Who Should Vote First on a Small Heterogeneous Sequential Jury?

Steve Alpern Bo Chen*
Warwick Business School, University of Warwick
Coventry, CV4 7AL, United Kingdom

May 25, 2014

Abstract

We consider a heterogeneous jury which must decide on a verdict of A or B (such as guilty or innocent) by a sequential majority vote. Jurors have private information in the form of integer signals, where positive signals indicate A is more likely and negative signals that B is more likely. A juror's strategy is a threshold (depending on previous voting, if any), where a juror votes A if his signal is higher than his threshold. Each juror's signal distribution is linear, with slope called his "ability", so that higher ability jurors are more likely to guess correctly between A and B. Using integer programming methods we show that the probability that a three-person jury comes to a correct verdict is maximized when the middle-ability juror votes first. In general, optimizing jurors must vote strategically, but when A and B are equiprobable and the abilities (b, c, a) in the voting order satisfy $a < b < c$, all jurors can vote naively for the alternative that has higher probability at time of voting. Our results have implications for larger juries and for optimizing line calls in sports such as tennis and badminton.
Keywords: jury — sequential voting — voting order — group decision

1 Introduction

A group of agents (jurors) must decide between two states of Nature, A and B (such as guilty or innocent). Their verdict is by majority rule in a sequential open vote, also known as "roll call" voting, and their common

*Corresponding author: b.chen@warwick.ac.uk; +44 24 76524755

aim is to maximize the probability that their verdict is correct. As the jurors have different "abilities" (or expertise, judgement, eyesight, etc.) to determine the correct state, the voting order might make a difference.

The main interest of this paper is to determine an optimal voting order, in terms of the set of abilities of the jurors. For reasons of a combinatorial nature, our techniques (integer programming) are restricted mainly to small juries. We find that, with three jurors of differing expertise, it is best to have the one of middle-level expertise vote first, and the order of the last two does not matter. On the other hand, when the alternatives are equally likely, and the voting sequence is middle ability followed by highest ability and finally lowest ability, all the jurors can afford to vote naively for the alternative which they deem most likely at the time of voting. These results apply as well to the last three jurors to vote in a larger jury.

In our model the *ability* of a juror is specified by the nature of his private information, in terms of the distribution of his *signal* within a set of ten odd values from -9 to $+9$, with higher values indicating A and lower ones B. We assume a linear distribution where the slope is proportional to a juror's ability a. Thus ability $a = 0$ is a horizontal line where every signal in the whole range comes equiprobably, and has no information content, while maximum ability $a = 4$ is most likely to give signals with high absolute values, with the most useful information. A strategy profile for the jury specifies how large a signal (*his threshold*) each juror requires in order to vote A. Individual thresholds depend on the prior voting sequence, if any. Thresholds are jointly optimized using integer programming or dynamic programming, as a function of the ability sequence of the jurors and the *a priori* probability of alternative A.

In our model of heterogeneous juries, jurors cannot optimize their decisions by simply assuming that they are pivotal. They must observe previous voting because it matters which jurors in the voting sequence have voted for A, not simply that half the other jurors have so voted.

The literature on using voting to amalgamate private information goes back at least as far as Condorcet [1], with his analysis of simultaneous voting. Surveys of various voting methods can be found in [2, 3]. Sequential voting has received less attention. Sequential voting is considered in [4, 5], but our problems of voting order and voting for the truth are not considered. The voting order of heterogenous experts is analyzed in [6, 7], however these experts care in part about their own reputations, rather than the correctness of the group decision, so the optimization problem (actually an equilibrium problem in their model) differs from ours. Unanimity rules are compared with majority rules in theoretical and experimental settings in [8, 9]. The result in [8] on increased wrong conviction rate when unanimous voting is required (compared to majority), is found in our model only when the jurors have the highest ability.

A recently published treatise on collective decision making [10] presents an extensive analysis of juries, discussing juror "quality" (including low quality due to drunkenness) and voting order. For the latter, several examples of historical anti-seniority rules are explicitly laid out. The historical importance of voting order is illustrated as far back as a Roman trial described by Cicero, where the defendant was given choice of open sequential voting or secret ballot. With his choice of the former, the order of voting proved seriously detrimental to his case: he was convicted by a majority of two.

The question of optimal ordering of heterogeneous experts or jurors is implicit in many rules of behavior. In tennis or badminton, the A/B call of In or Out is first made by a linesman and then can be overcalled by the umpire. A discussion of the optimal order of experts is given in [9], citing examples where courts follow either anti-seniority (increasing "ability" in our terminology) or seniority orders; respectively the ancient Sanhedrin and the contemporary American Supreme Court. Our model is simpler than other models in the literature in that our agents care only about the common good (getting the right verdict), rather than their own reputations, getting their preferred verdict, or minimizing their costs in obtaining quality signals. Thus we are concerned with simple optimization, rather than equilibria, and our analysis is not game theoretic.

We add a final note regarding our restriction to small jury size. Our methods of discrete optimization use integer programming to determine optimal threshold profiles for every fixed voting order. Since thresholds depend on previous voting, and the number of previous voting profiles potentially seen by the kth juror alone is 2^{k-1}, the number of threshold types (for just the last juror) is exponential in the size of the jury. This unfortunate fact restricts the applicability of our exact approach to small juries. Even for a jury of $n = 3$ jurors, since we have 11 potential threshold levels, and five possible voting histories $(-, A, B, AB, BA)$, there are $11^5 = 161,051$ threshold profiles to optimize over. However juries, as opposed to general electorates, are usually small. Three judges often decide a case (as in boxing and weight-lifting matches, X-Factor competitions) and sometimes only two (as in tennis or badminton line calls with overrule). Many legal decisions are determined by a three-judge panel, and appellate courts are often three tiered. Three-person juries are analyzed experimentally in [9]. Also, we show how some of our results can be applied to large juries by decomposing them into smaller ones, and how our results on ability ordering apply more generally to the last three jurors in the voting order.

The methodology adopted in this paper is neither experimental nor deductive. Rather we use computer based programming (integer programming or dynamic programming) to determine the set of optimal threshold profiles for each *a priori* probability of state A and each (ordered) sequence of juror abilities, and then exhaustive search to obtain comparative general results which hold for all or specified parameters. We use exact calculations (with fractional probability values) based on *Mathematica*. Hence these general results, which we call Propositions, have no proofs. The ones with true deductive proofs are called Theorems.

2 The model

There are two states of Nature, A and B (such as *guilty* and *innocent*), with the *a priori* probability θ of A. The symmetric case $\theta = 1/2$ is referred to as *of neutral alternatives*. A group of n agents (jurors) attempts to decide the true state of Nature by amalgamating their private information through sequential voting towards a verdict V of A or B. Their common aim is to maximize the probability Q that their collective verdict is the actual state of Nature.

We model the voting problem $\Gamma = \Gamma_{m,n}$ in terms of the minimum number of votes m (out of n) required for a verdict of A. We are primarily concerned with *majority voting*, where n is odd and a majority of $m = (n + 1)/2$ is sufficient for either alternative. We shall also sometimes relate our problem to one of *unanimous voting* $\Gamma_{n,n}$ (for say A), where one of the alternatives requires a unanimous vote. The agents will make their votes strategically (depending on θ, previous voting, and their private information). The jurors differ in their *ability* to discern the state of Nature, so the voting order may matter.

2.1 Signals and thresholds

We model the private information of the ith juror to vote as a *signal* s_i drawn from a fixed *signal set* $\mathcal{S} = \{-9, -7, \ldots, -1, +1, \ldots, +7, +9\}$ consisting of ten odd numbers. Positive signals will tend to indicate that A is true; negative signals, B. A higher positive signal gives a higher conditional probability of A, so is considered stronger. Similarly for negative signals.

When juror k comes up to vote, he votes A if his signal s_k is above a threshold $\tau_k = \tau_k(v_1, v_2, \ldots, v_{k-1})$, where $v_j \in \{A, B\}$ is the vote of the jth earlier juror, $j = 1, \ldots, k-1$. That is

$$v_k = A \text{ if and only if } s_k > \tau_k.$$

When there are $n = 3$ jurors, the threshold vector τ has five coordinates for majority voting:

$$\begin{aligned} \tau &= \left(\tau_1, \tau_2\left(B\right), \tau_2\left(A\right), \tau_3\left(AB\right), \tau_3\left(BA\right)\right) \\ &= \left(v, w, x, y, z\right), \end{aligned}$$

as $\tau\left(AA\right)$ and $\tau\left(BB\right)$ can be ignored since a majority has already been reached. Since signal indices are odd, we index the thresholds $\mathcal{T}$ by even numbers: $\mathcal{T} = \{-10, -8, \ldots, +10\}$. Note that threshold -10 is a certain vote for A and $+10$ is a certain vote for B. For example, given signal vector $s = (s_1, s_2, s_3) = (1, -3, 5)$ and threshold vector (strategy profile) $\tau = (2, -4, 6, -8, 10)$, the voting sequence is $v_1 = B$ (because $s_1 = 1 < 2 = \tau_1$), $v_2 = A$ (because $s_2 = -3 > \tau_2\left(v_1\right) = \tau_2\left(B\right) = -4$), and $v_3 = B$ (because juror 3 always votes B after BA as that threshold is $+10$). The voting sequence BAB determined by s and τ thus gives a majority verdict $V = V(BAB)$ of B.

2.2 Signal distributions and juror abilities

We now discuss the process by which individual jurors receive their private information in the form of signals. If Nature is in state A (resp. B), positive (resp. negative) signals should be more likely than negative (resp. positive) signals. Moreover, we want to have probabilities of A and B increase with signal strength, measured by the absolute value of $s \in \mathcal{S}$. The simplest such signal distributions are two linear functions f and g on $\mathcal{S}$ respectively for A and B with absolute values of their slopes proportional to juror's ability $a \in \Omega = \{0, 1, 2, 3, 4\}$. Hence a juror with a higher ability is more likely to guess the correct state of Nature. Specifically, a juror of ability $a \in \Omega$ receives signal $s \in \mathcal{S}$ with probability

$$\begin{cases} f_a(s) = \frac{1}{10} + a\left(\frac{s}{360}\right), & \text{if Nature is } A; \\ g_a(s) = f_a\left(-s\right), & \text{if Nature is } B. \end{cases}$$

Note that a juror of ability $a = 0$ essentially has no private information as $f_0(s) = g_0(s)$, so any signal he receives is equally likely to come from A or B. Also note that the maximum ability is chosen as $a = 4$ so that $f_4(-9) = g_4(9) = 0$ remain probabilities (i.e., non-negative number).

The cumulative distribution functions F_a and G_a, corresponding to the probability densities f_a and g_a, determine the probabilities that a juror of ability a will vote B. In particular, noting that thresholds are necessarily even, we have,

$$\Pr[\text{Event1}] = F_a(2j) = (5+j)\left(\tfrac{1}{10} - \tfrac{a(5-j)}{360}\right),$$
$$\Pr[\text{Event2}] = G_a(2j) = (5+j)\left(\tfrac{1}{10} + \tfrac{a(5-j)}{360}\right);$$

where Event1 (resp. Event2) denotes the event that, given the state of Nature is A (resp. B), a juror of ability a with threshold $\tau = 2j$ votes B. The significance of the ability parameter is easily demonstrated in the context of a jury of one. Here, a juror of ability a, faced with neutral alternatives, maximizes the probability Q_a of giving the right verdict by using a neutral threshold of $\tau_1 = 0$ with

$$Q(a) = \frac{G_a(0)}{F_a(0) + G_a(0)}.$$

Hence

$$\begin{aligned}(Q(0),\, Q(1), Q(2),\, Q(3),\, Q(4)) &= \left(\frac{1}{2}, \frac{41}{72}, \frac{23}{36}, \frac{17}{24}, \frac{7}{9}\right)\\ &\approx (0.50, 0.57, 0.64, 0.71, 0.78).\end{aligned}$$

2.3 Probability of correct verdict

Suppose we have three jurors of abilities a, b, c in order of voting. If we know their thresholds $\tau = (v, w, x, y, z)$, then for each state of Nature we can evaluate the probability of all voting sequences and consequently the probability Q that the verdict is the actual state of Nature, that is, the probability it is correct. If Nature is in state A, then the three voting sequences AA, ABA and BAA lead to a correct verdict; if Nature is in state B, then this holds for BB, BAB and ABB. For example, the probability that Nature is in state A and the voting sequence is ABA is given by $\theta(1 - F_a(v))F_b(x)(1 - F_c(y))$. More generally, we can write Q as the sum of the six possible ways of getting a correct verdict as follows:

$$\begin{aligned}&Q(\theta;\, a, b, c;\, v, w, x, y, z) =\\ &\quad \theta((1 - F_a(v))(1 - F_b(x)) + (1 - F_a(v))F_b(x)(1 - F_c(y))\\ &\quad + F_a(v)(1 - F_b(w))(1 - F_c(z))) + (1 - \theta)(G_a(v)G_b(w)\\ &\quad + G_a(v)(1 - G_b(w))G_c(z) + (1 - G_a(v))G_b(x)G_c(y)).\end{aligned}$$

For a fixed order in which jurors of different abilities may vote, we jointly optimize their thresholds and call the optimal probability of correct verdict $\bar{Q} = \bar{Q}\,(\theta;\, a, b, c)$. We solve the following integer program:

$$\bar{Q}\,(\theta;\, a, b, c) = \max_{v,w,z,y,z \in \mathcal{T}} Q(\theta;\, a, b, c;\, v, w, x, y, z).$$

Denote by $\bar{\tau} = \bar{\tau}(\theta; a, b, c) = (\bar{v}, \bar{w}, \bar{x}, \bar{y}, \bar{z})$ any optimal (strategic) threshold profile. The evaluation of $\bar{Q}$ for the case of three jurors may also be carried out via dynamic programming. For large juries, our discrete approach is not feasible, but some observations extended from the case $n = 3$ are mentioned later. Corresponding to our discrete signals, thresholds and juror abilities, we consider the *a priori* probability $\theta \in \{0.1, \dots, 0.9\}$.

In contrast to the optimal (strategic) threshold profile $\bar{\tau}$, it is useful to consider what is called *naive voting*, with naive thresholds $\tilde{\tau} = (\tilde{v}, \tilde{w}, \tilde{x}, \tilde{y}, \tilde{z})$ and correctness probability $\tilde{Q} = Q(\theta;\, a, b, c;\, \tilde{\tau})$. Naive voting is best defined recursively. The first juror votes A if his subjective probability of A, given *a priori* probability θ and his private information s_1, is at least $1/2$. For $\theta = 1/2$, this means that his naive threshold is 0. Suppose that the first $k-1$ jurors have chosen naive thresholds and the voting has gone in some sequence $(v_1, v_2, \dots, v_{k-1})$. Then the kth juror votes naively if he votes for the more likely state of Nature, given all this information and his signal s_k. His naive threshold $\tilde{\tau}_k\,(v_1, v_2, \dots, v_{k-1})$ is the even number between the lowest (odd-numbered) signal for which he votes A and the highest for which he votes B. Thus $\tilde{\tau}$ can be computed recursively (for any number of voters n). An interesting question is when (if ever) is $\tilde{\tau} = \bar{\tau}$ and hence $\tilde{Q} = \bar{Q}$. That is, when is naive voting optimal? This will be partially addressed in Proposition 5.

We note that instead of optimizing the probability Q of a correct verdict, we could specify positive costs C_1 and C_2 for verdict $V = A$ when Nature is B or verdict $V = B$ when Nature is A, and then minimize the expected cost. It turns out that our qualitative results in this case are not different from the case $C_1 = C_2$, which is identical to maximizing Q.

3 Optimal thresholds and need for strategic voting

This explanatory section gives some examples that demonstrate the necessity of strategic voting and motivate our results on optimal voting orders presented in the following section. For simplicity we assume neutral alternatives in this section, $\theta = 1/2$.

$a_1 \backslash a_2$	0	1	2	3	4
0	$(\mathbf{0,-10,10,0,0})$	$(\mathbf{0,-10,10,0,0})$	$(\mathbf{0,-10,10,0,0})$	$(0,-6,6,2,-2)$	$(0,-4,4,2,-4)$
1	$(\mathbf{0,-10,10,-2,2})$	$(\mathbf{0,-10,10,-2,2})$	$(2,-8,10,-2,0)$	$(0,-6,6,0,0)$	$(0,-4,2,2,-2)$
2	$(\mathbf{0,-10,10,-2,2})$	$(\mathbf{0,-10,10,-2,2})$	$(0,-8,8,-2,2)$	$(0,-4,4,0,0)$	$(0,-2,2,2,-2)$
3	$(\mathbf{0,-10,10,-4,4})$	$(\mathbf{0,-10,10,-4,4})$	$(\mathbf{0,-10,10,-4,4})$	$(0,-4,4-2,2)$	$(0,0,2,0,-2)$
4	$(\mathbf{0,-10,10,-6,4})$	$(\mathbf{0,-10,10,-6,6})$	$(0,-8,8,-4,4)$	$(0,-2,2-2,2)$	$(0,0,0,0,0)$

Table 1: Optimal thresholds after $(-, B, A, AB, BA)$ for voting order $(a_1, a_2, 4)$ with $\theta = 1/2$

For illustrative purposes we have listed in Table 1 an optimal threshold profile for each jury with a final voter of ability 4. Of particular interest (to be discussed below) are the zero threshold profile for homogeneous abilities $(4, 4, 4)$, the skewed (asymmetric) profile for $(1, 2, 4)$, and the extreme (± 10) thresholds for the second voter in the bold profiles.

3.1 Homogeneous jurors

Apart from sequential voting, our model includes two elements not ordinarily found in the literature: multiple (rather than binary) signals and heterogeneous jurors (of differing abilities). It is the latter assumption that creates the rich and counterintuitive flavor of our model. This can be easily demonstrated by simply calculating the five homogeneous thresholds uniquely as $\bar{\tau}(1/2, a, a, a) = (0, 0, 0, 0, 0)$ for all abilities $a \in \Omega$ (except that the uniqueness does not hold for ability $a = 0$ for an obvious reason), as indicated in Table 1 for the case of $a = 4$. That is, each juror votes A exactly when he receives a positive signal. We can rephrase this as the following elementary observation.

Proposition 1 *Facing neutral alternatives, homogeneous jurors can vote optimally by voting naively without observing previous votes, as if they were voting in a secret (or simultaneous) ballot.*

Thus when jurors are homogeneous, each can indeed vote as if they are the pivotal voter, and ignore any prior voting that they witness. Furthermore, we have the following simple results concerning conviction rate and verdict errors when compared with unanimous voting system $\Gamma_{3,3}$, which are further detailed in Table 2.

Proposition 2 *For homogeneous jurors with optimal voting strategies, majority rule leads to a higher conviction rate but lower rate of wrong acquittals* *when compared with unanimous voting rule. On the other hand, majority rule has a lower rate of wrong conviction for the highest-ability jury ($a = 4$), but otherwise a higher rate (for $a < 4$).*

We note that the lower rate of wrong conviction for majority rule was a major finding of [8].

Ability	P_m	P_u	E_m	E_u	F_m	F_u
1	.5	.398	.397	.311	.397	.514
2	.5	.417	.297	.242	.297	.407
3	.5	.449	.206	.184	.206	.285
4	.5	.494	.126	.136	.126	.148

Table 2: Comparison of voting rules for homogeneous jurors: strategic sequential majority (m) vs. unanimous (u): P conviction rate, E rate of conviction error, F rate of acquittal error

3.2 Symmetric strategy profiles

Table 1 gives some optimal profiles for the neutral alternative case $\theta = 1/2$, where there is an obvious symmetry between the two states of Nature, A and B. To exploit this symmetry, we define the transposition function given by $\hat{A} = B$ and $\hat{B} = A$. We can extend this to partial voting histories $v = (v_1, v_2, \ldots, v_{k-1})$, where $v_i \in \{A, B\}$, $1 \leq i \leq k-1$ and $k = 1, \ldots, n$, by defining $\hat{v} = (\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_{k-1})$ and to thresholds by $\hat{\tau}(v) = -\tau(\hat{v})$. For example, if profile τ votes A after previous voting B with signal $s_2 > 4$, then $\hat{\tau}$ votes B after previous voting A with signal $s_2 < -4$. Clearly for $\theta = 1/2$ the profiles τ and $\hat{\tau}$ yield the same correctness probability Q. Thus if τ is optimal for some ability parameters, so is $\hat{\tau}$. Thus optimal solutions come in pairs (just like conjugate pairs for quadratic equations). We call a profile τ symmetric if $\tau = \hat{\tau}$, or equivalently if $\tau(v) = -\tau(\hat{v})$ for all voting histories v. In particular, for symmetric thresholds τ we have $\tau_1 = \tau(\text{no history}) = -\tau_1$, i.e., $\tau_1 = 0$. One question we may naturally ask is whether for neutral alternatives $\theta = 1/2$ there is always an optimal threshold $\bar{\tau}$ that is symmetric.

To answer this question, look at the entry in Table 1 for the optimal threshold for abilities $(1, 2, 4)$: $(2, -8, 10, -2, 0)$. The transposed threshold is of course also optimal, but these are the only two. So in particular there is no optimal threshold profile for $(1, 2, 4)$ with first coordinate $\tau_1 = 0$, no symmetric optimal threshold. Thus, despite the symmetric nature of the optimization problem, the first juror must skew his vote to either B (by requiring a signal more than 2 to vote A) or to A (requiring a signal less than -2 to vote B). Given this example, it is perhaps surprising to note the following general result, which says that such skewing is not necessary for particular voting orders.

Proposition 3 *If $\theta = 1/2$, the jurors' abilities are labeled $a \leq b \leq c$ and the juror of middle ability b votes first, then there is an optimal threshold $\bar{\tau}$ that is symmetric.*

In Table 1 we have listed a symmetric threshold profile whenever one of these is optimal, and note that for the ability ordering $(2, 1, 4)$ where the middle ability juror votes first, there is indeed a symmetric threshold. We will see later that voting orders with middle-ability jurors first are of more special interest.

3.3 Two yokels and a boffin

Some special combinations of abilities $\{a, b, c\}$ lead to particularly intuitive results. Suppose the jurors have abilities $a, b < c$, where two of the jurors have low abilities compared with the third. With no offense intended we call the situation *two yokels and a boffin* (2Y1B), when an optimal strategy (threshold profile) has the second yokel always vote opposite to the first one, thus canceling out his vote. Such situations are highlighted in bold in Table 1 with the second yokel's thresholds $(-10, 10)$. This leaves the real decision up to the boffin. However, in some cases the boffin may obtain useful information from the vote of the first yokel.

Note that the 2Y1B phenomena demonstrated in Table 1 depend on voting orders, whose optimality we will address in the next section. To get some intuition for our main results, let us consider voting order $(1, 0, 4)$ in the 2Y1B context. We compare naive voting, where each juror votes for the alternative that he believes is most likely, with strategic (optimal) voting. In naive voting, the first juror (smart yokel) believes that A is more likely if and only if he gets a positive signal, so his threshold is 0. The second juror has a meaningless signal, as his ability is 0. Therefore, whatever the first yokel votes for will be copied by the second yokel. Thus with naive voting, juror 1 (with low ability) is the sole determinant of the verdict. The boffin is never even consulted (assuming voting ceases after a majority is reached)! This is a miniature example of an information cascade, which is avoided by strategic voting as seen below.

With strategic voting, the optimal thresholds for voter 2 (second row and first column of Table 1) require him to always vote the opposite of voter 1. This leaves the verdict up to the boffin, juror 3, of ability 4. Clearly it is better for the boffin of ability 4 to make the decision than the yokel of ability 1, an improvement from 0.57 to at least 0.78 as calculated at the end of the section on signals distributions and juror abilities. Of course in some cases (where he gets a weak signal), the boffin may improve further by taking into account the vote of the first voter.

When we are in the 2Y1B situation, the vote of the second yokel carries no information; but the vote of the first yokel does (if his ability is not 0). If the boffin votes *after* the first yokel, he can go with the first yokel's vote if his own signal is very small, say ± 1. (This argument would be even cleaner if we allowed a neutral signal of 0). So in order for the boffin to obtain useful information from the smart yokel, we need two conditions:

- The smart yokel must vote before the complete yokel, and
- The smart yokel must vote before the boffin.

Hence the only ordering in which the boffin can make use of the information contained in the smart yokel's vote is:

- The smart yokel votes first.

Note that in our example this means that the juror of middle ability votes first in the optimal ordering. This turns out to be true generally.

3.4 One complete yokel

We conclude this section with the analysis of a special case that will have applications later. The *a priori* probability of A is an arbitrary θ. The three abilities of the jurors, in voting order, are $a, 0, b$, where $a, b > 0$. That is, the middle voter is a complete yokel. Suppose that juror 1 votes A. If juror 2 votes A, the verdict is A and the last juror never gets to vote. Clearly the jury can do at least as well if juror 2 votes B and leaves the voting up to the last juror, because the last juror can always vote A and do the same as in the previous case. (Of course if the subjective probability of A for the last juror is less than 1/2 he can vote B and do better.) The same reasoning applies if juror 1 votes B. We can also check this argument by establishing via integer programming that

$$\bar{Q}(\theta, a, 0, b) = \max_{v,y,z \in \Omega} Q(\theta;\, a, 0, b;\, v, -10, 10, y, z),$$

or equivalently $(\bar{w}, \bar{x}) = (-10, 10)$. Summarizing this last argument, we have shown the following lemma.

Lemma 4 *For any θ and positive abilities a and b, the voting ability order $(a, 0, b)$ has optimal thresholds for the second juror of $(\bar{w}, \bar{x}) = (-10, 10)$. That is, he always votes against the fist juror.*

4 Main results

In this section we present our main results on the optimal voting orders for jurors of differing abilities (Proposition 5).

Given a *set* of abilities $\{a, b, c\}$ for three jurors, with $a \leq b \leq c$, what voting order maximizes the probability $\bar{Q}$ of obtaining the correct verdict, and how much is this optimal probability? Our approach is to calculate $\bar{Q}$ for various orderings of the jurors and attempt to spot patterns of optimality. We then check these patterns by exhaustive search over all voting orders and values of θ to see where they hold.

To give the reader a small taste of the pattern recognition problem, Table 3 provides some results about how the verdict correctness probability $\bar{Q}$ and the corresponding optimal threshold profile are dependent on the voting order of the three jurors of different (unordered) abilities $\{1, 2, 3\}$ and $\{1, 2, 4\}$. As some patterns only appear in the alternative-neutral case of $\theta = 1/2$ we take this and $4/5$ for our *a prioi* probability of A. Some observations from Table 3 that we have calculated to hold for general parameters are stated in our main results below:

Proposition 5 *For the majority voting problem $\Gamma = \Gamma_{2,3}$ of three jurors with linear signal distributions, a priori probability θ of A and an arbitrary set of three juror abilities $a, b, c \in \Omega$ labeled so that $a \leq b \leq c$, we have*

1. *The probability $\bar{Q}$ of a correct verdict is maximized when the **middle-ability juror votes first**. The order of the last two jurors does not affect $\bar{Q}$. Thus the two optimal orderings are (b, a, c) and (b, c, a).*

2. *When jurors vote naively, it is always optimal for the weakest to vote last. Thus the optimal ordering is either (b, c, a) or (c, b, a).*

3. *With neutral alternatives $\theta = 1/2$ and the voting order (b, c, a), where $a < b < c$, voting naively is optimal for maximizing Q. For other voting orders naive voting is suboptimal.*

θ, Ability Order		$\bar{Q}$	$\tilde{Q}$
4/5,	$(2,1,3)$	**.827**	.800
4/5,	$(2,3,1)$	**.827**	.815
4/5,	$(1,2,3)$	.825	.800
4/5,	$(1,3,2)$	.825	.8169
4/5,	$(3,1,2)$	.825	.815
4/5,	$(3,2,1)$	.825	**.8172**
1/2,	$(2,1,4)$	**.794**	.639
1/2,	$(2,4,1)$	**.794**	**.794**
1/2,	$(1,2,4)$	.781	.718
1/2,	$(1,4,2)$	.781	.781
1/2,	$(4,1,2)$	.780	.778
1/2,	$(4,2,1)$	.780	.778

Table 3: Comparison of ability orders

4.1 Overrules in tennis

As an application of Proposition 5, we consider the problem faced by the tennis line-calling jury consisting of a linesman and umpire. Here, ability might be related to eyesight, concentration and blinking rate. Signals might depend on more immediate phenomena such as sun position, player position (blocking view) and weather. The linesman votes first (A "In" or B "Out") and then the umpire can overrule. In practice, the umpires are instructed to overrule only when very sure of the call, but we analyze the problem in terms of the linesman/umpire team maximizing the probability of a correct verdict, in line with our previous analysis. To conform to our earlier notation, we always have the umpire vote (A or B), even if he is agreeing with the linesman. The umpire determines the verdict. If we have two people of different abilities (e.g., unequal eyesight), what roles should we assign them to? It turns out that by introducing a third referee, who is blind and votes after the linesman and before the umpire, and requiring a majority verdict, we can reduce the tennis problem to our earlier context.

Theorem 6 *Given two referees with distinct abilities a and b we should assign the line-calling to the one with lower ability and the umpiring to the one with higher ability.*

Proof. Consider the three-person jury majority game Γ with ability ordering $(a, 0, b)$, where a complete yokel votes in the second position. By Lemma 4, we know that the yokel will optimally vote against the first juror, leaving the verdict up to the third juror. Thus the tennis jury game is equivalent to this majority game Γ. From Proposition 5 we know that the middle ability juror should vote first in Γ, whose ability in the set $\{a, 0, b\}$ is the weaker of a and b. Hence the linesman should be the juror of ability $\min\{a, b\}$ and the umpire should have ability $\max\{a, b\}$. $\square$

5 Larger juries

We now show how some results for three jurors can be used to analyze larger juries of $n > 3$ jurors. Note that with unanimous voting (for say A), each juror i (of ability a_i) has a single threshold strategy $\tau_i = \tau_i(A, A, \ldots, A)$ as for the other voting sequences the verdict is already decided. So the probability Q of getting the correct verdict (with *a priori* probability θ of A) with unanimous voting is given by

$$Q = \theta \prod_{i=1}^{n} \left(1 - F_{a_i}(\tau_i)\right) + (1 - \theta)\left(1 - \prod_{i=1}^{n} \left(1 - G_{a_i}(\tau_i)\right)\right).$$

Observe that this is the same formula as for the case of *simultaneous* unanimous voting for B, when each juror simply compares their private signal s_i to their threshold τ_i (regardless of other jurors' votes). Therefore, *voting order does not matter.* Consequently, we have the following elementary result.

Theorem 7 *In sequential unanimous voting, the voting order of the jurors does not affect the correctness probability Q.*

Now consider sequential majority voting when $n = 3$. After the first vote by juror 1, the subproblem faced by jurors 2 and 3 is either $\Gamma_{2,2}$ on unanimous voting for A (if juror 1 chose B) or $\Gamma_{1,2}$ on unanimous voting for B (if juror 1 chose A). So in either case Theorem 7 demonstrates that the voting order of the last two jurors does not matter. This "explains" the similar observation for linear signals demonstrated computationally in Proposition 5. We will state this result more generally in the last part of Theorem 8.

Next consider the majority voting problem $\Gamma_{m,n} = \Gamma_{k+1,\,2k+1}$ when number of jurors $n = 2k + 1$ is odd and greater than 3 and signals have linear distribution. Suppose the first $n - 3 = 2(k - 1)$ jurors have voted, with a difference d (necessarily even) between the number of votes for A and B. Consider the subproblem faced by the last three jurors. If $|d| \geq 4$ then the verdict is already settled. If $d = 0$, then each alternative has received $k - 1$ votes, so the remaining subproblem is $\Gamma_{2,3}$, the same as the majority voting problem for three jurors (with some ex ante probability θ' for A, which depends on the prior voting sequence), which as we showed in part 1 of Proposition 5 is best handled for linear signals with the middle-ability juror voting first. Finally, if $|d| = 2$ then, depending on the sign of d, the subproblem is a unanimous voting problem for one of the alternatives, and by Theorem 7, the voting order does not matter. So the ordering is either irrelevant or should be middle ability first, leading to the following theorem.

Theorem 8 *Consider sequential majority voting with an odd number of jurors having linear signals. In an optimal voting order, the last three to vote should be ordered with the middle-ability voter first. More generally, for any signal distributions, the order of the final two jurors does not affect the optimal probability of a correct verdict.*

We observe that actually our argument proves a stronger version of the above proposition, namely that *even after observing the first $n - 3$ votes*, we would never want to change the order of the last three from an ordering with the middle ability first.

6 Concluding remarks

Using a simple model of sequential voting, with three heterogeneous jurors differentiated by their ability to discern the true state of Nature, we show that the probability of reaching the correct verdict is maximized when the middle-ability juror is the first to vote. This ordering applies as well to the relative abilities of the last three jurors in a larger jury. We also show that the jurors can afford to vote naively between equiprobable alternatives if and only if the voting order of abilities $a < b < c$ is (b, c, a), which is shown to be the unique optimal order when restricted to naive voting.

It would be useful to find optimal orderings for larger juries and for voting schemes $\Gamma_{m,n}$ other than simple majority.

References

[1] Condorcet, M.J.A.N.C. (1785). *Essai sur l'Application de l'Analyse à la Prababilité des Decisions Rendues à la Pluralité des Voix* (Impr. Royale, Paris); reprinted (1972) (Chelsea, New York).

[2] Feddersen, T. and Pesendorfer, W. (1999). Elections, information aggregation, and strategic voting. *Proceedings of the National Academy of Sciences* 96(19): 10572–10574.

[3] Easley, D. and Kleinberg, J. (2010). *Networks, Crowds and Markets.* Cambridge University Press.

[4] Feddersen, T.J. and Pesendorfer, W. (1997). Voting behaviour and information agregation in elections with private information. *Econometrica* 65 (5): 1029–1058.

[5] Dekel, E. and Piccione, M. (2000). Sequential voting procedures in symmetric binary elections. *Journal of Political Economy* 108 (1): 34–55.

[6] Scharfstein, D. and Stein, J. (1990). Herd behavior and investment. *American Economic Review* 80 (3): 465–479.

[7] Ottaviani, M. and Sorensen, P. (2001). Information aggregation in debate: who should speak first? *Journal of Public Economics* 81 (3): 393–421.

[8] 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.

[9] Ali, S., Goeree, J., Kartik, N., and Palfrey, T.R. (2008). Information amalgamation by voting. *American Economic Review: Papers & Proceedings* 98 (2): 181–186.

[10] Jon Elster (2013). *Securities Against Misrule: Juries, Assemblies, Elections.* Cambridge University Press, New York. ISBN 978-1-107-03173-9, 324 pages.

[11] Alpern, S., Gal, S., and Solan, E. (2010). A sequential selection game with vetoes. *Games and Economic Behavior* 68 (1): 1–14.