Votes Based on Protracted Deliberations

William S. Neilson
Department of Economics
University of Tennessee
Knoxville, TN 37996-0550
wneilson@utk.edu

Harold Winter
Department of Economics
Ohio University
Athens, OH 45701
winter@ohio.edu

September 2006

Abstract:

We analyze group decision-making in situations in which members discuss the value of a continuous random variable and then take an up-or-down vote based on their assessments of the continuous variable. Applications include jury deliberations and tenure and promotion or partnership votes. We determine conditions under which the group can reach a consensus on the value of the continuous variable, conditions under which the group fails to reach an agreement on the up-or-down vote, and when the outcome of a successful up-or-down vote matches the consensus on the continuous variable. We also consider the impact of unanimous vs. nonunanimous voting rules.

JEL codes: D71, D72, K41

Keywords: Deliberations, consensus, committee voting, factions, unanimity

We are grateful to two anonymous referees and seminar participants at Duke University for helpful comments. Neilson thanks the Private Enterprise Research Center at Texas A&M for financial support.

1. Introduction

Before they vote on promoting an associate to partnership status, the senior members of a law firm first discuss the strengths of a candidate's record. The process by which members of an academic department decide to promote and tenure a colleague shares the same pattern of a discussion and then a vote. In the same way, before voting on whether of not to pass a bill out of a legislative subcommittee, the members of the subcommittee debate the merits of the proposed legislation. Finally, when a jury enters the deliberation room, they have the option of discussing the strength of the evidence before voting whether to convict or acquit the defendant. All of these decision-making processes share one important feature. In every process, the members of the group discuss the value of a continuous variable, such as the strength of a promotion candidate's record or the strength of the evidence against a defendant, and then take an up-or-down vote based on the outcome of that discussion.

The purpose of this paper is to construct a model of the group decision-making process in which a yes-no vote on an issue follows a discussion of its merits. This problem has already begun to receive attention in the economics literature and elsewhere. Li, Rosen, and Suen (2001) construct a model based on Crawford and Sobel's (1982) analysis of strategic information transmission. In the Li et al. model, group members not only differ in their initial information but also in their preferences over the outcome, and because of the conflicting preferences group members tend to bias their messages in line with their preferences. This leads to more disagreements. Glazer and Rubenstein (2001, 2003) model the attempts of two agents to persuade a listener. Like the Li et al.

paper, their work concentrates on the choice of message to send.[1] In their papers, the two debaters have multiple arguments they could possibly make but are constrained to making only a subset of the possible arguments. Feddersen and Pesendorfer (1998) and Coughlan (2000) focus on the optimal voting rule for a juror. In equilibrium a juror votes as if he is pivotal, and being pivotal imparts different information about the signals of the other jurors depending on the verdict rule. Finally, a large literature analyzes group decisions in terms of power indices (e.g. Straffin, 1994; and Saari and Sieberg, 2001). These papers concentrate on the power a group member has based on the number of ways in which he or she might be pivotal.[2]

While these papers analyze many of the important aspects of the types of decisions we are interested in, three features have yet to receive attention. First, discussions are often protracted with the same member contributing information many times. Second, there may be more than one vote before a verdict is reached, especially in jury deliberations. Finally, influence is differential within the group, in that a member may listen to some of his colleagues but not others when updating his or her assessment. While the Glazer and Rubinstein papers allow for discussion of different types of information so that the debate can last more than one round, the driving force in their papers is the constraint that makes it impossible for the debaters to reveal all of their information. In the Li et al., Feddersen and Pesendorfer, and Coughlan papers, all information is revealed in the first round of discussion, and further discussion serves no purpose. This is a characteristic of any Bayesian

---

[1] The Glazer and Rubenstein model is also better suited for analyzing the choices made by the lawyers during a trial rather than the deliberation by jurors after the lawyers finish.

[2] In addition to all of these models with communication, Klevorick and Rothschild (1979) analyze a model of jury deliberation with no communication at all. In their work the deliberation process is a Markov process with the jury's vote as the state. They show how the transition matrix determining the evolution of votes can be estimated, and use it to illustrate the majority persuasion hypothesis, i.e. the first ballot almost always decides the outcome of the verdict, and how the unanimity rule affects the amount of deliberation.

treatment of the problem. Because all of the information is revealed in one round of discussion, nothing is gained by voting more than once.[3] Finally, because Bayesian updating is used, every group member is influenced by every other group member.

We propose a non-Bayesian model to capture these three features of deliberation. The model allows the outcome of the vote to be affected by the discussion, in which case it becomes important that during the discussion each member of the group may or may not influence all of the others. Consequently, the model also accounts for the ways in which members of the group interact with each other. We are then able to determine how the outcome of the process depends on the views that the group members bring with them to the discussion, the ways in which the group members interact, and institutional factors like the timing and number of votes and whether or not unanimity is required.

The model has two components corresponding to the two different aspects of the group decision-making process. The first component captures the interactions of the members when discussing the value of the continuous variable, such as the merits of a promotion candidate or the strength of the evidence in a criminal trial. The second component specifies the voting procedure of the group, including the timing of the votes, the size of the majority required for a decision, and what happens if the required majority is not reached. All voting is assumed to be sincere.[4]

The model of the interactions of the group members in their discussions of the value of the

---

[3] Aumann and Hart (2003) analyze a cheap talk game in which communication can be persistent. Their games differ from the setting used here, though, in that here the problem is purely information aggregation without any preference conflicts, and there the problem entails the resolution of potential preference conflicts.

[4] Coughlan (2000) finds that equilibrium votes are sincere in a wide class of settings.

continuous variable is based on the consensus model of DeGroot (1974). In that model, each group member has an initial assessment of the value of the continuous variable, and the discussion proceeds through a number of periods. In each period, all group members report their current assessments of the value and then update their assessments based on the reports.[5] A member may place more weight on some of his peers than on others when updating his assessment, and may even place no weight at all on the reports of some members of the group. DeGroot's work is tied to the idea of a consensus, in which all members of the group agree on the value of the continuous variable, i.e. they all have identical assessments at some point in time.[6]

DeMarzo et al. (2003) build on DeGroot's work by adding structure consistent with individuals updating beliefs and use the added structure to derive several intriguing results. Each member receives a noisy signal of the value of a variable, and when the noise terms are normally, independently, and identically distributed a linear updating rule is optimal with the weight assigned to each member's report depending on the accuracy of that member's signal. With the extra structure they are able to show that deliberation leads to "persuasion bias," which arises, for example, when member A updates his beliefs based on the reports of both B and C, but fails to correct for the fact that B's report reflects C's prior report.[7] They also establish the phenomenon of unidimensional

---

[5] In Li et al. (2001), reports of assessments are garbled because group members have conflicting interests. We assume that assessments are truthfully revealed, which occurs when there are no conflicts of interest. DeMarzo et al. (2003) also make this assumption.

[6] DeGroot's (1974) paper has been applied to situations besides the process of arriving at a consensus. For example, Razavi (1984) uses the model to discuss OPEC decision making, Conlisk et al. (2001) use it to model how individual purchase decisions are influenced by purchase decisions of "neighbors" in a social network, and Friedkin (2001) applies DeGroot's model to the formation of social norms.

[7] Garicano and Posner (2005) suggest how the work of DeMarzo et al. (2003) can provide a contributing explanation of intelligence failures before the Iraq War.

opinions, which states that if one member has an extreme assessment along one dimension of a multi-dimensional issue, he or she necessarily has an extreme assessment, in the same direction, along every dimension.

We build on the models of DeGroot (2003) and Demarzo et al. (2003) in two ways. First, we are as interested in when groups disagree as much as when they agree, and so we are more interested in the failure to reach a consensus than were DeGroot and DeMarzo et al. Second, in our model the consensus and the decision are two different entities, with the consensus being an agreement on the value of the continuous variable and the decision being the result of an up-or-down vote. Thus, we can address the following question: If there is a consensus, and if there is a decision, do the two agree? The answer, as it turns out, depends on whether or not the vote must be unanimous.

A consensus exists if everyone in the group can be persuaded that the continuous variable takes on the same value. The vote agrees with the consensus if the decision is consistent with this agreed-upon value of the continuous variable. For example, if every partner in a law firm can be persuaded that the strength of a promotion candidate takes on a certain value, and if that value is above the standard for promotion, once the consensus is reached every partner would vote for promotion. But, if a vote is taken before all of the partners are convinced that the candidate exceeds the promotion standard, and if unanimity is not required, the candidate might get passed over for promotion. When unanimity is required, a decision that is reached will match the consensus if one exists. However, with unanimity cases can arise in which the group never successfully reaches a decision. The model is able to shed light on the forces that must be in place for that to occur.

Besides the above-mentioned result on unanimity linking the consensus to the decision, our

most basic results can be summarized as follows. A consensus exists if the group contains at least one member who influences everyone else, whether directly or through a chain of other members (A influences B, who influences C, and so on). The group reaches a decision if, at the time of a vote, all members of the group have assessments of the continuous variable on the same side of some threshold value. This does not require the existence of a consensus. For example, in a promotion case two members could always disagree on the exact strength of the candidate's record but still agree that he or she is good enough to promote. The only circumstance in which the timing of the votes, the voting rule, and the patterns of influence do not matter to the outcome of the decision process is when all of the group members' initial assessments are on the same side of the threshold. So, unless everyone agrees from the outset, the outcome of the process is open to manipulation.

2. Deliberation, Consensus, and Disagreement

In this section we review and extend a key result from DeGroot's work. To begin with, each of n individuals receives his own private signal s_i of the value of a variable σ about which a consensus must be reached, and let $\boldsymbol{s} = (s_1,...,s_n)$. It is possible for σ and each s_i to be univariate or multivariate, as long as they all have the same dimensions.[8], all of the results hold For example, in a criminal jury trial jurors deliberate about the strength of the evidence against the defendant, σ, and s_i would be juror i's assessment of the strength of the evidence against the defendant after arguments are completed. In a promotion case, σ would correspond to a vector of the candidate's attributes, and s_i would be committee member i's assessment of the candidate's strengths before discussion

[8]Our notation and terminology is for the univariate case. None of the results depend on the univariate assumption, however.

begins.[9]

Deliberation proceeds by rounds. In round t, each individual reports his latest assessment $s_i^{(t-1)}$ of σ, and hears the reports of others, where the superscript in parentheses denotes the period in which the assessment is made.[10] Individual i's new assessment of the value of σ is a convex combination of the values reported in that round. Let $A = [a_{ij}]$ be an n x n matrix with a_{ij} measuring the importance that individual i places on the report of individual j.[11] Individual i's new assessment of σ is given by

$$s_i^{(t)} = \sum_{j=1}^{n} a_{ij} s_j^{(t-1)}.$$

We require that $0 \leq a_{ij} \leq 1$ for all $i,j \in \{1,...,n\}$ and all $t = 1,2,...$, and that

$$\sum_{j=1}^{n} a_{ij} = 1$$

for all $i = 1,...,n$, so that $s_i^{(t)}$ is a convex combination of $(s_1^{(t-1)},...,s_n^{(t-1)})$. Then, in matrix notation,

$$\boldsymbol{s}^{(t)} = A\,\boldsymbol{s}^{(t-1)}.$$

---

[9] It is possible to replace the multivariate case with a univariate one if all group members share the same real-valued "evaluation function" (similar to a utility function) and then deliberate over the values arising from the evaluation function. Alternatively, the parameters of the evaluation function could also be the subject of deliberation.

[10] Accordingly, $s_i^{(0)}$ is individual i's initial assessment, $s_i^{(0)} = s_i$.

[11] In DeMarzo et al. (2003), a_{ij} is determined by the relative precision of j's private signal and whether or not i listens to j. While our approach allows for this structure, none of the results rely on it, so we let A be any nonnegative matrix whose rows sum to one.

It follows that $\boldsymbol{s}^{(t)} = A^t \boldsymbol{s}$, where A^t is the t-th power of the matrix A.

The matrix A is called an *influence matrix*, because when $a_{ij} > 0$ it means that individual i is *directly influenced by* individual j in making an assessment of σ. (In the same circumstances, we also say that j *directly influences* i.) It is assumed that $a_{ii} > 0$ for $i = 1,...,n$, so that individuals always directly influence themselves. The i-th row of the matrix shows who influences individual i, while the i-th column shows whom i influences.

Even if i is not directly influenced by j, j's assessment may be incorporated indirectly into i's. To see how this works, suppose that there are only three individuals, and consider the following influence matrix:[12]

$$A = \begin{bmatrix} 1/2 & 1/2 & 0 \\ 1/3 & 1/3 & 1/3 \\ 1/3 & 1/3 & 1/3 \end{bmatrix}.$$

Individual 1 is directly influenced by individual 2 but not by individual 3, while individuals 2 and 3 are directly influenced by everyone. After the first period, individual 1's assessment is given by $s_1^{(1)} = .5s_1 + .5s_2$, and individual 3's initial signal is ignored. However, after the second period, 1's assessment is a convex combination of all three signals because

$$A^2 = \begin{bmatrix} 5/12 & 5/12 & 2/12 \\ 7/18 & 7/18 & 4/18 \\ 7/18 & 7/18 & 4/18 \end{bmatrix}.$$

[12] Although our model allows for any number of individual voters, we will often use a three person example to highlight the intuition of our results.

So, even though 1 is not directly influenced by 3, 3's signal is incorporated indirectly through 2. In general, if i is not directly influenced by j, but there exists a sequence $k_1,...,k_m$ such that i is directly influenced by k_1, k_1 is directly influenced by k_2, and so on, with k_m directly influenced by j, then it is said that *i is indirectly influenced by j*. Overall, individual i is said to be *influenced* by individual j if j directly *or* indirectly influences i.

We are interested in the eventual outcome of the group deliberation. The vectors $\boldsymbol{s}, \boldsymbol{s}^{(1)}, \boldsymbol{s}^{(2)},...$ form a sequence, so it makes sense to talk about limits. By construction,

$$\lim_{t\to\infty} \boldsymbol{s}^{(t)} = \lim_{t\to\infty} A^t \boldsymbol{s}.$$

Because every element of A is between 0 and 1, and since all of the diagonal elements are strictly positive, $\lim_{t\to\infty} A^t$ exists and is denoted by A^∞. Consequently, the limit of the sequence $\boldsymbol{s}, \boldsymbol{s}^{(1)}, \boldsymbol{s}^{(2)},...$ also exists, and is denoted $\boldsymbol{s}^*$.

The vector $\boldsymbol{s}^*$ represents the end of the deliberation process because further deliberation produces no change in the assessed values: $A\boldsymbol{s}^* = A \cdot A^\infty \boldsymbol{s} = A^\infty \boldsymbol{s} = \boldsymbol{s}^*$. There are two possibilities. The first is that every element of $\boldsymbol{s}^*$ is the same, in which case the group has reached a *consensus* given by the scalar s^*, where $\boldsymbol{s}^* = (s^*,...,s^*)$. The second is that $\boldsymbol{s}^*$ is not a constant vector, in which case the group *disagrees*. We are interested in two issues. First, under what circumstances does a group reach a consensus, and under what circumstances does it disagree? Second, if a group reaches a consensus, what are the eventual weights placed on the different members' initial signals? In other words, do some members of the group have a bigger impact on the consensus assessment than others?

The following is our principle result on the existence of a consensus:[13]

Proposition 1. If at least one member of the group influences everyone, the group reaches a consensus.

A consensus is reached when the group is sufficiently "connected" so that at least one individual's signal is factored into everyone's eventual assessment.[14] It is not necessary for communication to be two-way for a consensus to be reached, nor that everyone be influenced by everyone else. Rather, all that is needed is for some member's signal to reach everyone in the group.

The example given above fits the requirements of Proposition 1, so a consensus is reached. The matrix A^{∞} is

$$A^{\infty} = \begin{bmatrix} 2/5 & 2/5 & 1/5 \\ 2/5 & 2/5 & 1/5 \\ 2/5 & 2/5 & 1/5 \end{bmatrix}.$$

The consensus assessment is given by $s^* = 0.4s_1 + 0.4s_2 + 0.2s_3$.

In the above example, all of the rows of A^{∞} are identical. This turns out to be the general case when a consensus is reached. Let $s^*(\boldsymbol{s})$ denote the consensus assessment that is reached when the initial signal vector is $\boldsymbol{s}$.

---

[13] All proofs are collected in an appendix.

[14] DeMarzo et al. (2003, Theorem 1), restrict attention to the case in which the group is "strongly connected," that is, every member influences every other member.

Proposition 2. Suppose that at least one member of the group influences everyone. Then there exists a vector $\boldsymbol{w}$ such that $s^*(\boldsymbol{s}) = \boldsymbol{w}\cdot\boldsymbol{s}$ for all $\boldsymbol{s}$.

The vector $\boldsymbol{w}$ is a vector of *consensus weights*, and they reflect the impact that each member has on the consensus assessment. It is straightforward to demonstrate that $w_i \in [0,1]$ and $\sum w_i = 1$, so that the w_i's truly are weights. They can be found by the formula

$$\sum_{j=1}^{n} w_j a_{ji} = w_i,$$

which holds because $A^\infty A = A^\infty$, and the above expression simply states this for each element.

In cases where a consensus is reached, it is now possible to characterize group behavior using the consensus weights in $\boldsymbol{w}$ rather than the elements of the limit matrix A^∞. Furthermore, the magnitudes of the consensus weights tell something about the importance of the different group members for the final consensus.

Consider the influence matrix shown below.

$$A = \begin{bmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix}.$$

Member 1 is only directly influenced by himself, member 2 is directly influenced by himself and member 1, and member 3 is directly influenced by himself and member 2. However, no member directly influences everyone. Still, player 1 influences everyone, because 3 is directly influenced by

2 and 2 is directly influenced by 1. According to Proposition 1 the group reaches a consensus, and the resulting consensus weights are given by $\boldsymbol{w} = (1,0,0)$. Member 1 is given all of the weight in the consensus, in which case we say he is *decisive*. This occurs because member 1 is not directly influenced by anyone else, as shown by the next proposition.

Proposition 3. If everyone is influenced by member *i*, but member *i* is not directly influenced by any other members, then member *i* is decisive.

Another example illustrates what happens when no one is directly influenced by a member. Consider the influence matrix shown below.

$$A = \begin{bmatrix} 2/3 & 1/3 & 0 \\ 1/2 & 1/2 & 0 \\ 0 & 1/4 & 3/4 \end{bmatrix}.$$

No one is directly influenced by member 3. Members 1 and 2 are directly influenced by each other, and member 3 is directly influenced only by member 2, not by member 1. Since member 2 directly influences everyone, a consensus is reached, and $\boldsymbol{w} = (0.6, 0.4, 0)$. As this example shows, when no one listens to a member, that member's weight in the consensus assessment is zero, even if no single member is decisive. It is possible to prove a more general proposition, though.

Proposition 4. If everyone is influenced by member *i*, but member *i* is not influenced by member *j*, then the consensus is independent of *j*'s signal.

If everyone is influenced by member *i*, he has positive weight in the consensus. If member *i* is not influenced by member *j*, then *j* cannot possibly have positive weight in the consensus, because otherwise member *i* would have to incorporate *j*'s signal, which he does not. Consequently, *j*'s consensus weight is zero. In the above example, everyone is influenced by both 1 and 2, but neither of them is influenced by 3, so 3's consensus weight is zero.

So far we have shown that a consensus is reached when at least one member influences everyone else, and that in the consensus a member who influences everyone else but is not directly influenced by anyone else receives a consensus weight of one and a member who does not directly influence anyone else receives a consensus weight of zero. We have not yet said anything about when a group disagrees. The contrapositive of Proposition 1 states that if the group disagrees, there must be at least two members who do not influence each other. The converse of this statement is also true, as shown by the next proposition.

Proposition 5. The group disagrees if and only if at least two members of the group are not influenced by the same individual.

A simple example illustrates the result. Suppose, once again, that the group contains three members, and that the influence matrix is given by

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2/3 & 1/3 \\ 0 & 1/3 & 2/3 \end{bmatrix}.$$

Neither members 2 nor 3 are influenced by member 1, who in turn is influenced by neither of the other two. So, according to Proposition 5, the two subgroups should disagree. It is straightforward to compute

$$A^{\infty} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{bmatrix},$$

so that member 1's final assessment is the same as his initial assessment, but members 2 and 3 average their assessments. Unless they all get the same initial assessment, then, the three members will not agree on their final assessments, and they fail to reach a consensus.

To summarize, we have established the following results concerning the existence and characteristics of a consensus. First, a consensus exists if there is at least one member of the group who influences everyone, either directly or indirectly. Conversely, no consensus exists if and only if the group contains at least two members who are not influenced by a common member. If a consensus does exist, it is a weighted average of the members' initial assessments, and these weights reflect each member's importance, or eventual influence, in the discussions. A member is decisive if the consensus assessment always matches his initial assessment, and a decisive member has a consensus weight of one. For a member to be decisive, he must influence everyone but be influenced by no one. At the other extreme, a member has a consensus weight of zero if he does not influence the member that influences everyone. All of these results pertain to general discussions that try to reach a consensus about the value of a parameter. The goal of this paper is somewhat different, however, as it concerns decisions in which the outcome is determined by a vote based on discussions

of the parameter. Before moving on to an examination of voting procedures, though, we provide an alternative characterization of consensus in the next section.

3. Factions

An alternative approach to modeling consensus can be based on the idea of factions. A *faction* is a subgroup in which every member of the subgroup is influenced by every other member of the subgroup, and no member of the subgroup is influenced by anyone outside of the subgroup. Different factions, then, are isolated from each other in terms of influence, and no signal can pass from one faction to another. It is important to note that a faction need not be a proper subgroup; rather, it is possible for the entire group to be a faction.

Our first result states that no one can be in more than one faction.

Proposition 6. No individual can be a member of two different factions.

The important feature of a faction is that everyone who influences a given member of the faction must also be in that faction. Suppose that individual i is influenced by individual j, and that j is a member of some faction, call it C_k. If i is in a faction, he must also be in faction C_k because of the requirement that no member of a faction is influenced by anyone outside of the faction. However, i may or may not be a member of a faction. If i influences j, then i is in C_k. But, if i does not influence j, then the requirement that every member of a faction influence every other member of that faction is violated, and so i cannot be a member of any faction. An individual who is in the group but is not a member of any faction is a member of the *fringe*.

The key distinction between the fringe and a faction has to do with influence. Members of factions are not influenced by anyone outside of their own faction, but members of the fringe must be influenced by someone outside of the fringe, as shown by the next proposition.

Proposition 7. If i is a member of the fringe then there exists some individual j not in the fringe who influences i.

Because a member of the fringe is, by definition, not a member of any faction, and because the factions cannot overlap, given the influence matrix A one can partition the group into a fringe F and a set of factions $C_1,...,C_m$ with $m \in \{1,...,n\}$ such that $C_i \cap C_j = C_i \cap F = \varnothing$ for all $i,j \in \{1,...,m\}$. One can then decompose the influence matrix into the canonical form

$$A = \begin{bmatrix} A_1 & 0 & \dots & 0 \\ 0 & \ddots & 0 & \vdots \\ \vdots & & A_m & 0 \\ B_1 & \dots & B_m & B_F \end{bmatrix} \tag{1}$$

The first m rows and columns of the above decomposition correspond to the factions $C_1,...,C_m$, and the final row and column correspond to the fringe group F. The submatrix A_i governs how members of faction C_i influence each other. Because no member of a faction is influenced by anyone outside of the faction, the only positive elements in any of the first m rows are on the diagonal. Fringe members are influenced by faction members, though, and so at least one of the matrices in $B_1,...,B_m$

must have positive elements.

Because each row in the submatrix A_k sums to one, A_k can be thought of as an influence matrix for faction k.[15] Furthermore, because every member of faction C_k influences every other member of C_k, by Proposition 1 every faction reaches its own consensus. The next proposition follows immediately from this observation.

Proposition 8. (i) The group reaches a consensus if and only if there is a single faction. (ii) A member i is decisive if he is a one-person faction and everyone else is in the fringe. (iii) Fringe members have no impact on the consensus assessment of the evidence.

Proposition 8 reformulates the results of Section 2 using the idea of factions. As it demonstrates, the key to reaching a consensus is having everyone who is not in the fringe contained in the same faction. Furthermore, members of the fringe play no role, and have no weight in the consensus assessment.

4. Voting with Deliberation

The consensus model presented above can easily be fit into a model in which the group votes after deliberating, such as with juries (which will be our main example throughout this section). Assume that each member of the group receives a private signal about the value of the variable in

[15]Using our terminology, a matrix as *irreducible* if every member influences every other member (e.g. Feller, 1968). Each submatrix A_k in the canonical form of (1) is irreducible, and Proposition 8 below states that a consensus exists if the canonical form possesses a single irreducible submatrix. It follows that the original matrix A need not be irreducible for a consensus to exist.

question and, during deliberation, they discuss these signals. A key difference, however, between a process like jury deliberation, for example, and the consensus model is that the task of a jury is to decide whether or not a defendant is guilty, not to reach a consensus about the strength of the evidence. While a juror's vote on the verdict is, presumably, related to his assessment of the strength of the evidence, the link must be made explicit.

Member i receives a signal s_i about the strength of the variable σ that is at issue. For ease of exposition, assume that σ is univariate, although that assumption has no impact on our results. The members deliberate for a number of periods and then vote after deliberation period t. Members can vote either "yes" or "no." There is a threshold σ^* such that if the member's assessment at time t satisfies $s_i^{(t)} \geq \sigma^*$ he votes "yes", and he votes "no" otherwise.[16] For example, in a criminal jury trial, σ would be the true strength of the evidence against a defendant, $s_i^{(t)}$ would be juror i's assessment of the strength of the evidence at time t, and σ^* would be the reasonable doubt standard. A juror votes "yes" to convict the defendant if he believes the strength of the evidence exceeds the reasonable doubt standard, and he votes "no" to acquit if he believes the evidence fails to meet the reasonable doubt standard.

A "verdict" is reached if a sufficient number of members vote the same way; otherwise the deliberation continues. The vote on a verdict may or may not come before the members have reached a consensus about the value of σ, and a verdict can occur even if a consensus has not been reached. A simple example illustrates this fact. Suppose that the influence matrix is the identity matrix, so that no member influences any other members. Obviously, the group fails to reach a

[16]We are assuming that all members have the same threshold σ^*. The model can be adapted to make the appropriate threshold also a subject of deliberation, as the consensus model can accommodate multi-dimensional decision variables.

consensus on the value of σ unless they all receive the same signal. Still, if $s_i < \sigma^*$ for all i, every member votes "no," and the group reaches a verdict the first time it votes.

An important feature of some deliberation/voting processes is that sometimes the group fails to reach a decision. Specifically, the group decides that they are deadlocked and stops deliberating without agreeing on a verdict. This is especially important in jury deliberation. To allow for this possibility, let a *deliberation rule* be a set of times $T = \{t_1, t_2, ..., t_F\}$ governing when the group votes. More specifically, for every period $t < t_1$, the group deliberates without voting. In period t_1, the group votes based on assessments $\boldsymbol{s}^{(t_1-1)} = A^{t-1}\boldsymbol{s}$, and if it reaches a verdict the process stops, otherwise it deliberates in that period and the process continues. It then deliberates in every period in the interval $t_1 < t < t_2$, and votes again in period t_2. Deliberation can continue until time t_F, at which point the group takes its final vote. If it fails to reach a verdict in period t_F, it declares itself deadlocked.

To see how a deliberation rule works, consider a jury trial in which σ is interpreted as the strength of the evidence and suppose that $T = \{5,10,20\}$ (i.e. $t_1 = 5$, $t_2 = 10$, and $t_3 = 20$). Under this rule, the jurors discuss the strength of the evidence without voting in rounds 1 through 4. They vote for the first time at the beginning of round 5, at which time their assessments of the evidence are given by $\boldsymbol{s}^4 = A^4\boldsymbol{s}$. If they reach a verdict, the trial ends. If not, they deliberate again in that period and in periods 6 through 9. They vote again at the beginning of period 10 according to their newest assessments, $\boldsymbol{s}^9 = A^9\boldsymbol{s}$. If they reach a verdict after voting in the tenth period, the trial ends. If not, they deliberate in that period and in periods 11 through 19. At the beginning of period 20 they take the final vote according to the assessments $\boldsymbol{s}^{19} = A^{19}\boldsymbol{s}$. If they reach a verdict, the trial ends with an acquittal or conviction. If they fail to reach a verdict, a hung jury is declared.

To begin presenting our results of the deliberation/voting process, the next proposition

establishes whether or not deliberation plays an important role in determining the outcome.

Proposition 9. The outcome of a deliberation/voting process is independent of the deliberation rule and the influence matrix if and only if all of the initial assessments are on the same side of the threshold σ^*.

According to Proposition 9, the *only* time the influence matrix and the deliberation rule do not matter for the outcome of the deliberation/voting process is when all members agree on the verdict without deliberation. For example, if, at the beginning of a meeting on whether or not to promote a junior faculty member, all members of the promotion committee enter the room thinking the candidate is unworthy of promotion, deliberation cannot change the fact that they will unanimously vote against promotion. Basically, when everyone is against promotion, no one is there to argue for promotion, so deliberation cannot matter. Furthermore, any verdict rule, unanimous or nonunanimous, will lead to the same verdict. So, unless the case is either so strong that all members agree at the outset to promote, or so weak that all members agree at the outset to dismiss, deliberation is a vital determinant of the outcome of the promotion meeting.

Even though a consensus on the value of σ is not required for a verdict, the consensus and the verdict are still linked, as shown by the next proposition.

Proposition 10. Suppose that a consensus s^* exists and that verdicts are unanimous. If $s^* > (<) \sigma^*$, and if the deliberation process lasts sufficiently long, the group reaches a "yes" ("no") verdict at some point in the deliberation process. Furthermore, there exists no deliberation rule that leads to

the opposite verdict.

Proposition 10 relates the binary variable corresponding to the verdict to a consensus about the continuous variable. If there is a consensus, which is a property of the influence matrix and therefore the group, and not of the value of σ itself, a verdict is reached and it is consistent with the consensus. So, for example, if half of the signals are above the threshold and half below, as long as the influence matrix is consistent with a consensus and deliberation continues long enough, a verdict is reached, and whether it is a "yes" or a "no" depends on whether the consensus assessment $s^* = \sum w_i s_i$ is above or below the threshold.

The members' consensus weights are important when the verdict is required to be unanimous. When a consensus exists, then according to Proposition 10 the verdict must match the consensus assessment. So, in particular, if a member is decisive any unanimous verdict must match his initial assessment. At the other end of the spectrum, if a member is given zero weight in the consensus the verdict is independent of his initial assessment. Consequently, when unanimity is required, member consensus weights matter for determining the verdict in the same way that they matter for determining the consensus assessment of the continuous variable.

The next proposition concerns the circumstances under which a deadlock occurs.

Proposition 11. If a deadlock occurs, there exist members i and j such that $s_i < \sigma^* \leq s_j$. Conversely, if the deliberation rule is given by $T = \{t_1, t_2, ..., t_F\}$, a deadlock occurs if at each $t \in T$ there exist members i and j such that $s_i^{(t-1)} < \sigma^* \leq s_j^{(t-1)}$.

As Proposition 11 shows, a *necessary* condition for a deadlock is for the group to contain two members whose initial assessments are on opposite sides of the threshold σ^*, so that they disagree on the issue at the outset. After all, deliberation moves the members' assessments closer together, and if all initial assessments are on the same side of the threshold deliberation cannot make them cross the threshold. The existence of two members whose initial assessments disagree is not a *sufficient* condition, though, because if the influence matrix meets the requirements for the existence of a consensus, the members will no longer disagree after sufficiently long deliberation. The sufficient condition for a deadlock is that in every period in which a vote is taken, two members disagree which side of the threshold the continuous variable lies. This might occur, for example, if the group would eventually reach a consensus but the deliberation process is too short, or if the group would not reach a consensus.

For juries in a criminal trial, deadlock corresponds to a hung jury. According to Proposition 11, if a jury hangs, it must be the case that at least one juror initially thought that the defendant was guilty and at least one initially thought the defendant was innocent. Conversely, to get a hung jury there must be (possibly different) jurors who disagree on the defendant's guilt at each point at which a vote is taken. This can occur even if the jury would eventually reach a consensus, as long as the final vote comes before everyone's assessment is on the same side of the reasonable doubt standard.

Deliberation rules can obviously differ in many ways. They can differ on the timing of the first vote, the timing of the last vote, and the timing of the votes in between. The next proposition concerns which features of the deliberation rule affect the outcome of the deliberation/voting process, and it is specific to unanimous voting rules.

Proposition 12. Assume that verdicts must be unanimous. Holding the influence matrix and the initial assessments of the members fixed, any two deliberation rules with the same timing of the final vote yield the same outcome.

Proposition 12 states that when verdicts are required to be unanimous, as with jury deliberations, the only component of the deliberation rule that matters for the outcome is the timing of the final vote.

According to Proposition 12, there are three reasons why one deliberation rule would lead to a verdict while the other leads to a deadlock: (i) the initial assessments are different under the two rules; (ii) the timing of the last vote is different under the two rules; and (iii) the influence matrix is different under the two rules. Because initial assessments are formed exogenously and prior to any meetings of the deliberative body, there is no reason to believe that they would depend on the deliberation rule, which is unknown until deliberation begins. This rules out reason (i). As for reason (ii), it may well be the case that group members give up sooner with some deliberation rules than with others, perhaps because when they vote often they do not see the vote changing, even though the assessments of the evidence are changing. Finally, with reason (iii) it may be the case that the influence matrix is different with different deliberation rules. For example, if members vote in the first period, they vote before reporting their assessments of the evidence, which could affect the influence matrix if members who voted yes only listen to other members who voted the same way. When the two factions do not listen to each other, a deadlock could ensue.

5.Applications

Unanimous vs. Nonunanimous Verdicts

The academic legal literature has devoted a fair amount of space to the issue of the deliberation process and nonunanimous jury verdicts.[17] Interestingly, identical reasons are often put forth to describe the costs *and* benefits of nonunanimous verdicts. Those in favor of nonunanimous verdicts stress their benefits–they speed up the deliberation process and reduce the number of hung juries by circumventing a small minority of hold-out jurors. Yet the authors of an experimental study that basically confirm these results (Hastie, *et al,* 1983), conclude that their findings support the continued use of a *unanimous* decision rule. Thus, there appears to be a presumption in the literature supporting unanimity that a lengthy and thoughtful deliberation process leads, in some sense, to more reliable verdicts, and so a trade-off exists between verdict "reliability" and deliberation costs.[18] Our model can help address this trade-off.

If a consensus exists, there is a unique value of the strength of the evidence that all twelve jurors can eventually be convinced is the true strength of the evidence. This is a way of aggregating the jurors' initial assessments. If a jury reaches a verdict, with a unanimous verdict rule the direction of the verdict is independent of the deliberation rule. Once all twelve jurors are on the same side of the reasonable doubt standard, the next vote taken leads to the same verdict if the next vote occurs

---

[17] For supporters of nonunanimous jury verdicts, see Amar (1995), Glasser (1997), Morehead (1998), and Rosen (1998). For critics, see Kachmar (1996), Osher (1996), Saks (1997), and Smith (1997).

[18] There are also constitutional issues involved with nonunanimous jury verdicts. The Supreme Court upheld the constitutionality of nonunanimous verdicts in two 1972 decisions. In *Johnson v. Louisiana* (92 S. Ct. 1620, 1972), the Court ruled that having a minority of three jurors voting to acquit does not violate the "proof beyond a reasonable doubt" standard the due process clause of the Fourteenth Amendment is interpreted as guaranteeing. In *Apodaca v. Oregon* (92 S. Ct. 1628, 1972), the Court ruled that a nonunanimous verdict does not violate the right to a trial by jury specified by the Sixth Amendment. For further discussion of the constitutional issues, see Abramson (1994, ch. 5).

immediately or many periods later. So, if deliberation is a good way to aggregate information, unanimity guarantees that the verdict will match the aggregated value. It may be in this sense that we can consider a unanimous verdict as reliable.

A nonunanimous verdict rule, on the other hand, can be quite sensitive to the deliberation rule. If a consensus exists, a nonunanimous rule may lead to a verdict that does not match the consensus. A jury that returns a verdict quickly may reach a different verdict than the same jury would reach if they deliberated without voting for a long time. Thus, if a consensus exists *and* a verdict is reached, one advantage of a unanimous rule over a nonunanimous one is that unanimity guarantees the verdict will match the consensus. The downside of unanimity, however, is that it will likely involve greater deliberation costs. So when a verdict can be reached, a trade-off exists between consensus-matching and deliberation costs. But what if a verdict cannot be reached?

With a unanimous rule, if a jury cannot agree on a verdict it is because of one of two reasons: (i) the deliberation rule does not allow for sufficiently long deliberation; or, (ii) the conditions for a hung jury stated in Proposition 11 exist, i.e. if there is always at least one juror on each side of the reasonable doubt standard, deliberation can never lead to a verdict. With a nonunanimous rule, both of these reasons play less of an impact. Thus, as is often argued, nonunanimous verdict rules can lower the probability of a hung jury and, therefore, lead to lower litigation costs when compared to a unanimous rule.

But a nonunanimous rule is not the only way to reduce the hung jury rate. From Proposition 12, with a unanimous verdict rule, deliberation rules with the same timing of the final vote yield the same outcome, be it a verdict or a hung jury. But to lower the probability of a hung jury, a corollary to that proposition is that if a judge enforces a sufficiently long deliberation, reason (i) for a hung

jury can practically be eliminated. Although it is largely common sense, the results of our model support the very common practice of judges sending deadlocked juries back to the deliberation room.

Our discussion of factions can also be used to shed some light on the issue of consensus and nonunanimous verdicts. We offer the following corollary to Proposition 10 (proof omitted):

Corollary to Proposition 10. Suppose that a verdict requires at least k votes and that there is a faction with at least k members. Let s^f be the faction consensus. If $s^f > (<) \sigma^*$, and if the deliberation process lasts sufficiently long, the group reaches a "yes" ("no") verdict at some point in the deliberation process.

Because every faction reaches a consensus, and since the faction is large enough to determine the outcome on its own, if sufficient time elapses before the first vote the group as a whole must reach the verdict determined by the faction. If the vote comes too early, though, some faction members might vote the opposite way, allowing the opposite verdict to arise. One difference, then, between a unanimous and a nonunanimous verdict rule is that if a consensus exists a unanimous rule guarantees that the verdict matches the consensus, while a nonunanimous rule does not even guarantee that the verdict matches the consensus of a controlling faction.

If a jury is hopelessly deadlocked, a nonunanimous rule may be the only way to circumvent the deadlock. Proponents of a nonunanimous rule often argue that such a rule can eliminate the "flake factor" (i.e. an irrational lone hold-out juror) from blocking a verdict. But this reasoning ignores the converse–the lone hold-out juror may have the correct assessment of the evidence. Critics of nonunanimous rules often highlight this last point by referring to the setting depicted in

the classic movie *12 Angry Men*.

In that movie, a vote was taken before any evidence was discussed, and Henry Fonda's character was the lone hold-out juror for acquittal. Had a nonunanimous rule been used, deliberation would have ended immediately with a conviction. But with a unanimous rule, as the evidence was continually discussed all twelve jurors eventually agreed to an acquittal. An overlooked point by those who use that example as an indictment against a nonunanimous rule, however, is that we never find out if Fonda's assessment of the evidence was in fact *correct*. Even his character states that he doesn't know if he is right, just that he has some doubts about the defendant's guilt.

In real world criminal cases we can rarely know for sure if a verdict is correct, or even if a consensus exists, but we do know that deliberation can move every jurors' assessment of the evidence to the same side of the reasonable doubt standard regardless of how many jurors initially disagree with each other. If the goal of a deliberation/vote process is to make a "good" decision, even if the definition of "good" is tenuous, one measure of the value of a decision is how many people agree with it, and the larger the majority favoring the decision the better.[19] After all, if everyone in a group votes the same way, it means that everyone in the group is convinced that the decision is the correct one. However, the process also has costs associated with it, the primary one being the value of the time spent deliberating.

The trade-off between the accuracy of the decision and the cost of making it can be captured by a two-component loss function, with one component increasing with the fraction of the group that disagrees with the decision, and the second component increasing with the time taken before a

[19] Note, however, that the number of members with assessments above the threshold implies nothing about how *far* the consensus assessment is above the threshold, and so the size of the majority does not necessarily imply that the evidence is stronger.

decision is reached. A designer can impact the expected value of the loss function in two ways, by setting the majority requirement for a decision and by setting the deliberation rule. The accuracy component of the loss function can be reduced by strengthening the majority requirement or by increasing the length of time before the first vote. The time cost component can be reduced by reducing the time before the first vote and by weakening the majority requirement. The existence of factions makes reaching a stronger majority requirement more difficult, and so the likelihood of group members forming factions should influence the designer's choice of majority requirements.

When designing voting rules for a jury in a criminal trial, because the jury pool is selected randomly from the population jurors are unlikely to belong to factions. Also, society places a relatively high weight on the quality of the decision, especially in order to avoid wrongful convictions. Finally, historically in the United States jury service has been regarded as a duty, and so relatively little weight has been placed on the value of jurors' time. Under these conditions, one would expect a criminal jury to face a strong majority requirement, and indeed most criminal juries face a unanimity requirement.

In contrast, consider promotion committees. Committee members' time is valuable, because if they are not in the meeting they can be engaging in the unit's productive core activities. Law partners, for example, could be billing their hours to clients instead of deliberating over the fate of a junior associate. Because the weight on the time cost component of the loss function is relatively high, one would expect low majority requirements with fairly short deliberations. In fact, most promotion committee meetings are reasonably short and the promotion decision is based on a simple majority vote.

In all, while consensus-matching and verdict accuracy can be related, they are not identical.

There is a growing literature on verdict rules and verdict accuracy. Feddersen and Pessendorfer (1998), with their strategic voting model, find that a unanimous rule is less accurate than a nonunanimous rule in term of wrongful verdicts. Coughlan (2000) refines the strategic voting model by allowing for hung juries and for minimal communication between jurors and finds a unanimous rule to be more accurate than a nonunanimous rule. Finally, Neilson and Winter (2002) demonstrate that if cases are continually retried until a verdict is reached, a unanimous rule tends to be more accurate than a nonunanimous rule. How these differences in the voting rules relate to the accuracy of the rules cannot be addressed by our deliberation model, just as current verdict accuracy models do not address deliberation dynamics. Future research may be able to combine these two lines of inquiry.

Jury Selection

One implication of our model adds a richness to the role of jury selection. The academic literature on jury selection largely focuses on peremptory challenges and broad juror biases pertaining to race and gender.[20] Since the Supreme Court's decision in *Batson v. Kentucky*, prosecutors and defense lawyers have been somewhat constrained in using peremptory challenges to eliminate jurors who are believed to have racial or gender biases toward defendants.[21] But even if the *Batson* decision has effectively limited the role of jury selection in some contexts, our model suggests that jury selection can consist of far more than simply trying to identify jurors who may either be pro- or anti-defendant. Also of great concern is how jurors are predicted to affect the

---

[20]For example, see Alschuler (1989), Underwood (1992), and Neilson and Winter (2000).

[21]*Batson v. Kentucky*, 106 S. Ct. 1712 (1986).

deliberation process, and we allow for these particular character traits through the consensus weights introduced in section 2. As an example of an implication of our model, we can predict that a defense attorney may be less concerned with striking a juror who is predicted to be severely anti-defendant but likely to have little weight in the deliberation process, compared to a juror who is predicted to be only slightly anti-defendant but apt to have a large weight during deliberation.

While the likelihood of being able to identify consensus weights during *voir dire* and the empirical validity of their role during deliberation remain open questions, the jury consultant profession has relied on their claim of being able to predict how different juror personality types may interact. For example, consider the flamboyant jury selection strategy known as the "poison pill" strategy:[22]

> [A jury consultant] deliberately picked jurors who would explode, who would hate each other. That's what you do in a criminal case when it is obvious that people are guilty. You go for personalities. Then, you hope the personalities will combust.

To the extent that a defense attorney may be interested in ending with a hung jury, this may be achieved by trying to identify and select two jurors who will satisfy the hung jury conditions stated in Proposition 11. Thus, our model is sufficiently general that it can even account for (but not be blamed for) the "poison pill" strategy.

---

[22]Strier and Shestowsky (1999) p. 444.

6. Concluding Comment

This paper presents a formal model of a deliberation process that ends with an up-down vote, and adds several contributions to the existing literature. It makes a modeling contribution by linking the existing model of consensus, which has a continuous decision variable, to a deliberation model, which has a binary decision variable. It also formally defines a deliberation rule, which is a crucial component of the model. Furthermore, because disagreement is an important outcome in many deliberation processes, it adds to the consensus model by not just identifying factors that make a group agree, but also factors that make one disagree.

Besides making a modeling contribution, the paper also adds to our understanding of deliberation/vote processes. It looks at when deliberation determines the outcome, and finds that deliberation is an important determinant of the outcome whenever group members disagree before the deliberation starts. This, of course, is the only interesting scenario and the one that all analyses of jury behavior consider, for example. It is shown that when a consensus exists, unanimous verdicts must match the consensus assessment of the variable under consideration, so deliberation aggregates the members' assessments, and the members who have large weights in the deliberation process are most significant in the determination of the outcome of the vote. The paper also identifies what must occur for the group to fail to reach a verdict and what factors in the deliberation process matter. For unanimous verdicts the only considerations that affect the outcome are the members' initial assessments, the manner in which they influence each other, and the timing of the final vote. Finally, the paper allows us to determine the impact of relaxing the unanimity requirement. Doing so removes the link between the consensus assessment and the outcome of the vote, and the verdict may disagree with the consensus assessment. Consequently, the information aggregation properties of

consensus are lost. Furthermore, the timing of early votes becomes important, where only the timing of the final vote matters when verdicts must be unanimous. Because of this, relaxing the unanimity requirement makes the process by which the group reaches a verdict much more critical.

Appendix

Proof of Proposition 1. Suppose, without loss of generality, that member 1 influences everybody. Consider any member $i \neq 1$. Since i is influenced by 1, there exists an integer $1 \leq k \leq n$ such that $a_{i1}^{(k)} > 0$, reflecting that member i's k-th period assessment incorporates member 1's original signal. Furthermore, $a_{i1}^{(m)} > 0$ for all $m > k$ because $a_{ii} > 0$; that is, in period k member i is influenced by member 1, and in period $m > k$ member i is influenced by himself, and his previous assessments were influenced by member 1, so the m-th assessment is, too. Since this is true for all $i \neq 1$, and since $a_{11} > 0$, the first column of A^n is strictly positive. By Theorem 1 in DeGroot (1974), the matrix A^∞ consists of n identical rows. Letting $\boldsymbol{b}$ denote this common row vector A^∞, it follows that $s_1^* = \ldots = s_n^* = \boldsymbol{b} \cdot \boldsymbol{s}$, and a consensus is reached.

Proof of Proposition 2. By the assumptions made on the matrix A, $a_{ij} \in [0,1]$ for all i,j, $a_{ii} > 0$ for all i, and $\sum_j a_{ij} = 1$ for all i. Consequently, A is an aperiodic Markov transition matrix. Suppose that member i influences everyone. Let $I = \{j | j \text{ influences } i\}$. Note that $i \in I$. First consider the case where $I = \{1,\ldots,n\}$, so that juror i is also influenced by everyone. Then the matrix A is ergodic, and by Theorem 8.4.12 in Berman and Plemmons (1979), the matrix A^∞ can be written

$$A^\infty = \begin{bmatrix} \boldsymbol{w} \\ \vdots \\ \boldsymbol{w} \end{bmatrix},$$

where $\boldsymbol{w}$ is the unique stationary probability distribution (row) vector of the Markov chain, and every element of $\boldsymbol{w}$ is positive.

Now consider the case where $I \neq \{1,\ldots,n\}$. Without loss of generality, assume that $I = \{1,\ldots,k\}$ for some $k < n$. Then the influence matrix A takes the form

$$A = \begin{bmatrix} B & 0_k \\ C & D \end{bmatrix},$$

where B, C, and D are submatrices, and 0_k denotes a k x k matrix of zeroes. By assumption, B is ergodic. The matrix A^t takes the form

$$A^t = \begin{bmatrix} B^t & 0_k \\ X_t & D^t \end{bmatrix},$$

where X_t is a matrix. Taking limits,

$$\lim_{t\to\infty} A^t = A^\infty = \begin{bmatrix} B^\infty & 0_k \\ X_\infty & D^\infty \end{bmatrix}.$$

Since B is an ergodic Markov transition matrix, there exists a unique stationary probability distribution (row) vector $\boldsymbol{v}$ such that

$$B^\infty = \begin{bmatrix} \boldsymbol{v} \\ \vdots \\ \boldsymbol{v} \end{bmatrix}.$$

Note that since the elements of D are all nonnegative and the rows sum to less than one, $D^\infty = 0_{n-k}$.

Finally, it remains to show that each row of X_∞ is equal to $\boldsymbol{v}$. Suppose not, so that some row of $X_\infty = \boldsymbol{y} \neq \boldsymbol{v}$. Then, in particular, there exists an integer $j \in \{1,...,k\}$ such that $y_j \neq v_j$. Choose the vector $\boldsymbol{s}$ so that $s_i = 0$ for $i \neq j$ and $s_j > 0$. Then individual i's final assessment is $s_i^*(\boldsymbol{s}) = v_j s_j$ when $i \in \{1,...,k\}$, but it is $s_m^*(\boldsymbol{s}) = y_j s_j$ for some individual $m \in \{k+1,...,n\}$. But, by Proposition 1, a consensus is reached, so $s_i^*(\boldsymbol{s}) = s_m^*(\boldsymbol{s})$, which provides a contradiction. Finally, to prove the proposition, let $\boldsymbol{w}$ be the n-vector given by $\boldsymbol{w} = (\boldsymbol{v}, 0,...,0)$.

Proof of Proposition 3. By Proposition 1, since everyone is influenced by member i, a consensus is reached. Since member i is not directly influenced by any other members, $a_{ii} = 1$ and $a_{ij} = 0$ for $j \neq i$. Note that $a_{ij}^{(2)} = \sum_k a_{ik}a_{ki}$, so $a_{ij}^{(2)} = 0$ for $j \neq i$ and $a_{ii}^{(2)} = 1$. By induction, $a_{ij}^{(t)} = 0$ for $j \neq i$ and $a_{ii}^{(t)} = 1$ for $t = 1,2,...$ Taking the limit as $t \to \infty$ yields $w_i = 1$ and $w_j = 0$ for $j \neq i$.

Proof of Proposition 4. By Proposition 1, there is a consensus. By Proposition 2, there exist weights $w_1,...,w_n$ such that $s^* = \boldsymbol{w}\cdot\boldsymbol{s}$. We wish to show that $w_j = 0$. Suppose instead that $w_j > 0$. Then since $w_j = \lim_{t\to\infty} a_{ij}^{(t)}$, there exists and integer $T > 0$ such that for all $\tau > T$, $a_{ij}^{(\tau)} > 0$. But then member i is influenced by member j in round τ, which is a contradiction.

Proof of Proposition 5. By Proposition 1, it is enough to show that if at least two members are not influenced by the same individual, the group disagrees. Suppose not, that is, suppose that a consensus is reached. Assume without loss of generality that members 1 and 2 are not influenced by the same individual. Partition the group of n individuals into $k \leq n$ subgroups such that no member of subgroup i influences any member of subgroup j for all $i \neq j$, and such that each subgroup contains at least one member who influences everyone in the subgroup. Since members 1 and 2 are not influenced by the same individual they cannot be in the same subgroup, and so $k \geq 2$. By Proposition 1, each subgroup reaches a consensus. One can construct k weight vectors, $\boldsymbol{w}^1,...,\boldsymbol{w}^k$, corresponding to each subgroup, with $w_m^j = 0$ when individual m is not a member of subgroup j, and $w_m^j > 0$ for at least one member of subgroup j. However, since a consensus is reached, by Proposition 2, $\boldsymbol{w}^i = \boldsymbol{w}^j$ for all i,j. But this is impossible because $w_m^j = 0$ for every member

m of subgroup i, and $w_m^{\,i} > 0$ for at least one member of subgroup i. This is a contradiction, and thus the group disagrees.

Proof of Proposition 6. Suppose, to the contrary, that i is a member of two different factions, call them C_1 and C_2. Let j_1 be a member of C_1 but not C_2 and j_2 be a member of C_2 but not C_1. Because i is influenced by j_1, and because i influences j_2, j_1 influences j_2. Then C_2 cannot be a faction because it has a member who is influenced by someone outside of C_2. This provides a contradiction.

Proof of Proposition 7. The proof proceeds by induction. Suppose that i is the only member of the fringe and that i is not influenced by any member of a faction. Then i is not influenced by any other group member and the single-member subgroup $\{i\}$ is a faction, providing a contradiction. Now suppose that when the fringe has m members, each member of the fringe is influenced by some member of a faction. When the fringe has $m + 1$ members, and i is a member of the fringe, i must be influenced by some other member of the fringe or else $\{i\}$ is a faction. Suppose that i is influenced by fringe member k. Construct the influence matrix A_{-i} by removing the i-th row and column from A and renormalizing by dividing every element of row $j \neq i$ by $1 - a_{ij}$ so that the rows sum to one. Then the influence matrix A_{-i} has a fringe with m members, and therefore each member of the fringe is influenced by some member of some faction. Consequently, with the original influence matrix A, member k is influenced by some member of a faction, and therefore i is influenced by some member of some faction.

Proof of Proposition 8. For (i), if there is a single faction, then by the definition of a faction and by Proposition 7, everyone is influenced by every member of the faction. By Proposition 1, a consensus exists. Conversely, suppose that the group has at least two factions. Then when individuals i and j are in different factions they are not influenced by the same individual, and by Proposition 5 the group disagrees.

For (ii), if everyone but i is in the fringe, then there is only one faction $\{i\}$. Therefore the group reaches a consensus, and by Proposition 3, since i is not influenced by any other member, i is decisive.

For (iii), if there is a consensus then there is only one faction. Consequently, everyone is influenced by some member of the faction, say member i. As a member of the fringe, j has no influence on i, and by Proposition 4 the consensus is independent of j's signal.

Proof of Proposition 9. Suppose that all initial assessments are on the same side of the reasonable doubt standard. If $s_i < \sigma^*$ for all i, then, since updated assessments are convex combinations of initial assessments, $s_i^{(t)} < \sigma^*$ for all i and all t. Consequently, for all influence matrices and all deliberation rules, the jurors all vote to acquit on the first vote. The reasoning is similar if $s_i \geq \sigma^*$ for all i.

Suppose that the outcome of a trial is independent of the deliberation rule and the influence matrix, but, contrary to the statement of the proposition, not all of the initial assessments are on the same side of the reasonable doubt standard. Then there exist jurors i and j such that $s_i < \sigma^* \leq s_j$, and let m be the number of jurors whose initial assessments are below the reasonable doubt standard. Assume, without loss of generality, that more than half of the jurors have initial assessments below the reasonable doubt standard. Let A be an influence matrix in which juror i is decisive, and let B be an influence matrix in which juror j is decisive. Let t_A be the smallest positive integral value of

t for which each element of $A^t\boldsymbol{s}$ is below the reasonable doubt standard, and let t_B be the smallest positive integral value of t for which each element of $B^t\boldsymbol{s}$ is at or above the reasonable doubt standard. Both t_A and t_B are strictly greater than one. If a verdict requires m or fewer votes and the deliberation rule specifies a vote at time $t = 1$, the outcome of the trial is an acquittal. If a verdict requires m or fewer votes, the influence matrix is B, and the deliberation rule is such that $t_1 \geq t_B + 1$, the trial results in a conviction. If a verdict requires $m + 1$ or more votes, the influence matrix is A, and the deliberation rule is such that $t_F = 1$, the trial results in a hung jury. If a verdict requires $m + 1$ or more votes, the influence matrix is A, and the deliberation rule is such that $t_1 \geq t_A + 1$, the trial results in an acquittal. If a verdict requires $m + 1$ or more votes, the influence matrix is B, and the deliberation rule specifies $t_1 \geq t_B + 1$, the trial results in a conviction. The existence of these possibilities contradicts the hypothesis that the outcome is independent of the influence matrix and the deliberation rule.

Proof of Proposition 10. Suppose first that $s^* > \sigma^*$. Since s^* is in the convex hull of the set $\{s_1^{(t)},...,s_n^{(t)}\}$ for all t, it must be the case that $s_i^{(t)} > \sigma^*$ for some i at every stage t of the deliberation. Consequently, at every $t \in T$, some member of the jury votes "guilty," and therefore there cannot be a unanimous "not guilty" verdict. Also, since $s_i^{(t)}$ converges to s^* for all i, there exists an integer $\tau > 0$ such that for every integer $t > \tau$, $s_i^{(t)} > \sigma^*$ for all i. Construct $T = \{\tau+2\}$, and under this deliberation rule the verdict is "guilty" on the first vote. The analysis of the case in which $s^* < \sigma^*$ is similar.

Proof of Proposition 11. Regarding the first statement, if $s_i < \sigma^*$ for all i, the jury reaches a "not guilty" verdict the first time a vote is taken, and if $s_i \geq \sigma^*$ for all i, the jury reaches a "guilty" verdict the first time a vote is taken. The second statement is obvious.

Proof of Proposition 12. Let T^J and T^K be two deliberation rules, with $t_F^J = t_F^K = t_F$. Since the influence matrix A and the vector of initial assessments $\boldsymbol{s}$ are the same under both rules, the time-t assessment $\boldsymbol{s}^{(t)} = A^t\boldsymbol{s}$ is the same under both deliberation rules. Now, suppose that deliberation rule J leads to a verdict, and assume that it is an acquittal. Then at some time t^* such that $t^* + 1 \in T^J$, $s_i^{(t^*)} < \sigma^*$ for all i. Since $s_i^{(t^*)}$ is a convex combination of $s_1^{(\tau)},...,s_n^{(\tau)}$ for all $\tau = 1,...,t^* - 1$, it must be the case that $s_j^{(\tau)} < \sigma^*$ for some j and some $\tau = 1,...,t^* - 1$. So, under the other deliberation rule, rule K, the jury could not have agreed to convict before time t^*. Also, under rule K the jury reaches a "not guilty" verdict at time $t' = \min\ \{t \in T^K \mid t \geq t^* + 1\}$, if it has not already reached one. The analysis for the case where rule J leads to a conviction is similar, so, if a verdict is reached, the two deliberation rules result in the same verdict.

It remains to show that if one deliberation rule results in a hung jury, so does the other one. Suppose that rule J results in a hung jury. Then it must be the case that for some jurors i and j,

$$s_i^{(t_F-1)} < \sigma^* \leq s_j^{(t_F-1)},$$

so that juror i votes to acquit in the final vote and juror j votes to convict. Since $s_k^{(t_F-1)}$ is a convex combination of $s_1^{(\tau)},...,s_n^{(\tau)}$ for all $\tau = 1,...,t_F - 2$, it follows that an inequality like the one above must

hold for some jurors i and j in every prior period. Consequently, there is no period in which the jurors would have reached a unanimous verdict, and the jury hangs under deliberation rule K as well.

References

Abramson, J., 1994. We the Jury. Basic Books: New York.

Alschuler, A.W., 1989. The supreme court and the jury: Voir dire, peremptory challenges, and the review of jury verdicts. University of Chicago Law Review 56, 153-233.

Amar, A., 1995. Reinventing juries: Ten suggested reforms. U.C. Davis Law Review 28, 1169-1194.

Aumann, R. J,. Hart, S., 2003. Long cheap talk. Econometrica 71, 1619-1660.

Berman, A., Plemons, R.J., 1979. Nonnegative Matrices in the Mathematical Sciences. Academic Press: New York.

Conlisk, J., Gong, J.C., Tong, C.H., 2001. Actions influenced by a social network. Journal of Evolutionary Economics 11, 277–305.

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

Crawford, V. P., Sobel, J., 1982. Strategic information transmission. Econometrica 50, 1431-1451.

DeMarzo, P.M., Vayanos, D., Zwiebel, J., 2003. Persuasion bias, social influence, and unidimensional opinions. Quarterly Journal of Economics 118, 909-968.

DeGroot, M., 1974. Reaching a consensus. Journal of the American Statistical Association 69, 118-121.

Feddersen, T.J., Pesendorfer, W., 1998. Convicting the innocent: The inferiority of unanimous jury verdicts. American Political Science Review 92, 23-35.

Feller, W., 1968. An Introduction to Probability Theory and Its Applications, Volume 1. John Wiley & Sons: New York.

Friedkin, N.E., 2001. Norm formation in social influence networks. Social Networks 23, 167–189.

Garicano, L., Posner, R.A., 2005. Intelligence failures: An organizational economics perspective. Journal of Economic Perspectives 19, 151-170.

Glasser, M.H.,1997. Letting the supermajority rule: Nonunanimous jury verdicts in criminal trials. Florida State University Law Review 24, 659-677.

Glazer, J., Rubinstein, A., 2001. Debates and decisions: On a rationale of argumentation rules. Games and Economic Behavior 36, 158-173.

Glazer, J., Rubinstein, A., 2004. On optimal rules of persuasion. Econometrica 72, 1715-1736.

Hastie, R., Penrod, S.D., Pennington, N., 1983. Inside the Jury. Harvard University Press: Cambridge.

Kachmar, J., 1996. Silencing the minority: Permitting nonunanimous jury verdicts in criminal trials. Pacific Law Journal 28, 273-310.

Klevorick, A.K., Rothschild, M., 1979. A model of the jury decision process. Journal of Legal Studies 8, 141-164.

Li, H., Rosen, S., Suen, W., 2001. Conflicts and common interests in committees. American Economic Review 91, 1478-1497.

Morehead, J.W., 1998. A 'modest' proposal for jury reform: The elimination of required unanimous jury verdicts. University of Kansas Law Review 46, 933-945.

Neilson, W.S., Winter, H., 2005. The elimination of hung juries: Retrials and nonunanimous verdicts. International Review of Law and Economics 25, 1-19.

Neilson, W.S. Winter, H., 2000. Bias and the economics of jury selection. International Review of Law and Economics 20, 223-250.

Osher, J., 1996. Jury unanimity in California: Should it stay or should it go? Loyola of Los Angeles Law Review 29, 1319-1370.

Razavi, H., 1984. An economic model of OPEC coalition. Southern Economic Journal 51, 419-428.

Rosen, J., 1998. After one angry woman. University of Chicago Legal Forum 1997, 179-195.

Saari, D.G., Sieberg, K.K., 2001. Some surprising properties of power indices. Games and Economic Behavior 36, 241-263.

Saks, M.J., 1997. What do jury experiments tell us about how juries (should) make decisions? Southern California Interdisciplinary Law Journal 6, 1-53.

Smith, D.G., 1997. Structural and functional aspects of the jury: Comparative analysis and proposals for reform. Alabama Law Review 48, 441-581.

Straffin, P.D. Jr., 1994. Power and stability in politics. In: Aumann, R.J., Hart, S., (Eds), Handbook of Game Theory, Volume 2. Elsevier Science B.V.: Amsterdam.

Strier, F., Shestowsky, D., 1999. Profiling the profilers: A study of the trial consulting profession, its impact on trial justice and what, if anything, to do about it. Wisconsin Law Review 1999,

441-499.

Underwood, B.D., 1992. Ending race discrimination in jury selection: Whose right is it anyway? Columbia Law Review 92, 725-774.