\}$. Conditional on $s \in S$, these signals are independently and identically distributed with $\Pr(t_i = 1|a) = \Pr(t_i = 2|b) = r \in \left(\frac{1}{2}, 1\right)$.

Let $\Omega = S \times T^I$ denote the state space characterising all *ex ante* uncertainty. Together with r, each $p \in \left[\underline{p}, \overline{p}\right]$ determines a probability over Ω. The (closed and convex) set of

[3]See Ellis (2016, §3.1) for further discussion of this issue, and the rationale for defining equilibrium in terms of behaviour strategies.

[4]Our notation (mostly) follows Ellis (2016) to facilitate comparison with his analysis.

probabilities over Ω determined by $\left[\underline{p},\overline{p}\right]$ is denoted by Π. After receiving their signal, a juror uses the *full Bayesian updating (FBU)* rule to update their beliefs: they update each element of Π using Bayes' rule to obtain a set of posterior probabilities on Ω (Fagin and Halpern, 1990; Jaffray, 1992). The posterior interval for the conditional probability $\Pr(a|t_i = t)$ is independent of i and denoted by $\Pi_t = [\underline{\pi}_t, \overline{\pi}_t]$, with generic element π_t. Since $\left[\underline{p},\overline{p}\right] \subseteq (0,1)$ it follows that $\Pi_t \subseteq (0,1)$.

Voters share a common utility function, $u : D \times S \to \mathbb{R}$, with $u(A,a) = u(B,b) = 1$, $u(A,b) = 0$ and $u(B,a) = -c$, where $c \geq 0$. Thus, A is the "correct" decision in state a and B is the "correct" decision in state b. Ellis' (2016) model (or rather, a special case with binary signals) is obtained by setting $c = 0$. When $c > 0$ convicting the innocent results in lower utility than acquitting the guilty.

Note that:

$$\pi u(B,a) + (1-\pi) u(B,b) \geq \pi u(A,a) + (1-\pi) u(A,b)$$
$$\Leftrightarrow \quad \pi \leq \frac{1}{2+c}.$$

The quantity

$$1 - \left(\frac{1}{2+c}\right) = \frac{1+c}{2+c}$$

is what Feddersen and Pesendorfer (1998) refer to as the "threshold of reasonable doubt"; it is the minimum probability of guilt ($s = b$) necessary to justify the decision to convict. In the absence of ambiguity, it is therefore optimal to vote for conviction iff the juror's posterior probability on $s = b$, after incorporating their private information *and the implications of pivotality*, weakly exceeds this threshold. As noted in Ellis (2016),[5] in the presence of ambiguity we can no longer condition on pivotality when determining optimal voting behaviour. We return to this point below.

A *voting problem* is a vector $V = \left(N, c, r, \underline{p}, \overline{p}\right)$, with $N \in \{1, 2, ...\}$, $c \geq 0$, $r \in \left(\frac{1}{2}, 1\right)$ and $0 < \underline{p} \leq \overline{p} < 1$. The set of all voting problems is denoted by $\mathcal{V}$.

2.2 Best responses

Each voting problem induces a voting game. Let σ_t^i denote the probability that $i \in I$ votes B after observing $t \in T$, and let $\sigma^i = (\sigma_1^i, \sigma_2^i)$ denote i's strategy. We focus on *symmetric* profiles, in which each voter follows the same strategy, so we mostly omit the i superscript in what follows. We therefore abuse notation and refer to $\sigma = (\sigma_1, \sigma_2)$ interchangeably as the strategy of a generic voter in a symmetric profile or as the symmetric profile itself. In the terminology of Feddersen and Pesendorfer (1998), a symmetric profile is *responsive* if $\sigma_2 \neq \sigma_1$ and *non-responsive* if $\sigma_1 = \sigma_2$.

[5] And elaborated in Pan (2019).

Consider a generic voter i who believes that each other voter follows the strategy $\sigma = (\sigma_1, \sigma_2)$. Let ρ_s^σ denote the probability that voter i's vote is pivotal conditional on being in state $s \in S$; let θ_s^σ denote the probability that i is not pivotal *and* that a correct decision is made, conditional on being in state $s \in S$. Since conviction requires unanimity, we have $\theta_a^\sigma = 1 - \rho_a^\sigma$, $\theta_b^\sigma = 0$,

$$\rho_a^\sigma = \left[r\sigma_1 + (1-r)\,\sigma_2\right]^N \tag{1}$$

and

$$\rho_b^\sigma = \left[(1-r)\,\sigma_1 + r\sigma_2\right]^N \tag{2}$$

Our notation modifies that of Ellis (2016) to make explicit the dependence of ρ_a^σ, θ_a^σ and ρ_b^σ on the common strategy, σ, of the other jurors.

After observing their private signal $t \in T$, voter i chooses σ_t^i according to the maxmin expected utility (MEU) rule. Hence:

$$\sigma_t^i \in \arg\max_{x\in[0,1]} \left[\min_{\pi_t \in \Pi_t} V\left(x, \sigma; \pi_t\right)\right] \tag{3}$$

where

$$\begin{aligned} V\left(x,\sigma;\pi_t\right) &= \pi_t\left[\rho_a^\sigma\left(1-x-cx\right)+\theta_a^\sigma-\left(1-\rho_a^\sigma-\theta_a^\sigma\right)c\right] + \left(1-\pi_t\right)\left[\rho_b^\sigma x+\theta_b^\sigma\right] \\ &= \pi_t\left[\rho_a^\sigma\left(1-x-cx\right)+\left(1-\rho_a^\sigma\right)\right] + \left(1-\pi_t\right)\rho_b^\sigma x \end{aligned}$$

Because the minimising posterior in (3) may vary with σ_t^i we can no longer condition on pivotality when determining best responses.

If $\sigma = (0,0)$, it is obvious that any $\sigma_t^i \in [0,1]$ satisfies (3), since $\rho_a^\sigma = \rho_b^\sigma = 0$. Ryan (2021) derives the best response correspondence on the domain of non-trivial symmetric profiles: $\sigma \neq (0,0)$. This is summarised by Figure 1, which reproduces Ryan (2021, Figure 1). In this figure,

$$\pi^*\left(\sigma\right) = \frac{\rho_b^\sigma}{\rho_b^\sigma + (1+c)\,\rho_a^\sigma} = \frac{1}{1+(1+c)\left(\rho_a^\sigma/\rho_b^\sigma\right)} \tag{4}$$

and

$$\hat{\sigma}^*\left(\sigma\right) = \min\left\{\sigma^*\left(\sigma\right), 1\right\}$$

where

$$\sigma^*\left(\sigma\right) = \frac{1}{\rho_b^\sigma + (1+c)\,\rho_a^\sigma} = \frac{\pi^*\left(\sigma\right)}{\rho_b^\sigma} \tag{5}$$

Figure 1 is used to determine best responses as follows. Suppose voter i believes that each rival uses strategy $\sigma \neq (0,0)$ and i has received signal $t \in T$. To identify i's optimal vote, we use σ to calculate $\pi^*(\sigma)$ and $\hat{\sigma}^*(\sigma)$ and locate the point $(\underline{\pi}_t, \overline{\pi}_t)$ in Figure 1.

Voter i's optimal response (σ_t^i) is determined by the coloured region into which $(\underline{\pi}_t, \overline{\pi}_t)$ falls, as indicated in the figure. For example, if $(\underline{\pi}_t, \overline{\pi}_t)$ lies in the green region, then it is optimal to choose $\sigma_t^i = 0$ (i.e., to vote "not guilty"). Along the boundaries between the green, pink and blue regions, multiple optimal values for σ_t^i may exist. Within the pink region, excluding its boundary, there is a unique optimal value for σ_t^i but this value, $\hat{\sigma}^*(\sigma)$, may be strictly between 0 and 1. In this case, randomisation is necessary for an optimal response; randomisation may be a valuable hedge against uncertainty.

The quantities $\pi^*(\sigma)$ and $\sigma^*(\sigma)$ have straightforward interpretations.[6]

The quantity $\pi^*(\sigma)$ is the posterior probability for which juror i would be indifferent about how to vote, given that all other jurors use strategy σ; that is, where $V(0, \sigma; \pi) = V(1, \sigma; \pi)$. This quantity is always well-defined, independent of t and contained in the interval

$$\left[\frac{1}{2+c}, \frac{r^N}{r^N + (1+c)(1-r)^N}\right] \subseteq (0, 1).$$

Thus, if $\underline{\pi}_t < \pi^*(\sigma) < \overline{\pi}_t$ (which corresponds to the pink rectangular region in Figure 1), the posteriors in the set Π_t "disagree" about the optimal vote.

The quantity $\sigma^*(\sigma)$ is the value of σ_t^i for which juror i gets the same utility conditional on each $s \in S$, given that all other jurors use strategy σ. This quantity is always well-defined, strictly positive and independent of t, but it may exceed 1: it may not be feasible to perfectly hedge against uncertainty.

Knowledge of $\pi^*(\sigma)$, $\sigma^*(\sigma)$ and the posterior interval for each $t \in T$ suffices to determine juror i's best response(s) to σ.

2.3 Equilibria

We use "equilibrium" as shorthand for a *symmetric (Bayesian) Nash equilibrium.* Thus, $\sigma = (\sigma_1, \sigma_2)$ is an *equilibrium* iff $\sigma_t^i = \sigma_t$ satisfies (3) for each $t \in T$.

The symmetric profile $\sigma = (0, 0)$ is an equilibrium, albeit a trivial one. Furthermore, we must have $\sigma_2 \geq \sigma_1$ in any non-trivial equilibrium. This follows from the fact that $(\underline{\pi}_1, \overline{\pi}_1) \gg (\underline{\pi}_2, \overline{\pi}_2)$: the point $(\underline{\pi}_1, \overline{\pi}_1)$ will lie *strictly to the northeast* of the point $(\underline{\pi}_2, \overline{\pi}_2)$ when plotted in Figure 1.

It is convenient to partition symmetric strategy profiles with $\sigma_2 \geq \sigma_1$ into the following seven categories, which are depicted graphically in Figure 2.

A: The non-responsive profile $\sigma = (0, 0)$ which guarantees acquittal.

C: The non-responsive profile $\sigma = (1, 1)$ which guarantees conviction.

I: The informative profile $\sigma = (0, 1)$.

[6] See Appendix A of Ryan (2021) for more details.

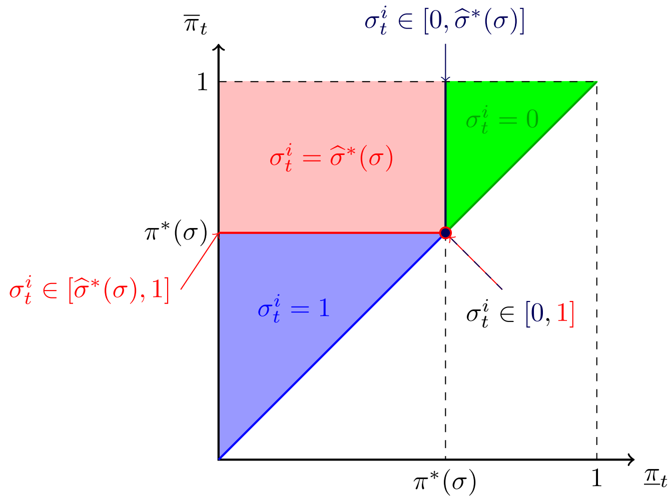

Figure 1: Optimal responses.

FP: Responsive profiles with $0 < \sigma_1 < \sigma_2 = 1$ (i.e., "not guilty" votes imply innocent signals but "guilty" votes do not imply guilty signals). These play a prominent role in Feddersen and Pesendorfer (1998).

DFP: "Dual" FP profiles with $0 = \sigma_1 < \sigma_2 < 1$ (i.e., "guilty" votes imply guilty signals but "not guilty" votes do not imply innocent signals).

MNR: Strictly mixed non-responsive profiles ($0 < \sigma_1 = \sigma_2 < 1$).

MR: Strictly mixed responsive profiles ($0 < \sigma_1 < \sigma_2 < 1$).[7]

The first three categories correspond to the three vertices in Figure 2; the next three to the edges excluding the vertices; and the final category to the interior of the triangle.

Feddersen and Pesendorfer (1998) characterise all equilibria of all voting problems with $\underline{p} = \overline{p}$ (i.e., non-ambiguous voting problems). They show that, fixing N, such voting problems have a generically unique non-trivial equilibrium and that this equilibrium is from category C, I or FP. There is also the non-generic possibility of a continuum of equilibria, comprising the union of categories I, DFP and A (i.e., all the profiles along the lefthand edge of the triangle). This scenario requires that a typical juror be indifferent between conviction and acquittal when all signals are known to be of the guilty variety.

[7] In Ryan (2021) this category bears the acronym "SMR", but this seems inconsistent with the acronym for the MNR category. We have chosen to restore consistency by replacing SMR with MR. (We could, of course, have replaced MNR with SMNR instead, but we chose the more economical path.)

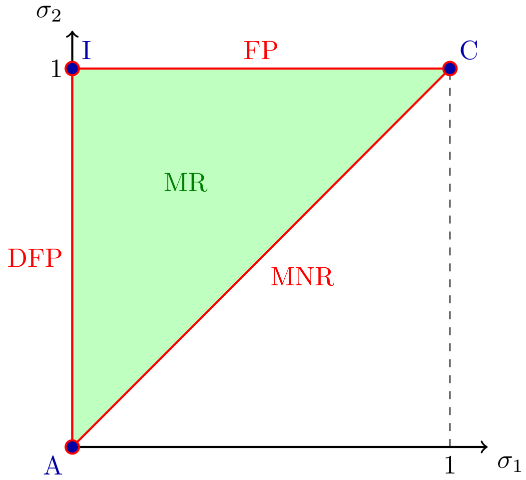

Figure 2: Partitioning strategies with $\sigma_2 \geq \sigma_1$.

Note that there cannot be any equilibrium in category MNR or in category MR in the absence of ambiguity.

The structure of the mapping from non-ambiguous voting problems to non-trivial equilibria is relatively simple and intuitive. For sufficiently low values of $p = \underline{p} = \overline{p}$ we have a unique non-trivial equilibrium at C. When guilt is extremely likely *a priori*, jurors vote to convict no matter what signals they receive. As p increases, the unique non-trivial equilibrium moves weakly westward along the top edge of the triangle towards the informative profile, I. As we move westward through the FP region, pivotality becomes increasingly compelling evidence of guilt so it must be balanced against increasing prior belief in innocence to sustain the indifference of a juror in receipt of an innocent signal. After reaching I, if p continues to increase it will eventually hit a (unique) value at which all profiles on the lefthand edge are equilibria; beyond that (i.e., for even higher values of p) there is no non-trivial equilibrium – only the trivial equilibrium A exists. The "speed" of this anti-clockwise progression around the triangle depends on the values of r, c and N.

Ryan (2021) provides a partial characterisation of the equilibria of voting problems with ambiguity (i.e., when $\underline{p} < \overline{p}$). Unlike the case of no ambiguity, he shows that multiplicity of non-trivial equilibria is a generic possibility,[8] and that we may encounter (again, generically) equilibria from the MR or MNR categories (and even from both at once). Ryan (2021) also observes that there can be at most one non-trivial equilibrium along the top edge of the triangle in Figure 1 (*ibid.*, Lemma 4.3); and that there are no equilibria along the lefthand edge when N is sufficiently large (*ibid.*, Lemma 4.4). In Section 3, we complete the characterisation of the equilibria of voting problems with ambiguity.

[8] All claims regarding genericity are for arbitrarily fixed N.

2.4 Decision quality

To evaluate decision quality, it is useful to map the symmetric profiles in Figure 2 into the associated state-conditional conviction probabilities:

$$\kappa_a^\sigma \equiv \Pr[d = C \mid s = a] = [r\sigma_1 + (1-r)\sigma_2]^{N+1}$$

$$\kappa_b^\sigma \equiv \Pr[d = C \mid s = b] = [(1-r)\sigma_1 + r\sigma_2]^{N+1}$$

This is done in Figure 3, whose derivation can be found in Appendix A. Note that the set of state-conditional conviction probabilities is convex.

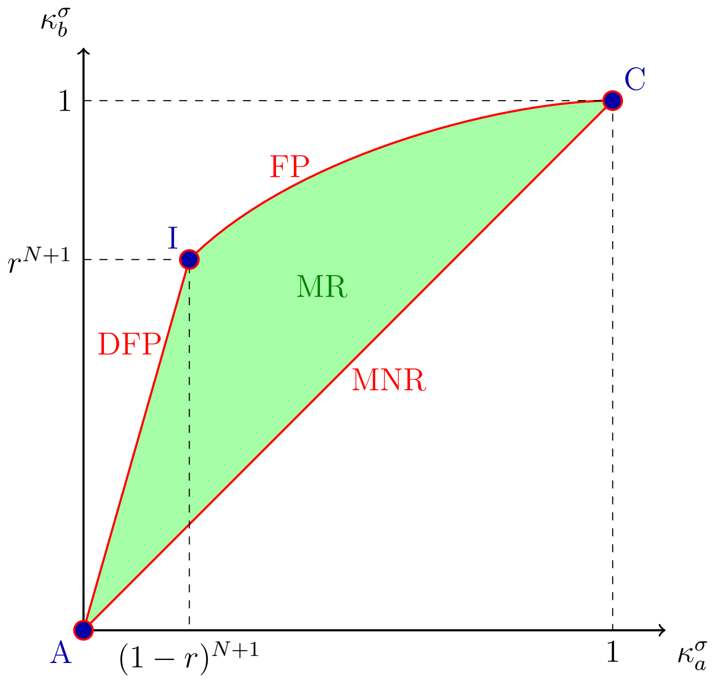

Figure 3: State-conditional conviction probabilities

We can use Figure 3 to evaluate decision quality, since κ_a^σ measures the probability of Type I error and $1-\kappa_b^\sigma$ the probability of Type II error. It is obviously desirable to have κ_a^σ as low as possible and κ_b^σ as high as possible, so the Pareto frontier comprises the northern and western boundaries of Figure 3 – the portion joining the origin (corresponding to profile A) to the point corresponding to profile C. As we move up this frontier, from A towards C, the probability of Type I error increases and the probability of Type II error diminishes.

From Feddersen and Pesendorfer's (1998) results, we observe that any equilibrium lies on this Pareto frontier when $\underline{p} = \overline{p}$ (no ambiguity). The *ex ante* optimal point on the frontier maximises

$$p\left[(1-\kappa_a^\sigma) - c\kappa_a^\sigma\right] + (1-p)\,\kappa_b^\sigma$$

when $\underline{p} = \overline{p} = p$. This function has linear contours with slope

$$\frac{p\,(1+c)}{1-p}.$$

According to McLennan (1998, Theorem 2) any *ex ante* optimal point is an equilibrium. Furthermore, it is not hard to show that when $\underline{p} = \overline{p}$ and a non-trivial equilibrium exists, any non-trivial equilibrium is *ex ante* optimal (see Proposition A.1 in Appendix A).

When $\underline{p} < \overline{p}$ (ambiguity is present), the natural criterion for evaluating *ex ante* expected decision quality is

$$\min_{p\in\left[\underline{p},\overline{p}\right]} \; p\left[(1-\kappa_a^\sigma) - c\kappa_a^\sigma\right] + (1-p)\,\kappa_b^\sigma$$

The minimising p depends on whether $\kappa_b^\sigma \gtreqless 1-(1+c)\,\kappa_a^\sigma$ so indifference contours are now kinked, with convex upper contour sets as depicted in Figure 4. In Section 3.2 we evaluate decision quality in the presence of ambiguity and establish that there exists an open set of voting problems (for any given N) in which all non-trivial equilibria are strictly inside the Pareto frontier. To do this, we first need a complete description of all equilibria for voting problems with ambiguity. That task is taken up in the next section.

3 Characterising equilibria under ambiguity

The mapping from each *non-ambiguous* voting problem (i.e., one with $\underline{p} = \overline{p}$) to its set of equilibria is detailed in Feddersen and Pesendorfer (1998) and briefly summarised above. When we expand the domain of this mapping to include the ambiguous voting problems, the structure of this mapping becomes substantially more complicated. We describe it here. We first state a series of results that construct pieces of this mapping. We then summarise the entire structure graphically and discuss its properties. The reader may wish to peek ahead at the final summary diagrams (Figures 6-7 and 10-12) before coming back to the Propositions. Since profile A is always an equilibrium (the so-called trivial equilibrium) we focus on the non-trivial equilibria.

First, let us recall the relevant results from Ryan (2021):

Proposition 3.1 (Ryan, 2021, Lemma 4.2) *The profile C is an equilibrium iff*

$$\overline{\pi}_1 \leq \frac{1}{2+c}$$

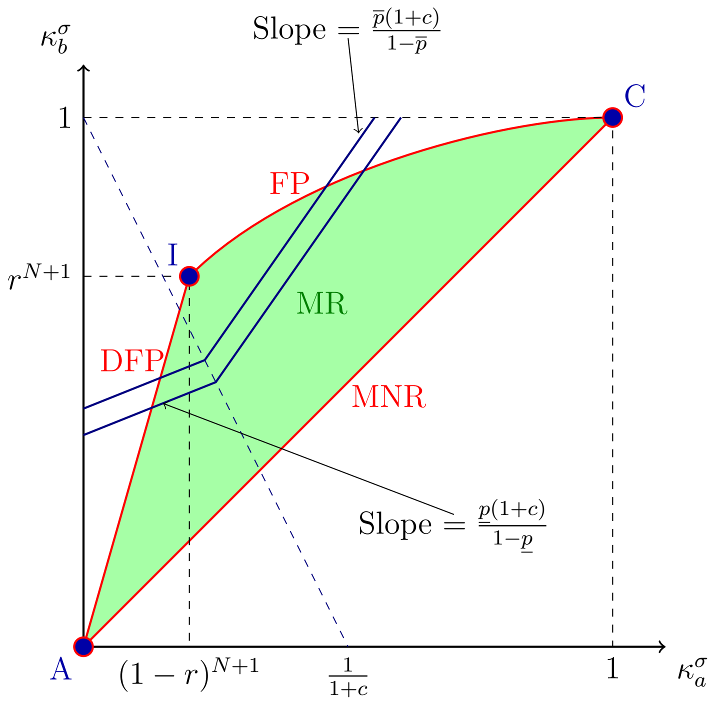

Figure 4: *Ex-ante* indifference curves.

In C, everyone votes to convict. Hence, $\rho_a^\sigma = \rho_b^\sigma = 1$ and therefore $\pi^*(\sigma) = \frac{1}{2+c}$.

Proposition 3.2 (Ryan, 2021, Proposition 4.1) *There exists an equilibrium in category MNR iff*

$$\underline{\pi}_1 \leq \frac{1}{2+c} \leq \overline{\pi}_2.$$

When such an equilibrium exists it always takes the form

$$\sigma_1 = \sigma_2 = \left(\frac{1}{2+c}\right)^{\frac{1}{N+1}}.$$

In an MNR equilibrium we have $0 < \sigma_1 = \sigma_2 < 1$, which implies $\sigma_1 = \sigma_2 = \sigma^*$ and $\rho_a^\sigma = \rho_b^\sigma = \sigma_1^N = \sigma_2^N$, leading to $\pi_t^* = \frac{1}{2+c}$. Then $x = \sigma^*((x,x))$ holds if and only if $x = \left(\frac{1}{2+c}\right)^{\frac{1}{N+1}}$. The voters are indifferent between voting to acquit or convict at $\pi^* = \frac{1}{2+c}$ and perfectly hedge the two states by choosing $\sigma_1 = \sigma_2 = \left(\frac{1}{2+c}\right)^{\frac{1}{N+1}}$.

Proposition 3.3 (Ryan, 2021, Proposition 4.2 [and its proof]) *If σ is a profile in the MR category then σ is an equilibrium iff $\underline{\pi}_1 \leq \overline{\pi}_2$ and σ satisfies one of the following sets of conditions:*

$$\pi^*(\sigma) = \underline{\pi}_1 \quad \textit{and} \quad \sigma_2 = \sigma^*(\sigma) \tag{MR1}$$

$$\pi^*(\sigma) = \overline{\pi}_2 \quad \textit{and} \quad \sigma_1 = \sigma^*(\sigma) \tag{MR2}$$

$$\pi^*(\sigma) = \underline{\pi}_1 = \overline{\pi}_2 \quad \textit{and} \quad \sigma_1 < \sigma^*(\sigma) < \sigma_2 \tag{MR3}$$

Furthermore, there exist constants α_1 and α_2 depending on parameters r, c and N,[9] *and satisfying*

$$\frac{1}{2+c} < \alpha_2 \leq \alpha_1 \leq \frac{r^N}{r^N + (1+c)(1-r)^N},$$

such that: (MR1) has a solution in the MR category iff

$$\frac{1}{2+c} < \underline{\pi}_1 < \alpha_1 \tag{i}$$

and any such solution is unique; (MR2) has a solution in the MR category iff

$$\frac{1}{2+c} < \overline{\pi}_2 < \alpha_2 \tag{ii}$$

and any such solution is unique; (MR3) has a solution in the MR category iff $\underline{\pi}_1 = \overline{\pi}_2$ and (i) holds, in which case (MR3) has a continuum of solutions in the MR category.

In MR equilibria, voters strictly randomise after both signals and the probability to convict is higher after a guilty than after an innocent signal: $0 < \sigma_1 < \sigma_2 < 1$. Consulting Figure 1, we see that this implies either $\underline{\pi}_1 = \pi^*(\sigma)$ or $\overline{\pi}_2 = \pi^*(\sigma)$ or both. We also see that, if $\underline{\pi}_1 = \pi^*(\sigma)$, we must have $\sigma_1 < \sigma_2 = \sigma^*(\sigma)$ and if $\overline{\pi}_2 = \pi^*(\sigma)$, we must have $\sigma_1 = \sigma^*(\sigma) < \sigma_2$.

Note that an equilibrium in the MR or MNR category can only exist if the posterior intervals Π_1 and Π_2 overlap: $\underline{\pi}_1 \leq \overline{\pi}_2$. It is clear that multiple such equilibria may co-exist, and the various possible combinations are summarised in Ryan (2021, Figure 3).

It remains to consider equilibria in categories I, FP and DFP. For this purpose, it is useful to define functions $g : [0,1] \to (0,1)$ and $h : [0,1] \to \mathbb{R}_{++}$ as follows:

$$\begin{aligned} g(\sigma_1) &= \pi^*((\sigma_1, 1)), \\ h(\sigma_1) &= \sigma^*((\sigma_1, 1)). \end{aligned}$$

[9] The constants α_1 and α_2 are defined (albeit implicitly) on pp. 569-571 of Ryan (2021).

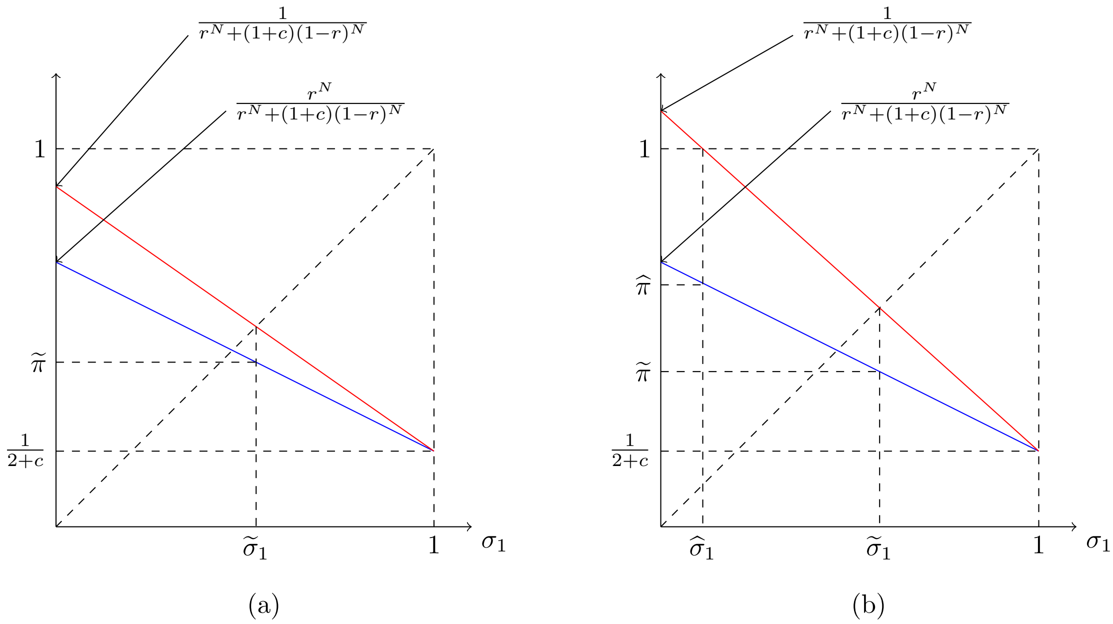

Figure 5: Functions g [blue] and h [red].

Function $g(\sigma_1)$ is the posterior probability that makes voters indifferent between voting to convict or acquit, and function $h(\sigma_1)$ the hedging strategy that delivers the same utility in each state, when all other voters use strategy $(\sigma_1, 1)$. It is easily verified that

$$g(1) = h(1) = \frac{1}{2+c}$$

$$g(0) = \frac{r^N}{r^N + (1+c)(1-r)^N} < h(0) = \frac{1}{r^N + (1+c)(1-r)^N}$$

and that each function is continuous and strictly decreasing. The functions g and h are schematically described in Figure 5. (They are non-linear functions, but since their curvature plays no role in what follows we have depicted them as if linear for convenience.) Note that $g(0) < 1$, while it is possible that $h(0) \geq 1$ as in Panel (b) of Figure 5. Each function has a unique fixed point, which lies in $(0,1)$. We use $\tilde{\sigma}_1$ to denote the fixed point of h and we define $\tilde{\pi} = g(\tilde{\sigma}_1)$. These quantities will play an important role in the analysis to follow. It will also be useful to define

$$\hat{\sigma}_1 = \begin{cases} h^{-1}(1) & \text{if } h(0) \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

and $\hat{\pi} = g(\hat{\sigma}_1)$. The quantities $\tilde{\sigma}_1$, $\tilde{\pi}$, $\hat{\sigma}_1$ and $\hat{\pi}$ are all functions of r, c and N.

We first identify the conditions under which profile I is an equilibrium.[10]

Proposition 3.4 *The informative profile I is an equilibrium iff* $\underline{\pi}_2 \leq g(0) \leq \underline{\pi}_1$ *and either (a)* $\overline{\pi}_2 \leq g(0)$ *or (b)* $h(0) \geq 1$.

The condition $\underline{\pi}_2 \leq g(0) \leq \underline{\pi}_1$ ensures that the juror is willing to acquit for any belief in the posterior interval following an innocent signal, but not for every belief in the posterior interval following a guilty signal. If $\overline{\pi}_2 \leq g(0)$, then a guilty vote is optimal for *all* beliefs in the posterior interval following a guilty signal; otherwise, the voter prefers to hedge. In the latter case, $h(0) \geq 1$ ensures that voting to convict is the closest one can get to a perfect hedge.

Next, we identify the structure of FP equilibria. In an FP equilibrium, the juror must be willing to randomise following an innocent signal. For this it is necessary that $(\underline{\pi}_1, \overline{\pi}_1)$ lie inside the pink rectangle in Figure 1, or on its boundary. If it lies in the interior, then an FP equilibrium must involve perfect hedging after an innocent signal: we must have $\sigma_1 = \sigma^*((0, \sigma_1))$, so $\sigma_1 = \tilde{\sigma}_1$. On the other hand, if we have an FP equilibrium with $(\underline{\pi}_1, \overline{\pi}_1)$ on the eastern boundary of the rectangle, then $g(\sigma_1) = \underline{\pi}_1$ and σ_1 may be below the perfect hedge value; while if we have an FP equilibrium with $(\underline{\pi}_1, \overline{\pi}_1)$ on the southern boundary of the rectangle, then $g(\sigma_1) = \overline{\pi}_1$ and σ_1 may be above the perfect hedge value.

Proposition 3.5 *For any voting problem, there is at most one equilibrium in the FP category. If* σ *is an equilibrium profile in the FP category then*

$$\sigma_1 = \begin{cases} g^{-1}(\underline{\pi}_1) & \text{if } \tilde{\pi} < \underline{\pi}_1 \\ \tilde{\sigma}_1 & \text{if } \underline{\pi}_1 \leq \tilde{\pi} \leq \overline{\pi}_1 \\ g^{-1}(\overline{\pi}_1) & \text{if } \overline{\pi}_1 < \tilde{\pi} \end{cases} \tag{6}$$

Since g is strictly decreasing, g^{-1} is well-defined and also strictly decreasing. When parsing (6) it is useful to recall that $\tilde{\pi}$ and $\tilde{\sigma}_1$ depend on r, c and N, but not on $\underline{p}$ or $\overline{p}$, and that $\tilde{\sigma}_1 = g^{-1}(\tilde{\pi})$.

The following result gives necessary and sufficient conditions for the existence of an equilibrium in the FP category.

Proposition 3.6 *There exists an equilibrium profile in category FP iff one of the following (mutually exclusive) conditions holds:*

(a) $\hat{\pi} \leq \underline{\pi}_1 < g(0)$; *or*

[10] Proofs of all results from this section are in Appendix B.

(b) $\tilde{\pi} \leq \underline{\pi}_1 < \hat{\pi}$ *and* $\overline{\pi}_2 \leq \underline{\pi}_1$*; or*

(c) $\underline{\pi}_1 < \tilde{\pi} < \overline{\pi}_1$ *and* $\overline{\pi}_2 \leq \tilde{\pi}$*; or*

(d) $(2+c)^{-1} < \overline{\pi}_1 \leq \tilde{\pi}$.

Combining Propositions 3.1, 3.4 and 3.6 we deduce the fact stated in Ryan (2021, Lemma 4.3) that there cannot be more than one equilibrium in the union of categories C, I and FP: for profile C to be an equilibrium it is necessary that no point in interval Π_1 exceed $(2+c)^{-1}$; for profile I to be an equilibrium it is necessary that no point in interval Π_1 fall below $g(0)$, which strictly exceeds $(2+c)^{-1}$; and for existence of the unique FP equilibrium it is necessary that Π_1 contains points below $g(0)$ or above $(2+c)^{-1}$.

Finally, we have the DFP profiles. Though dual in structure to the FP profiles, these have a very different equilibrium logic, since pivotality implies that all other jurors received guilty signals. In DFP equilibria, voters who receive an innocent signal vote to acquit, whereas voters who receive a guilty signal randomise. In this equilibrium, guilty votes can only come from jurors with guilty signals. If $\underline{\pi}_2 = g(0)$ then $(\underline{\pi}_2, \overline{\pi}_2)$ sits on the eastern boundary of the pink rectangle in Figure 1 and we have a DFP equilibrium. This is case (a) in Proposition 3.7. The other two cases cover the scenarios in which $(\underline{\pi}_2, \overline{\pi}_2)$ lies on the southern boundary of the rectangle - scenario (8) in case (b) - or in the interior, which gives scenario (9) in case (b). In the latter scenario, σ_2 is a perfect hedge.

Proposition 3.7 *There exists an equilibrium in the DFP category iff one of the following (mutually exclusive) conditions holds:*

(a) $\underline{\pi}_2 = g(0)$*; or*

(b) $\underline{\pi}_2 < g(0) \leq \min\{\underline{\pi}_1, \overline{\pi}_2\}$ *and* $h(0) < 1$.

If (a) holds then DFP profile σ *is an equilibrium iff*

$$\sigma_2 \leq h(0)^{\frac{1}{N+1}} \tag{7}$$

If (b) holds then DFP profile σ *is an equilibrium iff*

$$\overline{\pi}_2 = g(0) \quad \text{and} \quad \sigma_2 \geq h(0)^{\frac{1}{N+1}} \tag{8}$$

or

$$\underline{\pi}_2 < g(0) < \overline{\pi}_2 \quad \text{and} \quad \sigma_2 = h(0)^{\frac{1}{N+1}} \tag{9}$$

Only condition (a) can be satisfied in the absence of ambiguity, so DFP equilibria are non-generic. When ambiguity is present, case (b) opens the possibility of generic DFP equilibria. Since $g(0) \to 1$ as $N \to \infty$, we also confirm the result in Ryan (2021, Lemma 4.4) that no DFP equilibrium exists for sufficiently large juries.

Using the foregoing results, one can, in principle, identify **all** the equilibria of **any** given voting problem.[11] We have stated the equilibrium conditions in terms of the *posterior* probability intervals, as this allows a conveniently compact statement. However, it is not ideal for understanding the mapping from voting problems to the set of equilibria. Posterior probabilities bundle parameter $\underline{p}$ or parameter $\overline{p}$ with parameter r, and the latter usually appears separately in other parts of the equilibrium conditions. It is desirable, for the purpose of understanding equilibrium structure, to re-express the equilibrium conditions explicitly in terms of the "deep" parameters: r, c, N, $\underline{p}$ and $\overline{p}$. This translation is easily accomplished by using the fact that if π_t is the posterior probability given prior p and signal t, then:

$$\pi_1 \;=\; \frac{pr}{pr + (1-p)(1-r)}$$

and

$$\pi_2 \;=\; \frac{p(1-r)}{p(1-r) + (1-p)r}.$$

Hence, for any constant $k \in (0,1)$:

$$\pi_1 = k \quad \Leftrightarrow \quad p \;=\; \frac{1}{1 + \left(\frac{r}{1-r}\right)\left(\frac{1-k}{k}\right)} \tag{10}$$

and

$$\pi_2 = k \quad \Leftrightarrow \quad p \;=\; \frac{1}{1 + \left(\frac{1-r}{r}\right)\left(\frac{1-k}{k}\right)} \tag{11}$$

Therefore, any condition on $\overline{\pi}_1$ or $\overline{\pi}_2$ can be translated into a corresponding restriction on $\overline{p}$ and r; and likewise, any condition on $\underline{\pi}_1$ or $\underline{\pi}_2$ can be translated into a corresponding restriction on $\underline{p}$ and r. We will not record the translated forms of propositions here, but we will freely make use of these translated versions in the following sections.

The mapping from voting problems (i.e., from values for r, c, N, $\underline{p}$ and $\overline{p}$) to equilibria that emerges from our analysis is a somewhat forbidding structure. To keep things manageable, we will first analyse voting problems with disjoint posterior intervals ($\underline{\pi}_1 > \overline{\pi}_2$), and then the ones with overlapping intervals ($\underline{\pi}_1 \leq \overline{\pi}_2$). The former case includes the voting problems with no ambiguity ($\underline{p} = \overline{p}$), so is a natural starting point.

[11] Of course, quantities such as $\tilde{\pi}$ and $\hat{\pi}$ (Proposition 3.6), or α_1 and α_2 (Proposition 3.3), are implicit functions of parameters, and do not, in general, have an explicit formulation.

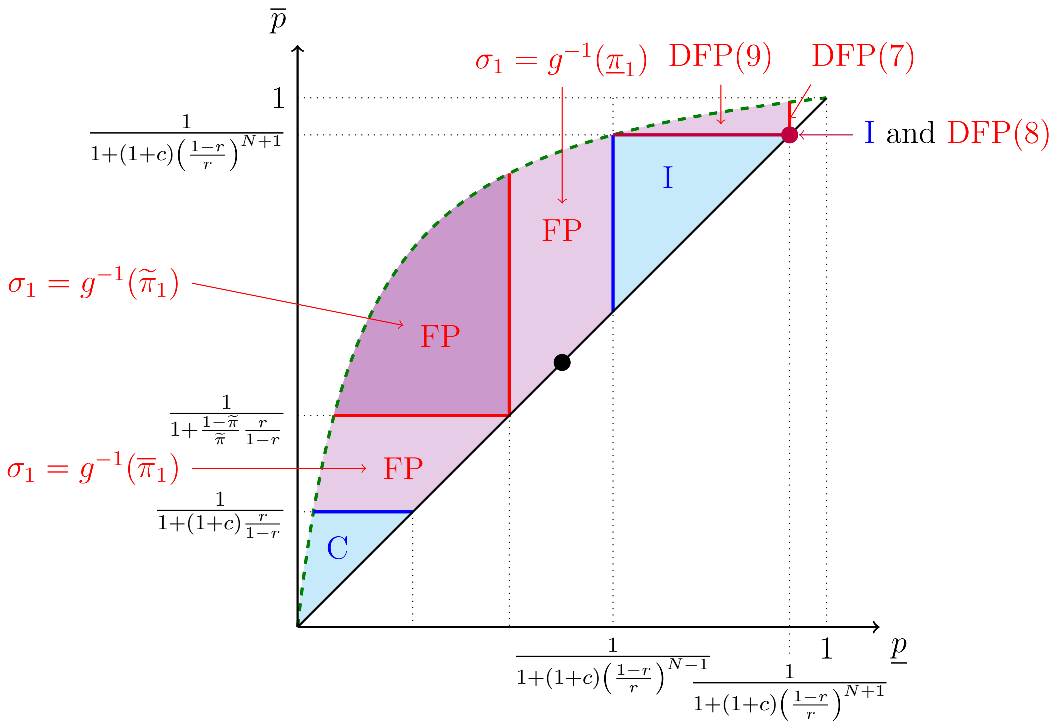

Figure 6: Equilibria when posteriors are disjoint and $h(0) < 1$.

3.1 Disjoint posterior set

When posteriors are disjoint there are no equilibria in the MNR or MR categories. This simplifies matters considerably.

Note that

$$\underline{\pi}_1 > \overline{\pi}_2 \quad \Leftrightarrow \quad \overline{p} < \frac{\underline{p}r^2}{\underline{p}r^2 + \left(1-\underline{p}\right)(1-r)^2} \tag{12}$$

For fixed $r \in \left(\frac{1}{2}, 1\right)$, the righthand side of (12) is a strictly increasing and strictly concave function of $\underline{p} \in (0,1)$, approaching zero as $\underline{p} \to 0$ and approaching unity as $\underline{p} \to 1$. It is depicted as the dashed green curve in Figures 6-7 below. Condition (12) is satisfied at points below this curve.

Figures 6-7 describe the mapping from all voting problems with disjoint posteriors to their non-trivial equilibria.[12] Figure 6 describes the situation for voting problems with

[12] Figures 6-7 are constructed by considering the special cases of Propositions 3.1, 3.4, 3.6 and 3.7 for voting problems satisfying $\underline{\pi}_1 > \overline{\pi}_2$, translating the equilibrium conditions using (10)-(11), and then depicting the corresponding regions of the parameter space graphically.

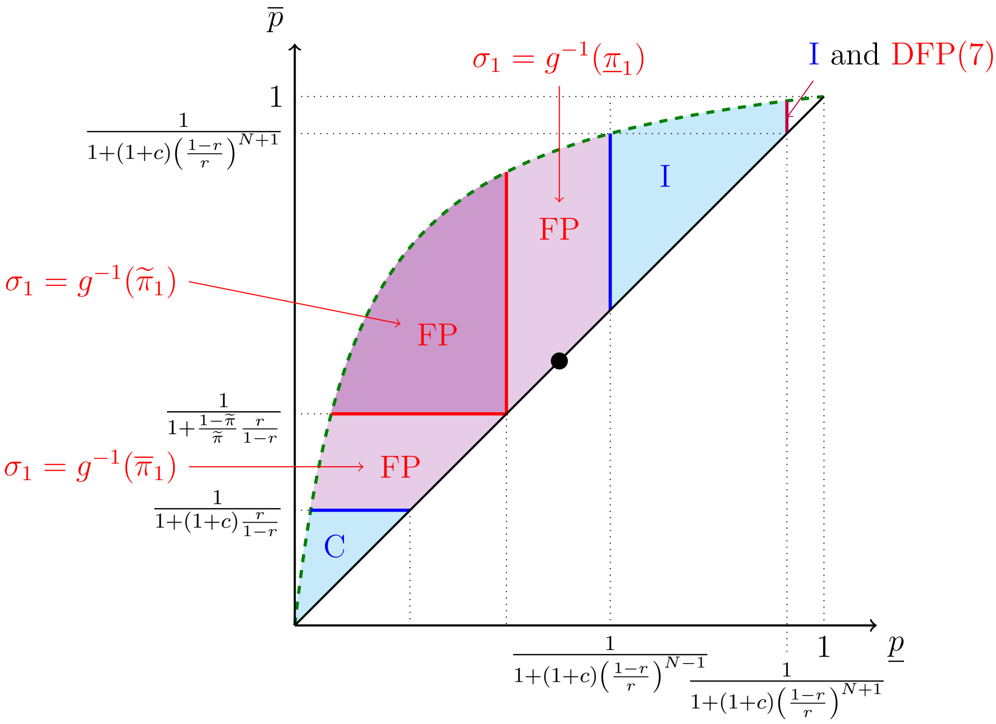

Figure 7: Equilibria when posteriors are disjoint and $h(0) \geq 1$.

$h\,(0) < 1$ and Figure 7 those with $h\,(0) \geq 1$.[13] In each figure, the values of $\underline{p}$ and $\overline{p}$ determine a point in the graph. Based on the location of this point, the coloured regions indicate the nature of any non-trivial equilibria. For example, the black dot in Figure 6 indicates a voting problem with $h\,(0) < 1$ and no ambiguity ($\underline{p} = \overline{p}$); it has a unique non-trivial equilibrium, which is of the FP variety with $\sigma_1 = g^{-1}\left(\overline{\pi}_1\right)$.

The boundaries of the coloured regions depend on the values of the parameters r, c, and N as indicated. In particular:

$$\frac{1}{2+c} < \tilde{\pi} < g\,(0)$$

[13]Condition (12) involves all parameters except N and c. Note that

$$h\,(0) < 1 \quad \Leftrightarrow \quad 1 + c \; < \; \frac{1 - r^N}{(1-r)^N}.$$

The lefthand side of this expression is strictly increasing in c, with range $[0, \infty)$; the righthand side is strictly increasing in N, being equal to 1 when $N = 1$ and tending to infinity as $N \to \infty$. Hence, given $r\left(\frac{1}{2}, 1\right)$, there is a strictly decreasing function $f : \{1, 2, ...\} \to [0, \infty)$ such that $h\,(0) < 1$ iff $c < f\,(N)$.

(see Figure 5) so

$$1+c \;=\; \frac{1-(2+c)^{-1}}{(2+c)^{-1}} > \frac{1-\tilde{\pi}}{\tilde{\pi}} > \frac{1-g(0)}{g(0)} \;=\; (1+c)\left(\frac{1-r}{r}\right)^{N}$$

and therefore

$$\frac{1}{1+\left(\frac{r}{1-r}\right)(1+c)} < \frac{1}{1+\left(\frac{r}{1-r}\right)\left(\frac{1-\tilde{\pi}}{\tilde{\pi}}\right)} < \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N-1}}.$$

In each figure, the dashed green curve passes through the point

$$\left(\frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N-1}}, \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N+1}}\right)$$

since it is easily checked that when

$$\underline{p} \;=\; \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N-1}}$$

the righthand side of (12) is equal to

$$\frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N+1}}.$$

Consider Figure 6 (where $h(0)<1$). If

$$\underline{p} \;>\; \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N+1}}$$

then only the trivial equilibrium exists – there is no non-trivial equilibrium. If

$$\underline{p} \;\leq\; \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N+1}}$$

there is unique non-trivial equilibrium unless

$$\frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N+1}} \in \left\{\underline{p}, \overline{p}\right\}.$$

When

$$\overline{p} \;\leq\; \frac{1}{1+\left(\frac{r}{1-r}\right)(1+c)}$$

this unique non-trivial equilibrium is of variety C; when

$$\frac{1}{1+\left(\frac{r}{1-r}\right)(1+c)} \;<\; \overline{p} \quad \text{and} \quad \underline{p} \;<\; \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N-1}}$$

it is of variety FP; when

$$\frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N+1}} \;<\; \overline{p} \quad \text{and} \quad \underline{p} \;\geq\; \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N-1}}$$

it is of variety I; and when

$$\underline{p} \;<\; \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N+1}} \;<\; \overline{p}$$

it is of variety DFP.

If

$$\underline{p} \;=\; \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N+1}}$$

there is a continuum of equilibria: an I equilibrium plus all the DFP equilibria described in (7). Finally, if

$$\overline{p} \;=\; \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N+1}}$$

there is an I equilibrium plus all the DFP equilibria described in (8). At the point where

$$\underline{p} \;=\; \overline{p} \;=\; \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N+1}} \tag{13}$$

the set of non-trivial equilibria comprises the I profile together with all the DFP profiles.

Figure 7 illustrates the equilibrium mapping when $h(0) \geq 1$ and is interpreted similarly. In this case the I region expands and the DFP region contracts. We now observe a DFP equilibrium iff

$$\underline{p} \;=\; \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N+1}}$$

in which case the set of non-trivial equilibria comprises the I profile together with all the DFP profiles. For all other voting problems with disjoint posteriors and $h(0) \geq 1$ there is at most one non-trivial equilibrium.

Looking along the diagonal of either figure, we recover the familiar mapping from parameters to (non-trivial) equilibria for the no-ambiguity case. In many respects, the case of ambiguity with disjoint posteriors cleaves to this no-ambiguity benchmark. In particular, we immediately deduce the following from Figures 6-7:

Theorem 3.1 *Generically, voting problems with disjoint posteriors have at most one non-trivial equilibrium. Moreover, every equilibrium generates state-conditional conviction probabilities on the Pareto frontier.*

However, some noteworthy novelties do emerge from our analysis.

First and foremost, ambiguity promotes the DFP equilibrium to a generic possibility – DFP equilibria exist for an open set of voting problems (see Figure 6). In such an equilibrium, jurors who receive an innocent signal vote "not guilty" while jurors in receipt of a guilty signal randomise. This Dual FP behaviour is arguably more natural than FP behaviour: it is intuitive that jurors whose utility functions reflect Blackstone's maxim might be more resolute when following innocent signals than guilty ones. However, in the absence of ambiguity, pivotality logic militates strongly against this intuition. Guilty votes can only come from jurors with guilty signals in a DFP equilibrium. In such an equilibrium, jurors must therefore be *indifferent* about convicting or acquitting a defendant when *all* $N+1$ jurors have guilty signals. This knife-edge condition, which, it is crucial to observe, is not affected by the equilibrium value of $\sigma_2 \in (0,1)$, ensures that the DFP scenario is non-generic in the absence of ambiguity: it requires parameters to satisfy (13).

If ambiguity is present, hedging incentives may undo this logic. This happens when $\underline{\pi}_2 < \pi^*(\sigma) < \overline{\pi}_2$ so that posteriors following a guilty signal "disagree" about which is the better vote: there is robust indecision rather than knife-edge indifference. If $\sigma^*(\sigma) < 1$ there is a perfect hedge against this uncertainty and it involves randomisation. When $\sigma = (0,\sigma_2)$ is a DFP profile, $\pi^*(\sigma) = g(0)$ and $\sigma^*(\sigma) = h(0)/\sigma_2^N$. Therefore, if $\underline{\pi}_2 < g(0) < \overline{\pi}_2$ and $\sigma_2 > h(0)^{1/N}$ juror i will choose $\sigma_2^i = \sigma^*(\sigma) \in (0,1)$ in response to σ following a guilty signal. Provided $h(0) < 1$ we have a DFP equilibrium when $\sigma_2 = h(0)^{1/(N+1)}$. From Figure 3 we see that this equilibrium has lower Type I error, and higher Type II error, than any non-trivial equilibrium that may be generically observed in the absence of ambiguity.

This observation suggests a more general comparative static question. We have seen that for voting problems with disjoint posteriors, non-trivial equilibria are generically unique when they exist. How is this unique non-trivial equilibrium affected by adding a small amount of ambiguity? Provided the change is small enough to keep posteriors disjoint, we can use Figures 6-7 to answer this question.

By "adding ambiguity" we mean going from a voting problem with $\underline{p} = \overline{p} = z$ to one with $\underline{p} = z - \varepsilon$ and $\overline{p} = z + \varepsilon$ for some $\varepsilon \in (0, \min\{z, 1-z\})$ small enough to maintain disjoint posteriors, holding all other parameter values fixed.[14] Graphically, this is a northwestly movement from a point on the diagonal in Figure 6 or Figure 7. Starting from a point in the C region, such a movement either keeps the voting problem in this region or moves it into the FP region. If we start from the FP region, we remain there.

[14] The assumption that z falls exactly in the middle of the prior interval is not essential to our analysis.

Note, however, that if we start from

$$\underline{p} = \overline{p} = z \; < \; \frac{1}{1 + \left(\frac{r}{1-r}\right)\left(\frac{1-\tilde{\pi}}{\tilde{\pi}}\right)}$$

(which corresponds to $\underline{\pi}_1 = \overline{\pi}_1 < \tilde{\pi}$) then this movement **lowers** the equilibrium value of σ_1 (since it raises $\overline{\pi}_1$ and function g^{-1} is strictly decreasing), while if we start from

$$\underline{p} = \overline{p} = z \; > \; \frac{1}{1 + \left(\frac{r}{1-r}\right)\left(\frac{1-\tilde{\pi}}{\tilde{\pi}}\right)}$$

(which corresponds to $\underline{\pi}_1 = \overline{\pi}_1 > \tilde{\pi}$) then this movement **raises** the equilibrium value of σ_1 (since it lowers $\underline{\pi}_1$). If

$$\underline{p} = \overline{p} = z \; = \; \frac{1}{1 + \left(\frac{r}{1-r}\right)\left(\frac{1-\tilde{\pi}}{\tilde{\pi}}\right)}$$

the movement has no effect on the equilibrium. This is a curiously non-monotonic – and testable – comparative static property of the model.

Continuing along the diagonal, if we start from a point in the I region, we may either remain there or move into the FP region, or possibly even into the DFP region (if $h(0) < 0$ and z and ε are high enough). Finally, if we start at the point where the voting problem with ambiguity has a continuum of DFP equilibria, adding ambiguity will either move us into the interior of the DFP region (if $h(0) < 1$), or into the I region (if $h(0) \geq 1$).

These comparative statics are summarised in Figure 8. The black dot in the FP region corresponds to the profile $(\tilde{\sigma}_1, 1)$, while the black dot in the DFP region corresponds to the profile $\left(0, h(0)^{1/(N+1)}\right)$. Purple arrows indicate the possible directions of movement as a result of adding ambiguity (starting from a non-ambiguous voting problem) while maintaining disjoint posteriors.[15]

We can also combine this type of comparative static analysis with Proposition A.1 in Appendix A to determine when an *ex ante* optimal symmetric profile is an equilibrium; and if not, the direction of the equilibrium distortion away from optimality. Consider Figure 9, which exhibits the unique *ex ante* optimal symmetric profile for a particular voting problem with ambiguity (the black dot on the Pareto frontier). The optimal profile is from the FP category. Is it an equilibrium? We may answer this question as follows. First, the same profile (let's denote it by σ') is *ex ante* optimal for the non-ambiguous voting game obtained by increasing $\underline{p}$ until it matches $\overline{p}$ while keeping all other parameters the same as before. It follows by Proposition A.1 that σ' is also the unique non-trivial

[15] When reading this figure recall that if we start from a non-ambiguous voting problem with a DFP equilibrium, then **all** DFP profiles, together with I profile, are equilibria. Starting from such a voting problem, adding ambiguity will either move us to the DFP profile $\left(0, h(0)^{1/(N+1)}\right)$ or to the I profile.

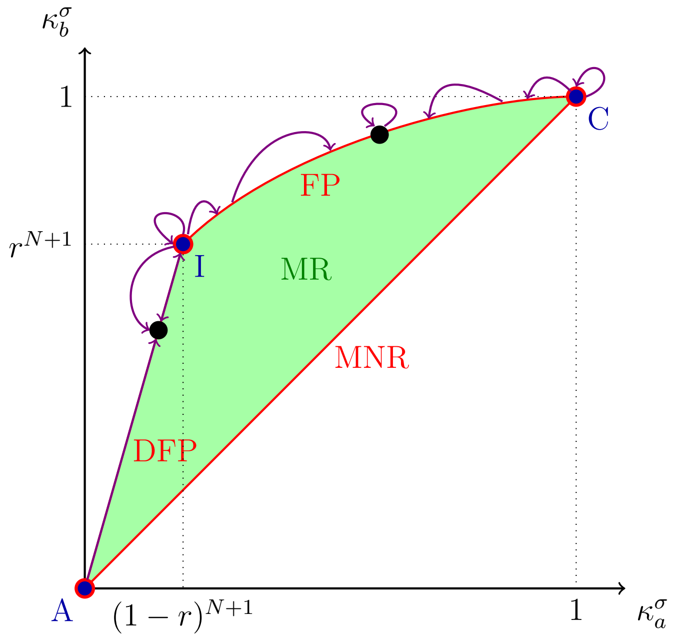

Figure 8: Effects of adding ambiguity on (generic) equilibrium decision quality.

equilibrium of this non-ambiguous voting problem. So, the question becomes, is σ' still an equilibrium for the original voting problem with ambiguity? Viewed in Figure 6 or 7 (as appropriate), the ambiguous problem lies due west of the non-ambiguous problem; and since the non-ambiguous problem is in the FP region, so is the ambiguous one. Therefore, whether σ' still is an equilibrium for the original voting problem with ambiguity depends on whether σ_1' is above or below $\tilde{\sigma}_1$; that is, whether $\overline{p}$ is below or above

$$\frac{1}{1+\left(\frac{r}{1-r}\right)\left(\frac{1-\tilde{\pi}}{\tilde{\pi}}\right)}.$$

Only if $\sigma_1' \geq \tilde{\sigma}_1$ will a westward movement preserve the equilibrium value of σ_1. If $\sigma_1' < \tilde{\sigma}_1$ the westward movement will increase the equilibrium value of σ_1 and therefore raise Type I error (lower Type II error) relative to the *ex ante* optimum.

In Figure 9, we have assumed that the *ex ante* optimal profile, σ', has state-conditional conviction probabilities $\left(\kappa_a^{\sigma'}, \kappa_b^{\sigma'}\right)$ located to the right of the kink in the indifference contour through that point. In principle, if c is close enough to zero, we might have an *ex ante* optimal profile, σ', that is in the FP category but $\left(\kappa_a^{\sigma'}, \kappa_b^{\sigma'}\right)$ sits to the **left** of the kink, so that σ' is *ex ante* optimal – and therefore an equilibrium – for the non-ambiguous voting problem obtained by reducing $\overline{p}$ until it matches $\underline{p}$ (rather than raising $\underline{p}$). In this case, moving from the non-ambiguous to the ambiguous problem is a **northward** (rather than westward) movement in Figure 6 or 7. It is easy to show that if $\left(\kappa_a^{\sigma'}, \kappa_b^{\sigma'}\right)$ sits to the left of

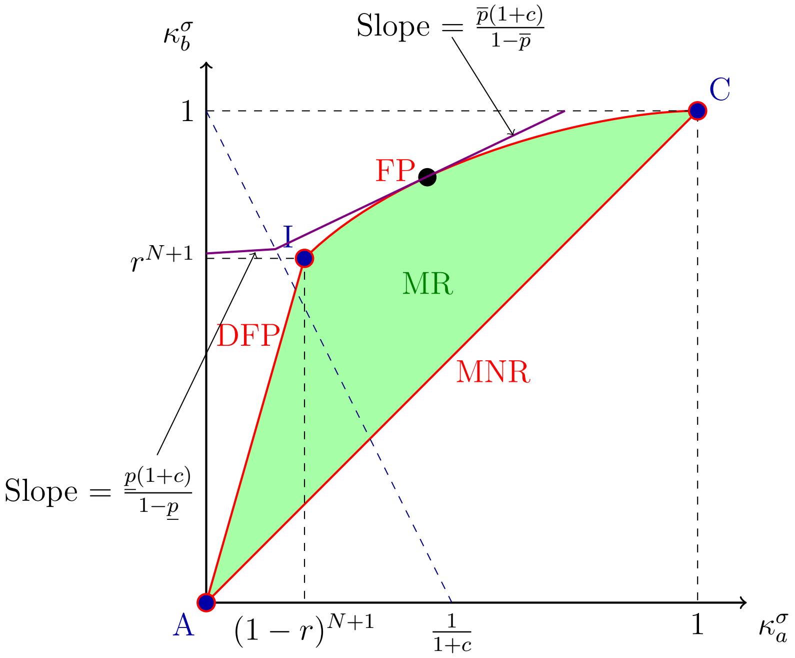

Figure 9: An *ex ante* optimal profile.

the kink (i.e., $\kappa_b^{\sigma'} < 1 - (1+c)\,\kappa_a^{\sigma'}$) then we must have $\sigma_1' < \tilde{\sigma}_1$.[16] Moving northward from such a point does not alter the equilibrium value of σ_1 so σ' continues to be an equilibrium. In other words, the *ex ante* optimal profile is necessarily an equilibrium in this case.

We can use the same sort of analysis to evaluate any point on the Pareto frontier: to assess whether such a point is an equilibrium when it is *ex ante* optimal, and if not, whether equilibrium constraints raise or lower Type I error relative to the *ex ante* optimum. As

[16] Suppose $\sigma' = (\sigma_1', 1)$ with $0 < \sigma_1' < 1$. Therefore:

$$\begin{aligned} & \kappa_b^{\sigma'} < 1 - (1+c)\,\kappa_a^{\sigma'} \\ \Leftrightarrow\quad & [(1-r)\,\sigma_1' + r]^{N+1} + (1+c)\,[r\sigma_1' + (1-r)]^{N+1} < 1 \\ \Rightarrow\quad & \sigma_1'\,[(1-r)\,\sigma_1' + r]^{N} + (1+c)\,\sigma_1'\,[r\sigma_1' + (1-r)]^{N} < 1 \\ \Leftrightarrow\quad & h(\sigma_1') > \sigma_1' \\ \Leftrightarrow\quad & \sigma_1' < \tilde{\sigma}_1 \end{aligned}$$

where the third line uses the fact that

$$\sigma_1' < \min\{\,(1-r)\,\sigma_1' + r,\ r\sigma_1' + (1-r)\}\,.$$

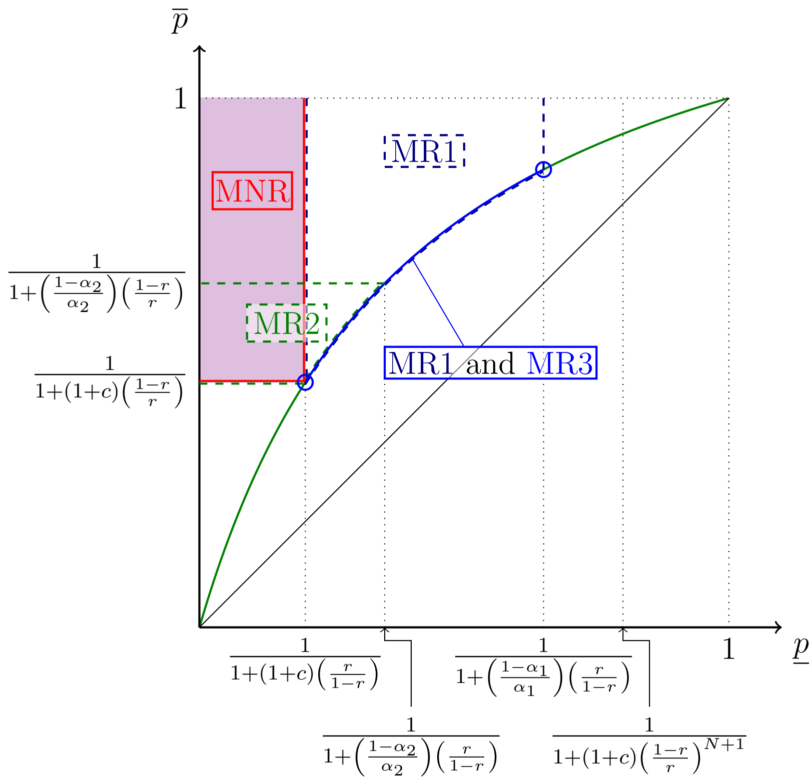

Figure 10: Strictly mixed equilibria with overlapping posteriors.

our example above suggests, a complex set of possibilities arises. Given the contentious, if conventional, measure of *ex ante* welfare, we do not attempt to summarise all the possibilities here. Rather, we move on to consideration of overlapping posteriors, where an entirely uncontroversial failure of decision quality can be shown to be possible – a situation in which **all** non-trivial equilibria generate state-conditional conviction probabilities **strictly inside** the Pareto frontier.

3.2 Overlapping posteriors

Only if posteriors overlap ($\underline{\pi}_1 \leq \overline{\pi}_2$) is it possible to observe equilibria of the strictly mixed varieties: MNR or MR. This possibility also creates the potential for a dramatic failure of decision quality to emerge (see Proposition 3.8).

Figure 10 identifies the parameter constellations where MNR and MR equilibria exist

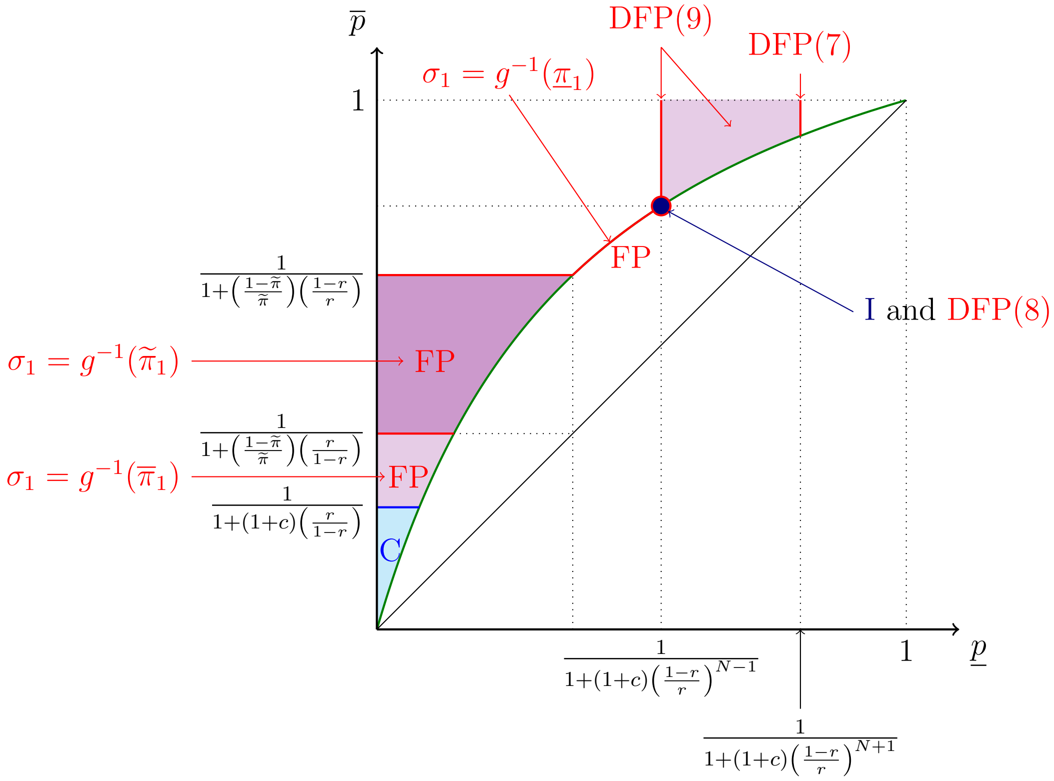

Figure 11: Other (not strictly mixed) equilibria with overlapping posteriors and $h(0) < 1$.

(cf., Figure 3 in Ryan, 2021).[17] Once again, the green curve is where $\underline{\pi}_1 = \overline{\pi}_2$ so posteriors overlap on or above this curve. The (unique) MNR equilibrium exists within the pink rectangular region, which includes all boundary points except those on the northern and western boundaries, which are outside the domain of the parameter space. An MR equilibrium of sub-type MR1 exists strictly between the vertical dashed purple lines; an MR2

[17]When reading this diagram, recall that

$$\frac{1}{2+c} < \alpha_2 \leq \alpha_2 \leq g\left(0\right),$$

which implies

$$\frac{1}{1+\left(\frac{r}{1-r}\right)(1+c)} < \frac{1}{1+\left(\frac{r}{1-r}\right)\left(\frac{1-\alpha_2}{\alpha_2}\right)} \leq \frac{1}{1+\left(\frac{r}{1-r}\right)\left(\frac{1-\alpha_1}{\alpha_1}\right)} \leq \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N-1}}.$$

Figure 10 illustrates the case where both weak inequalities hold strictly, but the picture is qualitatively the same even if neither does.

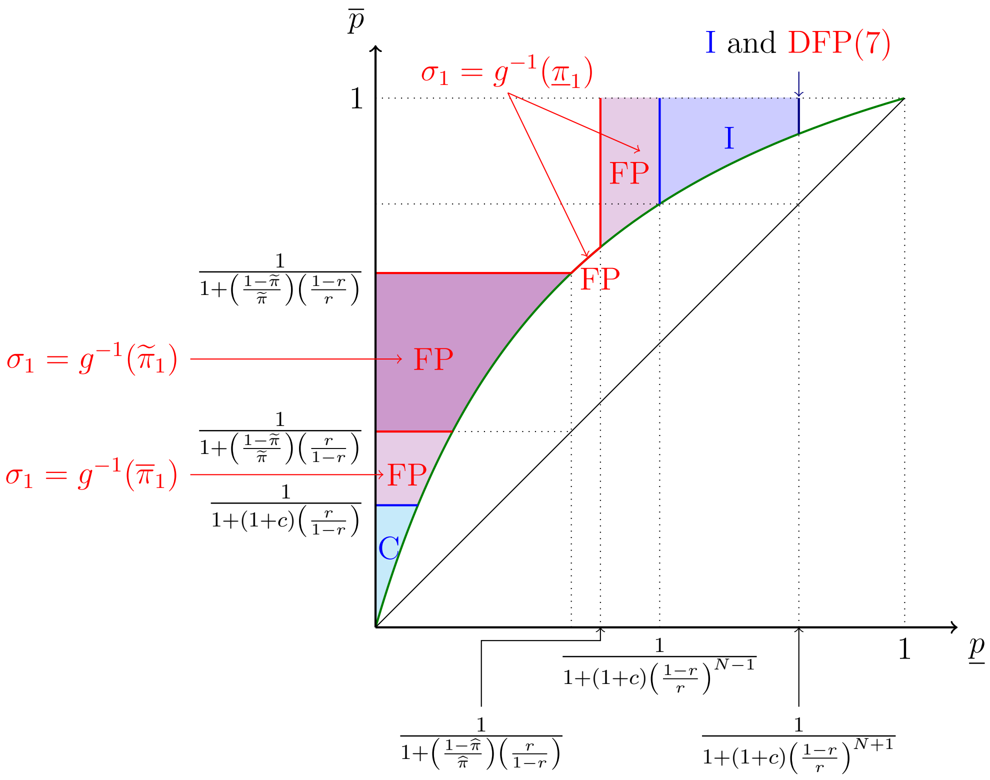

Figure 12: Other (not strictly mixed) equilibria with overlapping posteriors and $h(0) \geq 1$.

sub-type exists strictly between the horizontal dashed green lines;[18] and the continuum of MR3 sub-type equilibria along the indicated portion of curve where $\underline{\pi}_1 = \overline{\pi}_2$.

A key implication of Figure 10 is the existence of an open set of voting problems which possess both an MNR and MR2 equilibrium. With disjoint posteriors, multiplicity of non-trivial equilibria is more than a non-generic possibility.

In the interests of readability, rather than encumber Figure 10 with additional shading to indicate regions where C, I, FP or DFP equilibria exist, we provide the latter information in a separate figure – or rather, two separate figures: one for voting problems with $h(0) < 1$ and another for voting problems with $h(0) \geq 1$. These are Figures 11 and 12. When reading Figure 12, recall that $\tilde{\pi} < \hat{\pi} \leq g(0)$. Figure 12 is drawn assuming $\hat{\pi} < g(0)$. If $\hat{\pi} = g(0)$ then

$$\frac{1}{1 + \left(\frac{1-r}{r}\right)\left(\frac{1-\hat{\pi}}{\hat{\pi}}\right)} = \frac{1}{1 + (1+c)\left(\frac{1-r}{r}\right)^{N-1}}$$

so the rightmost shaded FP region vanishes: FP equilibria with $\sigma_1 = g^{-1}(\underline{\pi}_1) \neq g^{-1}(\tilde{\pi})$

[18]Recall that the MR1 and MR2 equilibria are also unique when they exist.

become non-generic, as in the $h(0) < 1$ case.

To obtain the complete mapping from the parameter space to the set of non-trivial equilibria, the reader should imagine Figure 10 superimposed on each of these figures. When imagining this superimposition, note that since

$$\frac{1}{2+c} < \tilde{\pi}$$

we have

$$\frac{1}{1+\left(\frac{1-r}{r}\right)(1+c)} < \frac{1}{1+\left(\frac{1-r}{r}\right)\left(\frac{1-\tilde{\pi}}{\tilde{\pi}}\right)}$$

which implies that the FP region strictly overlaps with the MNR and MR2 regions (irrespective of whether $h(0) < 1$ or $h(0) \geq 1$). This exacerbates the multiplicity issue noted above. We now see that there is an open set of voting problems with **three** non-trivial equilibria: an FP equilibrium with $\sigma_1 = g^{-1}(\tilde{\pi})$, the unique MNR equilibrium and the unique MR2 equilibrium. There also exists an open set of voting problems possessing an FP equilibrium with $\sigma_1 = g^{-1}(\tilde{\pi})$ together with the unique MR1 equilibrium.

On the other hand, we have not established anything about the value of $\tilde{\pi}$ in relation to α_1 or α_2. This should be borne in mind when reading the superimposed figures.

The most striking lesson from Figures 10-12 is the existence of an open set of voting problems for which **all** non-trivial equilibria are in the MR or MNR categories. Given any non-trivial equilibrium of any such voting problem, a Social Planner could find a symmetric strategy profile which strictly lowers the probability of both Type I and Type II error (see Figure 3). This occurs provided $\overline{p}$ is high enough and $\underline{p}$ low enough; that is, when prior ambiguity is sufficiently large (where "sufficiency" depends on the values of N, r and c).

Proposition 3.8 *The following are equivalent:*

(i) The following conditions are satisfied: $\underline{\pi}_1 < \overline{\pi}_2$,

$$\underline{p} < \min\left\{\frac{1}{1+\left(\frac{1-\alpha_1}{\alpha_1}\right)\left(\frac{r}{1-r}\right)}, \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N-1}}\right\}$$

and

$$\overline{p} > \max\left\{\frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)}, \frac{1}{1+\left(\frac{1-\overline{\pi}}{\overline{\pi}}\right)\left(\frac{1-r}{r}\right)}\right\}.$$

(ii) The voting problem has a non-trivial equilibrium and all non-trivial equilibria generate state-contingent conviction probabilities strictly inside the Pareto frontier of Figure 3.

Of course, the trivial equilibrium always exists and the reader may wonder whether this equilibrium is *ex ante* optimal in precisely the scenarios in which all non-trivial equilibria are strictly mixed. From Figure 4 it is clear that profile A is *ex ante* optimal if and only if it would also be *ex ante* optimal were the ambiguous prior to be replaced by a non-ambiguous prior equal to $\underline{p}$ (i.e., when $\overline{p}$ is reduced until it coincides with $\underline{p}$). Moreover, if profile A is *ex ante* optimal for this non-ambiguous voting problem, then it is also the **unique** equilibrium: see Proposition A.1 in Appendix A. Thus, if the non-ambiguous voting problem with prior equal to $\underline{p}$ has a non-trivial equilibrium, we may conclude that profile A is **not** *ex ante* optimal in the original (ambiguous) voting problem.

In fact, it is easy to see that the non-ambiguous voting problem does indeed have a non-trivial equilibrium whenever the original ambiguous problem has all non-trivial equilibria strictly mixed. Starting from any point in Figure 10 where a strictly mixed equilibrium exists, reducing the value of $\overline{p}$ until it coincides with $\underline{p}$ shifts that point due south to the diagonal. From Figures 11-12 we observe that the corresponding non-ambiguous voting problem will possess a non-trivial equilibrium provided

$$\underline{p} \ \leq \ \frac{1}{1+(1+c)\left(\frac{1-r}{r}\right)^{N+1}}$$

which is necessarily the case if the original ambiguous voting problem has a strictly mixed equilibrium.

4 Concluding remarks

In the absence of ambiguity: (i) all equilibria generate state-conditional conviction probabilities on the Pareto frontier; (ii) the *ex ante* optimal profile is an equilibrium; and (iii) any non-trivial equilibrium is *ex ante* optimal amongst symmetric profiles. None of these statements remains true once voting problems with ambiguity are admitted.

Ellis (2016) already showed that ambiguous prior beliefs can distort voting behavior under the majority rule, and may produce exotic equilibria with poor decision quality when posterior intervals overlap. The present paper reaches similar conclusions for voting under the unanimity rule. Our analysis has a more specialised signal structure than Ellis' (2016), but employs the more general utility specification of Feddersen and Pesendorfer (1998).

Because Ellis studies voting under the majority rule, and is concerned with information aggregation in the sense of Condorcet, he evaluates equilibria against the "correct expected winners" criterion: a voter is more likely to vote for the correct state than the incorrect one, conditional on either state. He identifies a condition under which there exists an equilibrium that violates this property – the "coin toss" equilibrium of the Introduction (*ibid.*, Proposition 1); and also shows that this condition precludes **any** equilibrium from having correct expected winners (*ibid.*, Theorem 1). The required condition is that voters "lack confidence": the interior of $\Pi_1 \cap \Pi_2$ contains $\frac{1}{2}$. Because of potential asymmetry in the Bernoulli utility function, our decision quality criterion is Pareto optimality with respect to Type I and Type II error probabilities. We observed in the previous section that when posteriors overlap, the unanimity rule can generate a Pareto sub-optimal for any non-trivial equilibrium. This is the natural analogue of Ellis' (2016) result. Moreover, our results, summarised in Figures 10-12, allow us to precisely determine those voting problems for which this pathology arises.

Importantly, our results differ sharply from those of Fabrizi *et al.* (2022). When voting takes place under the unanimity rule, ambiguity about p has very different implications to ambiguity about r. Adding ambiguity about r has no impact on the range of equilibria that one may observe, and typically lowers the probability of Type I error. Relative to the no-ambiguity benchmark, the impact of ambiguity is negligible from a descriptive standpoint and mildly positive form a normative standpoint. When ambiguity affects p, matters are very different, and this is so whether voting follows the majority rule or the unanimity rule. To the best of our knowledge, the impact of ambiguity about r on the majority voting game remains an open question.

Our analysis provides a complete characterisation of all symmetric equilibria. Such completeness is important for experimental testing, where precise model predictions are usually needed. This allows us to identify regions of the parameter space where non-trivial equilibria are unique, so that model comparative statics may be identified and, in principle, tested.[19] Our complete characterisation also allows us to evaluate whether the *ex ante* optimum can be achieved, and if not, the direction of the distortion.

References

Ali, S. N., Goeree, J. K., Kartik, N., and Palfrey, T. R. (2008). Information aggregation in standing and ad hoc committees. *American Economic Review*, 98(2), 181-86.

Anderson, L. R., Holt, C. A., Sieberg, K. K., and Oldham, A. L. (2015). An experimental study of jury voting behavior. In *The Political Economy of Governance* (pp. 157-178). Springer, Cham.

Austen-Smith, D., and Banks, J. S. (1996). Information aggregation, rationality, and the Condorcet jury theorem. *American Political Science Review*, 90(1), 34-45.

Coughlan, P. J. (2000). In defense of unanimous jury verdicts: Mistrials, communication, and strategic voting. *American Political Science Review*, 94(2), 375-393.

Ellis, A. (2016). Condorcet meets Ellsberg. *Theoretical Economics,* 11(3), 865-895.

[19]We are currently engaged in experimental work of exactly this sort.

Fabrizi, S., Lippert, S., Pan, A., and Ryan, M. (2022). A theory of unanimous jury voting with an ambiguous likelihood. *Theory and Decision*, 93, 399–425.

Fagin, R. and Halpern, J. (1990). A new approach to updating beliefs. *Proceedings of the 6th Conference on Uncertainty in AI* (pp.317-325). Elsevier, New York.

Feddersen, T., and Pesendorfer, W. (1998). Convicting the innocent: The inferiority of unanimous jury verdicts under strategic voting. *American Political Science Review*, 92(1), 23-35.

Gerardi, D., and Yariv, L. (2007). Deliberative voting. *Journal of Economic Theory*, 134(1), 317-338.

Goeree, J. K., and Yariv, L. (2011). An experimental study of collective deliberation. *Econometrica*, 79(3), 893-921.

Guarnaschelli, S., McKelvey, R. D., and Palfrey, T. R. (2000). An experimental study of jury decision rules. *American Political Science Review*, 94(2), 407-423.

Jaffray, J.-Y. (1992). Bayesian updating and belief functions. *IEEE Transactions on Systems, Man and Cybernetics*, 22, 1144-1152.

McLennan, A. (1998). Consequences of the Condorcet jury theorem for beneficial information aggregation by rational agents. *American Political Science Review*, 92(2), 413-418.

Pan, A. (2019). A note on pivotality. *Games*, 10(2), 24.

Ryan, M.J. (2021). Feddersen and Pesendorfer meet Ellsberg. *Theory and Decision*, 90(3), 543-577.

Siniscalchi, M. (2011). Dynamic choice under ambiguity. *Theoretical Economics*, 6, 379–421.

Young, H. P. (1988). Condorcet's theory of voting. *American Political Science Review*, 82(4), 1231-1244.

APPENDICES

A Characterising decision quality

In this appendix we describe the mapping $\sigma \mapsto (\kappa_a^\sigma, \kappa_b^\sigma)$ from symmetric profiles to state-conditional conviction probabilities (as depicted in Figure 3), and then prove an important result on the *ex ante* optimal symmetric profile(s) for voting problems without ambiguity.

Let $\Gamma \equiv \left\{ (z_1, z_2) \in [0,1]^2 \middle| \, z_1 \leq z_2 \right\}$ and consider the mapping $\sigma \mapsto (x(\sigma), y(\sigma))$ on this domain, where

$$x(\sigma) = r\sigma_1 + (1-r)\sigma_2 = (\kappa_a^\sigma)^{1/(N+1)}$$

and

$$y(\sigma) = (1-r)\sigma_1 + r\sigma_2 = (\kappa_b^\sigma)^{1/(N+1)}.$$

We describe this mapping as an intermediate step to characterising $\sigma \mapsto (\kappa_a^\sigma, \kappa_b^\sigma)$. It is evident that $(x(\sigma), y(\sigma)) \in \Gamma$ for any $\sigma \in \Gamma$, and that the mapping $\sigma \mapsto (x(\sigma), y(\sigma))$ is continuous and satisfies the following mixture-linearity property:

$$\lambda (x(\sigma), y(\sigma)) + (1-\lambda)(x(\sigma'), y(\sigma')) = (x(\lambda\sigma + (1-\lambda)\sigma'), y(\lambda\sigma + (1-\lambda)\sigma'))$$

for any $\lambda \in [0,1]$ and any $\sigma, \sigma' \in \Gamma$. The image of Γ is therefore the convex hull of the image of its extreme points (i.e., profiles A, I and C in Figure 2). Hence, the image of Γ is the convex hull of the points:

$$\{(0,0), \ (1-r, r), \ (1,1)\}.$$

This image is the subset of Γ consisting of the points on or below the graph of the piecewise linear function $f : [0,1] \to [0,1]$ defined as follows:

$$f(x) = \begin{cases} \left(\frac{r}{1-r}\right) x & \text{if } 0 \leq x \leq 1-r \\ r + \left(\frac{1-r}{r}\right)(x - (1-r)) & \text{if } 1-r \leq x \leq 1 \end{cases}$$

Now consider the mapping $\sigma \mapsto \left(x(\sigma)^{N+1}, y(\sigma)^{N+1}\right) = (\kappa_a^\sigma, \kappa_b^\sigma)$. It is easily verified that the image of Γ under this mapping is the set[20]

$$\left\{ \left(\lambda x^{N+1}, \lambda y^{N+1}\right) \,\middle|\, \lambda \in [0,1] \text{ and } (x,y) \text{ is on the graph of } f \right\} \qquad (14)$$

[20]Note that (x,y) is in the image of Γ under the mapping $\sigma \mapsto (x(\sigma), y(\sigma))$ iff $(x,y) = (\mu\overline{x}, \mu\overline{y})$ for some $\mu \in [0,1]$ and some $(\overline{x}, \overline{y})$ in the graph of f. Thus $(\kappa_a^\sigma, \kappa_b^\sigma) = \left(x^{N+1}, y^{N+1}\right)$ for some (x,y) in the image of Γ under the mapping $\sigma \mapsto (x(\sigma), y(\sigma))$ iff $(\kappa_a^\sigma, \kappa_b^\sigma) = \left(\mu^{N+1}\overline{x}^{N+1}, \mu^{N+1}\overline{y}^{N+1}\right)$ for some $\mu \in [0,1]$ and some $(\overline{x}, \overline{y})$ in the graph of f. Setting $\lambda = \mu^{N+1} \in [0,1]$ gives the result.

It follows that the image of Γ under the mapping $\sigma \mapsto (\kappa_a^\sigma, \kappa_b^\sigma)$ is the subset of Γ consisting of the points on or below the graph of function $g : [0,1] \to [0,1]$ defined by

$$g(z) = f\left(z^{1/(N+1)}\right)^{N+1} = \begin{cases} \left(\frac{r}{1-r}\right)^{N+1} z & \text{if } 0 \le z \le (1-r)^{N+1} \\ \left[r + \left(\frac{1-r}{r}\right)\left(z^{1/(N+1)} - (1-r)\right)\right]^{N+1} & \text{if } (1-r)^{N+1} \le z \le 1 \end{cases}$$

This is the region depicted in Figure 3. It is easily checked that $g'(z) > 0$ and $g''(z) < 0$ when $(1-r)^{N+1} < z < 1$, giving the strictly concave portion of the Pareto frontier. It is also straightforward to show that

$$\lim_{z \to (1-r)^{N+1}} g'(z) = \left(\frac{r}{1-r}\right)^{N-1} < \left(\frac{r}{1-r}\right)^{N+1}$$

which implies a kink in the Pareto frontier at the point $\left((1-r)^{N+1}, r^{N+1}\right)$ and, more importantly, convexity of the image of Γ under the mapping $\sigma \mapsto (\kappa_a^\sigma, \kappa_b^\sigma)$: see Figure 3.

When $\underline{p} = \overline{p} = p$ (no ambiguity), the *ex ante* expected utility of a typical juror from symmetric profile σ is

$$p\left[(1-\kappa_a^\sigma) - c\kappa_a^\sigma\right] + (1-p)\kappa_b^\sigma \tag{15}$$

The profile σ is *ex ante* optimal if $(\kappa_a^\sigma, \kappa_b^\sigma)$ maximises this linear function over the convex set in Figure 3. Note that if an FP profile is *ex ante* optimal (amongst symmetric profiles) it is the unique *ex ante* optimal profile; whereas a DFP profile is *ex ante* optimal iff the set of *ex ante* optimal profiles consists of all profiles in the union of categories A, DFP, and I. The condition for *ex ante* optimality of a DFP profile is easy to determine. Since (15) has linear contours with slope

$$\frac{p(1+c)}{1-p}$$

a DFP profile is *ex ante* optimal iff

$$\frac{p(1+c)}{1-p} = \left(\frac{r}{1-r}\right)^{N+1}$$

which is equivalent to

$$(1-p)\,r^{N+1} - p(1-r)^{N+1} c = p(1-r)^{N+1} \tag{16}$$

Condition (16) says that jurors are indifferent between conviction and acquittal when all $N+1$ signals are (known to be) of the guilty type. This is precisely the condition under which a DFP equilibrium exists (in the absence of ambiguity). So if a DFP profile is *ex ante* optimal, it is an equilibrium. Indeed, a more general result is easily shown using McLennan (1998, Theorem 2):

Proposition A.1 *If* $\underline{p} = \overline{p}$ *then any non-trivial equilibrium is* ex ante *optimal amongst symmetric profiles, and if no non-trivial equilibrium exists then the trivial equilibrium is* ex ante *optimal amongst symmetric profiles.*

Proof. In the absence of ambiguity, it is irrelevant whether players choose mixed strategies or behaviour strategies. McLennan (1998, Theorem 2) shows that the *ex ante* optimal symmetric profile must be an equilibrium. In any non-trivial equilibrium a player can achieve the same *ex ante* expected payoff as in the trivial equilibrium by deviating to the pure strategy that selects the "not guilty" vote irrespective of the signal. Hence, the equilibrium payoff from the non-trivial equilibrium must be at least as high as the trivial equilibrium payoff. It follows that if the voting problem has a non-trivial equilibrium, there exists a non-trivial equilibrium that is *ex ante* optimal amongst symmetric profiles, and if no non-trivial equilibrium exists then the trivial equilibrium must be *ex ante* optimal. Moreoever, the analysis in Feddersen and Pesendorfer (1998) shows that all non-trivial equilibria generate the same *ex ante* expected payoff in the absence of ambiguity, so if a non-trivial equilibrium exists then any such equilibrium is *ex ante* optimal amongst symmetric profiles. □

B Proofs for Section 3

The proofs in this section all follow the same basic logic. First, given a candidate equilibrium profile, σ, we determine $\pi^*(\sigma)$ and $\sigma^*(\sigma)$. We then use Figure 1 to determine the conditions on Π_1 and Π_2 under which σ satisfies the equilibrium conditions.

Proof of Proposition 3.4. If $\sigma = (0,1)$ then $\pi^*(\sigma) = g(0)$ and $\sigma^*(\sigma) = h(0)$. If $h(0) < 1$ then Figure 1 implies that σ is an equilibrium iff $\overline{\pi}_2 \leq g(0) \leq \underline{\pi}_1$. If $h(0) \geq 1$ then σ is an equilibrium iff $\underline{\pi}_2 \leq g(0) \leq \underline{\pi}_1$. □

Proof of Proposition 3.5. If $\sigma = (\sigma_1, 1)$ then $\pi^*(\sigma) = g(\sigma_1)$ and $\sigma^*(\sigma) = h(\sigma_1)$. Suppose $0 < \sigma_1 < 1$ and σ is an equilibrium. Then, by inspection of Figure 1, one of the following must hold: either (i) $\underline{\pi}_1 \leq g(\sigma_1) \leq \overline{\pi}_1$ and $\sigma_1 = h(\sigma_1)$; or (ii) $\underline{\pi}_1 = g(\sigma_1)$ and $\sigma_1 < h(\sigma_1)$; or (iii) $\overline{\pi}_1 = g(\sigma_1)$ and $\sigma_1 > h(\sigma_1)$. From Figure 5 we see that

$$g(\sigma_1) \gtreqless \tilde{\pi} \quad \text{as} \quad h(\sigma_1) \gtreqless \sigma_1$$

and $\sigma_1 = h(\sigma_1)$ iff $\sigma_1 = \tilde{\sigma}_1$. The result now follows. □

Proof of Proposition 3.6. If $\sigma = (\sigma_1, 1)$ then $\pi^*(\sigma) = g(\sigma_1)$ and $\sigma^*(\sigma) = h(\sigma_1)$.

Recall that

$$\hat{\sigma}_1 = \begin{cases} h^{-1}(1) & \text{if } h(0) \geq 1 \\ 0 & \text{otherwise} \end{cases}$$

and $\hat{\pi} = g(\hat{\sigma}_1)$. Note that $\hat{\sigma}_1 < 1$ and also:

$$\hat{\sigma}_1 > 0 \text{ iff } h(0) > 1.$$

Hence $\sigma^*(\sigma) \geq 1$ for some $\sigma_1 \in (0,1)$ iff $\hat{\sigma}_1 > 0$. From Figure 1 we immediately deduce:

(a) *If $\hat{\sigma}_1 > 0$ then $\sigma = (\sigma_1, 1)$ is an equilibrium with $\sigma_1 \in (0, \hat{\sigma}_1]$ iff $\underline{\pi}_1 = g(\sigma_1)$.*

Since $(0, \hat{\sigma}_1] = \emptyset$ when $\hat{\sigma}_1 = 0$, this gives case (a) in Proposition 3.6.

The remaining cases cover equilibria in which $\sigma_1 > \hat{\sigma}_1$. By inspection of Figure 1 we distinguish the following three exhaustive cases:

(b) *If $\sigma_1 \in (\hat{\sigma}_1, 1)$ satisfies $\underline{\pi}_1 = g(\sigma_1) \geq \overline{\pi}_2$ and $\sigma_1 \leq h(\sigma_1)$ then $\sigma = (\sigma_1, 1)$ is an equilibrium.*

(c) *If $\underline{\pi}_1 < g(\tilde{\sigma}_1) < \overline{\pi}_1$ and $\overline{\pi}_2 \leq g(\tilde{\sigma}_1)$ then $\sigma = (\tilde{\sigma}_1, 1)$ is an equilibrium.*

(d) *If $\sigma_1 \in (\hat{\sigma}_1, 1)$ satisfies $\overline{\pi}_1 = g(\sigma_1)$ and $\sigma_1 \geq h(\sigma_1)$ then $\sigma = (\sigma_1, 1)$ is an equilibrium.*

These are readily mapped to the corresponding cases in Proposition 3.6, which also makes it apparent that they are mutually exclusive. To make this mapping, recall from Figure 5 that

$$h(\sigma_1) \gtreqless \sigma_1 \quad \text{as} \quad \sigma_1 \lesseqgtr \tilde{\sigma}_1$$

and $g(\tilde{\sigma}_1) = \tilde{\pi}$ (and also recall that $\hat{\pi} = g(0)$ by assumption). □

Proof of Proposition 3.7. If $\sigma = (0, \sigma_2)$ then $\pi^*(\sigma) = g(0)$ and $\sigma^*(\sigma) = h(0)/\sigma_2^N$. Hence, $\sigma^*(\sigma) \geq 1$ iff $\sigma_2 \leq h(0)^{1/N}$, and

$$\sigma_2 \gtreqless \sigma^*(\sigma) \quad \text{as} \quad \sigma_2 \gtreqless h(0)^{\frac{1}{N+1}}.$$

If $h(0) \geq 1$ then $\sigma^*(\sigma) \geq 1$ for all $\sigma_2 \in (0,1)$. From Figure 1 it follows that if $h(0) \geq 1$ and $\sigma_2 \in (0,1)$, then $\sigma = (0, \sigma_2)$ is an equilibrium iff $\underline{\pi}_2 = g(0)$.

Suppose $h(0) < 1$. By inspection of Figure 1, there are three possibilities:

(i) If $\sigma_2 \in \left(0, h(0)^{1/(N+1)}\right]$ and $\underline{\pi}_2 = g(0)$ then $\sigma = (0, \sigma_2)$ is an equilibrium.

(ii) If $\underline{\pi}_2 < g(0) < \overline{\pi}_2$ and $\underline{\pi}_1 \geq g(0)$ then $\sigma = \left(0, h(0)^{1/(N+1)}\right)$ is an equilibrium.

(iii) If $\sigma_2 \in \left[h(0)^{1/(N+1)}, 1\right)$ and $\overline{\pi}_2 = g(0) \leq \underline{\pi}_1$ then $\sigma = (0, \sigma_2)$ is an equilibrium.

Hence, there exists $\sigma_2 \in (0,1)$ such that $\sigma = (0, \sigma_2)$ is an equilibrium iff condition (a) or (b) of Proposition 3.7 holds. The remaining claims follow directly; in particular, (7) follows by noting that $h(0)^{1/(N+1)} \geq 1$ when $h(0) \geq 1$ and that $\sigma_2 \in (0,1)$ in any DFP profile. □