Consistency and Communication in Committees

Inga Deimen*   Felix Ketelaar   Mark T. Le Quement

Institute for Microeconomics, University of Bonn, Germany.

Abstract

We generalize the classical binary Condorcet jury model by introducing a richer state and signal space, thereby generating a concern for consistency in the evaluation of aggregate information. We analyze truthtelling incentives in simultaneous pre-vote communication in heterogeneous committees and find that full pooling of information followed by sincere voting is compatible with a positive probability of ex post conflict in the committee.

Keywords: Committees, Voting, Information aggregation, Cheap talk
JEL Classification: D72, D82, D83

1   Introduction

This paper considers a deliberation and voting model in which rich state and signal spaces combine with a binary action space. A committee consisting of privately informed agents with known heterogeneous preferences engages in simultaneous information exchange prior to voting. Our information structure generates a concern for consistency in the aggregation of individual signals; a given signal is interpreted differently depending on how it matches other available evidence. We find that in contrast to the classical model featuring binary state and signal spaces, full information sharing and sincere voting can constitute an equilibrium although agents with some probability disagree ex post.

Consider the example of a jury aiming at determining whether a defendant is guilty or innocent. If guilty, he must have committed the crime at one specific point in time, for example on one particular day of a given week. If innocent, he must have been engaged in some activity at the moment of the crime; for example working, watching TV, or doing sports. Different days of the week constitute mutually exclusive variants of the guilty state while different activities constitute mutually exclusive variants of the innocent state.

Jurors gather evidence through a trial hearing which generates private signals. Before deciding whether to acquit or convict, jurors retire to deliberate and share their private signals. Consider two possible scenarios. In the first scenario, half of the jurors received a signal indicating that the defendant committed the crime on Monday, while the other half received a signal indicating that he committed the crime on Wednesday. In the second scenario, all jurors

---

received a signal indicating Monday. The latter scenario is more consistent than the first and therefore provides more convincing evidence of guilt. Jurors do not as such care about the time at which the crime was committed but wish to establish with sufficient certainty whether the defendant is guilty or innocent. More consistent profiles yield stronger evidence.

The core elements of the above description apply to many other situations. Consider a group of investment bankers that contemplates investing in shares of a large manufacturer, e.g. Chrysler. Committee members need to assess whether Chrysler will avoid bankruptcy in the near future. This may happen if either the US Federal State provides a bailout package or if some private company (e.g. Fiat) decides to step in. On the other hand, if Chrysler does go bankrupt, this may happen according to different chapters of the bankruptcy code. Another example is that of a board of directors that seeks to predict whether a Democrat or a Republican will win the next US presidential election. Different Democratic (Republican) candidates constitute different variants of the Democratic (Republican) state.

We incorporate the key features of the above examples into a model of collective decision making. There are two basic states, each of which splits into a set of substates. Each signal is informative with respect to a basic state and a particular substate. In this context, the consistency of signals matters as illustrated above; more consistent signals provide stronger evidence for the corresponding state. Members of a heterogeneous committee communicate via cheap talk before voting on a binary outcome. In contrast to results obtained in the classical binary signal setup,[1] we find that the truthful communication and sincere voting equilibrium (TS equilibrium) is virtually always compatible with a positive probability of ex post conflict among agents.

The intuition for our result comes out clearly when compared to the classical binary signal model.[2] In the latter, in the putative TS equilibrium, pivotality at the communication stage pins down uniquely the information held by the remaining committee members. Disagreement about the optimal decision rule implies that there is always at least one agent for whom this pivotal profile implies a suboptimal decision on the equilibrium path. Consequently, this agent profitably deviates and bends the decision rule in his favored direction.

In our model, the set of pivotal profiles is not a singleton anymore: different signal profiles can yield similar posteriors because the conditional probability of guilt depends on two aspects, the total number of signals indicating respectively guilt or innocence as well as the consistency of signals within these subsets. A smaller total number of guilty signals can be compensated by a higher degree of consistency among guilty signals. The multiplicity of pivotal scenarios in our model allows two effects to come into play. First, there exist pivotal profiles for which all agents agree with the decision taken on the equilibrium path. At these profiles a deviation is disadvantageous (*consensus effect*). Second, the impact of an agent's announcement depends on the signals of the whole committee. He faces uncertainty as to which announcement is more or less consistent with other agents' signals and is thus suitable to shift the outcome in a desired direction (*uncertainty effect*).

Building on the theory of strategic voting as information aggregation (E.g. Austen-Smith and Banks (1996), Feddersen and Pesendorfer (1998), Bhattacharya (2013)), different classes of contributions have analyzed communication in heterogeneous committees by modifying the baseline model (Coughlan (2000)). Austen-Smith and Feddersen (2006), Meirowitz (2007), and

---

[1]See Coughlan (2000), Austen-Smith and Feddersen (2006), Meirowitz (2007), Van Weelden (2008).

[2]See in particular Coughlan (2000).

Le Quement (2013) examine the implications of preference uncertainty. Van Weelden (2008), Hummel (2012), and Le Quement and Yokeeswaran (2015) analyze alternative communication protocols. Gerardi et al. (2009), Gerardi and Yariv (2007), and Wolinsky (2002) adopt a mechanism design approach. Additionally, experimental work has analyzed communication behavior in groups (E.g. Goeree and Yariv (2011) and Dickson et al. (2008)). Our contribution lies in the introduction of a novel information structure that captures the idea that individual information is interpreted within the context of aggregate information.

2   The model

A jury of n agents, $n \in \mathbb{N}$, $n \geq 3$ has to decide whether to acquit (A) or convict (C) a defendant. The defendant is either innocent (I) or guilty (G). Both innocence and guilt occur in finitely many different variants $i_1, ..., i_{m_I}$ and $g_1, ..., g_{m_G}$, respectively, with $m_I, m_G \in \mathbb{N}$, $m_I, m_G \geq 2$.[3] The state space is hence given as $\Omega = I \cup G$ where $I = \{i_1, ..., i_{m_I}\}$ and $G = \{g_1, ..., g_{m_G}\}$. We denote the state by ω and say that the defendant is innocent if $\omega \in I$ and guilty if $\omega \in G$. The state is drawn from a publicly known prior distribution such that $P(\omega \in I) + P(\omega \in G) = 1$, $P(i_l) = \frac{P(\omega \in I)}{m_I} \; \forall l \in \{1, ..., m_I\}$, and $P(g_l) = \frac{P(\omega \in G)}{m_G} \; \forall l \in \{1, ..., m_G\}$.

The jury implements an action $a \in \{A, C\}$ by voting according to a voting rule $k \in \{1, ..., n\}$. Each agent $j \in \{1, ..., n\}$ casts a vote in favor of one of the two actions. If the number of votes cast for conviction is greater than or equal to k, the defendant is convicted while otherwise he is acquitted.

The utility of agent j from action a conditional on $\omega \in \Omega$ is given by

$$u_j(a, \omega) = \begin{cases} 0 & \text{if } (a, \omega) \in \{(A, I)\,(C, G)\} \\ -q_j & \text{if } (a, \omega) = (C, I) \\ -(1 - q_j) & \text{if } (a, \omega) = (A, G). \end{cases}$$

The commonly known preference parameter $q_j \in (0, 1)$ characterizes the relative importance assigned to the two types of errors.[4] As agent j maximizes expected utility, he prefers conviction over acquittal if and only if the probability of the defendant being guilty exceeds the cut-off q_j. We mainly focus on the case of a committee featuring two preference types $q_H < q_D$ referred to as hawks and doves.

Prior to the voting stage, each agent receives a private signal $s \in S = \Omega$. Signals are i.i.d. across agents conditional on ω: If $\omega = i_l$ for some $l \in \{1, ..., m_I\}$, then

$$\begin{aligned} P(s = i_l | \omega = i_l) &= \lambda \cdot \frac{p}{\lambda + (m_I - 1)}, \\ P(s = i_r | \omega = i_l) &= \frac{p}{\lambda + (m_I - 1)} \quad \forall r \in \{1, ..., m_I\},\, r \neq l, \\ P(s = g_r | \omega = i_l) &= \frac{1 - p}{m_G} \quad \forall r \in \{1, ..., m_G\}, \end{aligned}$$

---

[3] The case of $m_I = m_G = 1$ corresponds to the classical model analyzed in Coughlan (2000) and others.

[4] Existing results (Austen-Smith and Feddersen (2006)) indicate that truthful communication is easier to achieve if there is uncertainty about preference types. By assuming observable preference types, we isolate the specific truthtelling incentives that are inherent to our model.

with $p \in \left(\frac{1}{2}, 1\right)$ and $\lambda > 1$. If $\omega = g_l$ for some $l \in \{1, ..., m_G\}$, respective expressions apply after permuting i and g as well as I and G.[5] Applying Bayes' law, p measures up to priors the probability that the signal correctly reveals whether the defendant is innocent or guilty. $\lambda = \frac{P(s=i_l|\omega=i_l)}{P(s=i_l|\omega=i_r)} = \frac{P(s=g_l|\omega=g_l)}{P(s=g_l|\omega=g_r)}$ measures the relative informativeness of signals with respect to the particular variant of innocence or guilt. We finally assume that

$$\frac{P(s=i_l|\omega=i_r)}{P(s=i_l|\omega=g_t)} \geq 1 \quad \forall l, r \in \{1, ..., m_I\}, \, t \in \{1, ..., m_G\} \quad \text{and}$$

$$\frac{P(s=g_l|\omega=g_r)}{P(s=g_l|\omega=i_t)} \geq 1 \quad \forall l, r \in \{1, ..., m_G\}, \, t \in \{1, ..., m_I\},$$

which is equivalent to requiring a lower bound $p \geq \max\left\{\frac{m_G-1+\lambda}{2m_G-1+\lambda}, \frac{m_I-1+\lambda}{2m_I-1+\lambda}\right\}$ on the informativeness of signals with respect to innocence and guilt.

A collection of signals constitutes a *signal profile* $\sigma = (x_1, ..., x_{m_I}, y_1, ..., y_{m_G})$ where x_r denotes the number of i_r-signals, $r \in \{1, ..., m_I\}$ and y_l denotes the number of g_l-signals, $l \in \{1, ..., m_G\}$. Conditional on signal profile σ, the posterior probability of guilt is given by

$$\beta(\sigma) \equiv P(\omega \in G|\sigma) = \frac{P(\omega \in G) \cdot P(\sigma|\omega \in G)}{P(\omega \in G) \cdot P(\sigma|\omega \in G) + P(\omega \in I) \cdot P(\sigma|\omega \in I)}.$$

In terms of utilities agents only care whether $\omega \in I$ or $\omega \in G$, hence the number $\beta(\sigma)$ is a sufficient statistic for the preferred action of each individual agent for any signal profile σ.

The timing of the game is as follows: Nature draws ω from f. Each agent receives a private signal s. Subsequently, agents simultaneously send a public cheap talk message $m \in M = S$.[6] Finally, each agent casts a vote, the action $a \in \{A, C\}$ is implemented according to the voting rule, and payoffs are realized.

Our equilibrium concept is Perfect Bayesian Equilibrium. We focus on the so-called *TS equilibrium*: agents *truthfully* reveal their private information by sending a message $m = s$ at the communication stage and (correctly) believe that others communicate truthfully as well. Subsequently, agents vote *sincerely* by voting for conviction if and only if $\beta(\sigma) \geq q_j$.

3 A simple example

Consider a three persons committee consisting of two doves and one hawk. The voting rule is $k = 2$, i.e. simple majority. Let $I = \{i_1, i_2\}$ and $G = \{g_1, g_2\}$. Referring to the example given in the introduction, think of i_1 as "innocent and working", i_2 as "innocent and doing sports", g_1 as "guilty on Monday", and g_2 as "guilty on Wednesday". Aggregate signal profiles are ordered as follows

$$\begin{matrix} \beta(3,0,0,0) \\ \beta(0,3,0,0) \end{matrix} < \cdots < \begin{matrix} \beta(1,0,0,2), \beta(1,0,2,0) \\ \beta(0,1,0,2), \beta(0,1,2,0) \end{matrix} < \begin{matrix} \beta(0,0,1,2) \\ \beta(0,0,2,1) \end{matrix} < \begin{matrix} \beta(0,0,0,3) \\ \beta(0,0,3,0) \end{matrix}.$$

Suppose q_H, q_D are such that a dove favors conviction if and only if the aggregate signal profile is either $(0,0,0,3)$ or $(0,0,3,0)$ while a hawk in addition favors conviction for profiles $(0,0,1,2)$

---

[5]Note that whenever $\lambda > 1$ our signal generating process is not reducible to a process that generates i.i.d. signals conditional on I and G. It is, however, reducible to a process that generates correlated signals conditional on I and G.

[6]See end of Section 3 for a comment on sequential communication.

and $(0, 0, 2, 1)$, so that the group disagrees for these two profiles. While both groups require three g-signals to prefer conviction, doves furthermore require these g-signals to be consistent. In this setting the TS equilibrium exists, as we now show.

An agent never has an incentive to deviate from sincere voting as this weakly decreases the probability that his favored decision ensues. This holds true independently of his announcement at the communication stage. Moreover, given that the voting rule is simple majority doves can always enforce their favored decision and thus have no incentive to deviate from truthtelling.

We now analyze the truthtelling incentives of the hawk. The hawk's announcement is pivotal if the remaining two agents hold signal profiles $(0, 0, 2, 0)$ or $(0, 0, 0, 2)$. In the first (second) case, a g_1- (g_2-) announcement triggers conviction while any of the remaining announcements causes acquittal. Given the symmetry of the model, conditions ensuring truthtelling when the hawk holds an i_1- or an i_2-signal are identical modulo an exchange of subscripts. We can therefore without loss focus on deviation incentives conditional on an i_1-signal. An equivalent argument applies when the hawk holds a g_1- or a g_2-signal.

Let the hawk hold an i_1-signal and be pivotal at the communication stage. The aggregate signal profile is then $(1, 0, 2, 0)$ or $(1, 0, 0, 2)$. In either case, the committee decision given the true signal profile is acquittal and coincides with the decision favored by the hawk. Accordingly, he has no incentive to deviate from truthtelling. Here, the consensus effect is the source of truthtelling: a hawk holding an i_1-signal fully agrees with the doves on the preferred action in all pivotal scenarios.

Assume next that the hawk holds a g_1-signal and is pivotal in the communication stage. The signal profile of the entire committee is then either $(0, 0, 1, 2)$ or $(0, 0, 3, 0)$. The hawk disagrees with the acquittal ensuing from truthtelling at $(0, 0, 1, 2)$ while he agrees with the conviction ensuing from truthtelling at $(0, 0, 3, 0)$. If the hawk deviates to announcing some i-signal, the signal profile observed by others at the voting stage is either $(1, 0, 2, 0)$, $(1, 0, 0, 2)$, $(0, 1, 0, 2)$ or $(0, 1, 2, 0)$, in all cases leading to an undesired acquittal. If the hawk deviates to a g_2-announcement, the signal profile observed by the remaining agents at the voting stage is $(0, 0, 0, 3)$ or $(0, 0, 2, 1)$. The deviation beneficially overturns an acquittal in the first case but adversely overturns a conviction in the second case. The hawk thus faces uncertainty about the impact of his statement. Among the two pivotal profiles, $(0, 0, 0, 2)$ incentivizes lying while $(0, 0, 2, 0)$ incentivizes truthtelling. We call this the uncertainty effect. Which incentive dominates depends on the relative likelihood assigned to these two profiles, the latter itself depending on the probability assigned to the states g_1 and g_2. An agent holding a g_1-signal assigns a higher probability to state g_1 than to state g_2 and accordingly to profile $(0, 0, 2, 0)$ than to profile $(0, 0, 0, 2)$. The signal profile that incentivizes truthtelling is thus considered more likely than the one that incentivizes lying. Hence the hawk never prefers to announce a g_2-signal. We conclude that the TS equilibrium exists despite the existence of signal profiles generating ex post conflict.

We close with two remarks. By the same arguments as above, the TS equilibrium also exists under unanimity when $k = 3$. Moreover, the TS equilibrium continues to exist under sequential communication if the hawk speaks first. Indeed, the hawk's incentives then replicate those arising under simultaneous communication while doves still determine the outcome and hence have no incentives to deviate.[7]

_____

[7] The existence of the TS equilibrium under unanimity and sequential communication stands in contrast to

4   Analysis of the TS equilibrium

The example of Section 3 shows that the TS equilibrium can exist despite potential disagreement after full pooling of information. In what follows, we provide an equilibrium analysis for the general model.

For any signal $s \in S$, let σ_s denote the signal profile that consists of one signal s only. Moreover, for a given agent j, we denote the signal profile of all other agents by σ_{-j}. The following lemma addresses the effect of shifting mass from one entry of σ to another. This replicates the change in beliefs of other agents achievable by misreporting a signal in the putative TS equilibrium.

Lemma 1. *For any signal profile $\sigma = (x_1, ..., x_{m_I}, y_1, ..., y_{m_G})$, the function $\beta(\sigma)$ is invariant under any permutation of x-entries and any permutation of y-entries of σ. Moreover, the following inequalities hold:*

$$\beta(\sigma + \sigma_{g_r}) > \beta(\sigma + \sigma_{i_l}) \quad \forall l \in \{1, ..., m_I\}, r \in \{1, ..., m_G\}, \tag{4.1}$$

$$\beta(\sigma + \sigma_{g_l}) \geq \beta(\sigma + \sigma_{g_r}) \quad \forall l, r \in \{1, ..., m_I\}, y_r \leq y_l, \tag{4.2}$$

$$\beta(\sigma + \sigma_{i_l}) \leq \beta(\sigma + \sigma_{i_r}) \quad \forall l, r \in \{1, ..., m_I\}, x_r \leq x_l. \tag{4.3}$$

Conditions (4.2) and (4.3) hold with equality if and only if respectively $y_l = y_r$ or $x_l = x_r$.

Proof. See Appendix. □

Lemma 1 shows that three factors determine the posterior probability of guilt; an increase in the total number of g-signals and in the consistency of the profile of g-signals leads to an increase in the posterior probability of guilt. An increase in the consistency of the profile of i-signals has the opposite effect.

Note that if $\lambda = 1$ or $m_G = m_I = 1$ the impossibility result shown in Coughlan (2000) applies: the TS equilibrium exists if and only if either 1) at least k agents ($n - k$ agents) favor conviction (acquittal) for any realization of signals or 2) all agents favor the same action for any realization of signals. The existence of the TS equilibrium beyond these trivial cases thus requires both an enlarged state and signal space as well as the assumption that signals are informative with respect to the variant of innocence or guilt that applies.

Introducing terminology, we say that hawks have *critical mass* if the number of hawks is weakly greater than k, so that hawks are sufficiently many to impose conviction whenever they wish. Otherwise, doves have critical mass. We call a signal profile σ a *conflict profile* if conditional on σ hawks and doves disagree on the preferred action, that is, if $q_H < \beta(\sigma) < q_D$. We impose the following simple assumption on preferences of hawks and doves.

Assumption 1 (No partisans). *The preferred action of each agent depends on the aggregate signal profile. Moreover, hawks require less than the maximal possible evidence of guilt to prefer conviction and doves require less than the maximal possible evidence of innocence to prefer acquittal.*

---

the impossibility results of Austen-Smith and Feddersen (2006) and Van Weelden (2008).

Let q denote the threshold of the group that has critical mass and consider an agent j from the group that does not have critical mass. We denote the set of signal profiles σ_{-j} at which an i_r-report by agent j triggers an acquittal while a g_l-report triggers a conviction by

$$Piv_{i_r,g_l}(q) \equiv \{\sigma_{-j} : \beta(\sigma_{-j} + \sigma_{i_r}) < q \wedge \beta(\sigma_{-j} + \sigma_{g_l}) \geq q\}.$$

Our main result reads as follows.

Theorem 1. *Let k be non-unanimous and impose Assumption 1.*
a) Assume hawks have critical mass. The TS equilibrium exists if and only if

$$q_D \leq \hat{q}_D(q_H) \quad \equiv \quad \frac{1}{P(\sigma_{-j} \in Piv_{i_r,g_l}(q_H)|s_j = g_l)} \sum_{\sigma_{-j} \in Piv_{i_r,g_l}(q_H)} P(\sigma_{-j}|s_j = g_l) \cdot \beta(\sigma_{-j} + \sigma_{g_l}).$$

b) Assume doves have critical mass. The TS equilibrium exists if and only if

$$q_H \geq \hat{q}_H(q_D) \quad \equiv \quad \frac{1}{P(\sigma_{-j} \in Piv_{i_r,g_l}(q_D)|s_j = i_r)} \sum_{\sigma_{-j} \in Piv_{i_r,g_l}(q_D)} P(\sigma_{-j}|s_j = i_r) \cdot \beta(\sigma_{-j} + \sigma_{i_r}).$$

c) If hawks have critical mass and $q_D = \hat{q}_D(q_H)$ there exists at least one conflict profile. If doves have critical mass and $q_H = \hat{q}_H(q_D)$ there exists at least one conflict profile.

Proof. See Appendix. □

Theorem 1 provides a general existence result for the TS equilibrium. Part a) states the existence of a critical dove type $\hat{q}_D(q_H)$ such that the TS equilibrium exists if and only if $q_D \in (q_H, \hat{q}_D(q_H)]$. The threshold $\hat{q}_D(q_H)$ corresponds to the probability of guilt conditional on all pivotal profiles σ_{-j} where a truthful g_l-report of agent j leads to conviction while an i_r-report leads to acquittal. Part b) states the corresponding result for the case where doves have critical mass. Part c) yields the fundamental qualitative statement that the TS equilibrium is compatible with the existence of conflict profiles. It stands in stark contrast to Coughlan's impossibility result.

We outline the main steps of the proof of Theorem 1 in what follows. A first observation is that agents of the type that has critical mass never have an incentive to deviate as their preferred action given aggregated information is always implemented. A second observation is that under a non-unanimous voting rule, an agent of the type that does not have critical mass is never pivotal at the voting stage, irrespective of whether he deviated at the communication stage. Assuming hawks have critical mass, given sincere voting a dove's vote can only influence the outcome if hawks vote for acquittal. The reciprocal argument holds for the case of doves having critical mass. In both cases, however, *all* other agents will vote unanimously and thus the agent's vote cannot be pivotal. A third observation is that a hawk never has an incentive to misreport a g-signal as an i-signal while a dove never has an incentive to misreport an i-signal as a g-signal. This immediately follows from Lemma 1, Equation (4.1). A deviation of the above described type would only further worsen the implemented decision rule.

A fourth observation is that no agent has an incentive to misreport an i-signal as a different i-signal, or to misreport a g-signal as a different g-signal. Here, the uncertainty effect is key. Consider a juror holding signal i_r and contemplating announcing i_l instead, for $l \neq r$. The set of signal profiles σ_{-j} splits into pairs of profiles that are identical up to a simple permutation of the numbers of i_r- and i_l-signals. If both reports trigger identical actions for both these

profiles, the reporting decision is irrelevant. Otherwise, a truthful i_r-report will lead to an acquittal for the profile that features more i_r-signals and lead to a conviction for the other profile. Deviating to an i_l -report overturns both outcomes. If the defendant is guilty, both profiles are equally likely to occur and agent j is thus indifferent between the two reports. If the defendant is innocent, the profile that is more consistent with agent j's own signal is more likely than the other one given $\lambda > 1$. Hence agent j has an incentive to trigger acquittal for the former profile rather than the latter and thus to tell the truth. Deviations from one g-signal to another are ruled out by a similar argument.

Given the four above observations, the only deviations that remain to be excluded involve doves reporting an i-signal instead of a g-signal and hawks reporting a g-signal instead of an i-signal. These deviations have a clear effect on the outcomes via Equation (4.1). Here, the multiplicity of pivotal profiles allows the consensus effect to provide truthtelling incentives. While hawks and doves have diverging interests for some pivotal profiles, their preferred outcome coincides for others. This (partial) consensus is more pronounced the smaller $q_D - q_H$. Accordingly, we get an upper bound for q_D in Part a) and a lower bound for q_H in Part b). As for Part c), a type who is indifferent between truthtelling and lying necessarily faces pivotal profiles that incentivize truthtelling and pivotal profiles that incentivize deviating. For the sake of concreteness, consider a dove holding a g_l-signal and let hawks have critical mass. Let σ be a signal profile such that a hawk prefers conviction precisely for those signal profiles that yield at least as much evidence for the defendant being guilty as σ does. Whenever $q_D > \beta(\sigma)$, the profile σ is a conflict profile and incentivizes lying. On the other hand, Assumption 1 guarantees the existence of another signal profile $\tilde{\sigma}$ that satisfies $\beta(\tilde{\sigma}) > \beta(\sigma)$ and $\beta(\tilde{\sigma} - \sigma_{g_l} + \sigma_{i_r}) < \beta(\sigma)$. An example of such a profile $\tilde{\sigma}$ is one that has the same total number of i- and g-signals as σ but is either slightly less consistent with respect to its i-signals or slightly more consistent with respect to its g-signals. Now, if $\beta(\tilde{\sigma}) > q_D > \beta(\sigma)$ then $\tilde{\sigma}$ incentivizes truthtelling and this incentive dominates the deviation incentive from σ if q_D is sufficiently close to $\beta(\sigma)$.[8]

We finally sketch how our results generalize to arbitrary preference types. For $j \in \{1, ..., n\}$ let q_j be juror j's preference parameter and assume without loss of generality $q_1 \leq ... \leq q_n$. Let $k \in \{2, ..., n-1\}$ denote the voting rule. The TS equilibrium then implements juror k's optimal decision rule. For any juror $j < k$, the implemented decision rule is (weakly) "dovish" as $q_j \leq q_k$ while for any juror $j > k$ the decision rule is (weakly) "hawkish" as $q_j \geq q_k$. The insights from the two type case suggest that the TS equilibrium exists if and only if for all $j \in \{1, ..., n\}$ we have $q_j \in [\hat{q}_H(q_k), \hat{q}_D(q_k)]$ with $\hat{q}_H(q), \hat{q}_D(q)$ defined as in Theorem 1.

The problem is that a juror is now pivotal at the voting stage if exactly $k-1$ other jurors prefer conviction given the aggregate signal profile.[9] Consider a juror $j < k$ holding an i-signal. If this juror reports a g-signal instead, an irreversible conviction is triggered only if in addition to juror k also juror $k+1$ prefers to convict based on the reported evidence. Indeed, if only jurors 1 to k prefer conviction based on reports, juror j can veto a conviction. As a consequence, for profiles σ_{-j} where a particular g-announcement causes jurors 1 to k to

---

[8]Numerical simulations show that the parameter area that is compatible with the existence of the TS equilibrium in our model is typically larger than in the binary model in Coughlan (2000). Moreover, the number of conflict profiles compatible with the TS equilibrium becomes large when committees increase, contrasting e.g. Le Quement (2013).

[9]This is ruled out in the two-type setup by excluding unanimous voting rules.

favor a conviction based on reported evidence, a deviation to this announcement can only be advantageous as juror j can implement his favored decision at the voting stage.

Given these considerations, the following result holds. The TS equilibrium exists if and only if $q_j \in \left[\underline{q}\left(q_k, q_{k+1}\right), \overline{q}\left(q_k, q_{k-1}\right)\right]$ for all $j \in \{1, ..., n\}$ where the noteworthy aspect is that the bounds now depend on two preference types instead of only one. Furthermore, if jurors k and $k+1$ favor the same action for each signal profile we have $\underline{q}\left(q_k, q_{k+1}\right) = \hat{q}_H\left(q_k\right)$. Likewise, if jurors k and $k-1$ favor the same action for each signal profile we have $\overline{q}\left(q_k, q_{k-1}\right) = \hat{q}_D\left(q_k\right)$. It follows that if jurors $k-1$, k, and $k+1$ share the same optimal decision rule the TS equilibrium exists if and only if $q_j \in \left[\hat{q}_H\left(q_k\right), \hat{q}_D\left(q_k\right)\right] \ \forall j \in \{1, ..., n\}$. These insights provide some guidance regarding the optimal composition of heterogeneous committees with an eye to maximizing truthtelling incentives. In a committee too polarized for the TS equilibrium to exist, the inclusion of moderate agents endowed with decision power through a suitably chosen voting rule can help to overcome lying incentives. However, a single moderate agent will not suffice to ensure truthtelling.

5   Conclusion

In our collective decision model with pre-vote communication, a positive probability of ex post disagreement among agents is frequently compatible with the existence of the truthful communication and sincere voting equilibrium. The driving forces underlying our positive result are the consensus and uncertainty effects, both of which originate in the multiplicity of pivotal scenarios at the communication stage. The latter feature follows from the role played by consistency given our information structure. From a conceptual perspective, the key and novel feature of our information structure is that a given signal is interpreted differently depending on other available information; meaning is determined in context. We find this aspect worth exploring within other communication games.

Acknowledgments

Appendix

Proof of Lemma 1. We write $x \equiv \sum_{t=1}^{m_I} x_t$ and $y \equiv \sum_{t=1}^{m_G} y_t$. The statements follow immediately from $p \geq \max\left\{\frac{m_G - 1 + \lambda}{2m_G - 1 + \lambda}, \frac{m_I - 1 + \lambda}{2m_I - 1 + \lambda}\right\}$ and

$$\beta\left(\sigma\right)=\left[1+\frac{P\left(\omega\in I\right)}{P\left(\omega\in G\right)}\cdot\frac{m_G}{m_I}\cdot\left(\frac{p\cdot m_I}{\left(1-p\right)\cdot\left(\lambda+m_I-1\right)}\right)^x\cdot\left(\frac{\left(1-p\right)\cdot\left(\lambda+m_G-1\right)}{p\cdot m_G}\right)^y\cdot\frac{\sum_{r=1}^{m_I}\lambda^{x_r}}{\sum_{l=1}^{m_G}\lambda^{y_l}}\right]^{-1}.$$

Note that if $\lambda=1$ then $\beta\left(\sigma\right)$ only depends on x and y. $\qquad\square$

Proof of Theorem 1. Observations 1 to 3 from the main text are obvious. To show observation 4, suppose first that agent j who is not of critical mass type holds a signal $s=i_l$ and considers to report $m=i_r$ with $r\neq l$, $l,r\in\{1,...,m_I\}$. Consider two candidates for σ_{-j}, namely $\hat{\sigma}=(x_1,...,x_l,...,x_r,...,x_{m_I},y_1,...,y_{m_G})$ and $\hat{\sigma}_{x_l\longleftrightarrow x_r}=(x_1,...,x_r,...,x_l,...,x_{m_I},y_1,...,y_{m_G})$ and assume without loss of generality that $x_l\geq x_r$. We compare the expected utility of the reports $m=i_l$ and $m=i_r$ conditional on $\sigma_{-j}\in\{\hat{\sigma},\hat{\sigma}_{x_l\longleftrightarrow x_r}\}$. If both reports $m=i_l$ and $m=i_r$ trigger identical actions, the reporting decision does not matter. In particular, this is the case if $x_l=x_r$ by Equation (4.3). If the reports trigger different actions, then $x_l>x_r$ and thus $m=i_l$ will trigger acquittal for $\sigma_{-j}=\hat{\sigma}$ and conviction for $\sigma_{-j}=\hat{\sigma}_{x_l\longleftrightarrow x_r}$ while $m=i_r$ will trigger conviction for $\sigma_{-j}=\hat{\sigma}$ and acquittal for $\sigma_{-j}=\hat{\sigma}_{x_l\longleftrightarrow x_r}$, again by Equation (4.3). Hence

$$Eu\left[m=i_l|\sigma_{-j}\in\{\hat{\sigma},\hat{\sigma}_{x_l\longleftrightarrow x_r}\}\right]-Eu\left[m=i_r|\sigma_{-j}\in\{\hat{\sigma},\hat{\sigma}_{x_l\longleftrightarrow x_r}\}\right]$$

$$=\quad-\sum_{t=1}^{m_G}P\left(\omega=g_t|s=i_l\right)\cdot P\left(\sigma_{-j}=\hat{\sigma}|\omega=g_t,\sigma_{-j}\in\{\hat{\sigma},\hat{\sigma}_{x_l\longleftrightarrow x_r}\}\right)\cdot(1-q)$$

$$+\sum_{t=1}^{m_G}P\left(\omega=g_t|s=i_l\right)\cdot P\left(\sigma_{-j}=\hat{\sigma}_{x_l\longleftrightarrow x_r}|\omega=g_t,\sigma_{-j}\in\{\hat{\sigma},\hat{\sigma}_{x_l\longleftrightarrow x_r}\}\right)\cdot(1-q)$$

$$-\sum_{t=1}^{m_I}P\left(\omega=i_t|s=i_l\right)\cdot P\left(\sigma_{-j}=\hat{\sigma}_{x_l\longleftrightarrow x_r}|\omega=i_t,\sigma_{-j}\in\{\hat{\sigma},\hat{\sigma}_{x_l\longleftrightarrow x_r}\}\right)\cdot q$$

$$+\sum_{t=1}^{m_I}P\left(\omega=i_t|s=i_l\right)\cdot P\left(\sigma_{-j}=\hat{\sigma}|\omega=i_t,\sigma_{-j}\in\{\hat{\sigma},\hat{\sigma}_{x_l\longleftrightarrow x_r}\}\right)\cdot q$$

$$=\quad\frac{\frac{P(\omega\in I)}{m_I}\cdot\frac{p}{\lambda-1+m_I}}{\frac{P(\omega\in I)}{m_I}\cdot p+\frac{P(\omega\in G)}{m_G}\cdot(1-p)}\cdot\frac{-\lambda^{x_r+1}-\lambda^{x_l}+\lambda^{x_r}+\lambda^{x_l+1}}{\lambda^{x_r}+\lambda^{x_l}}\cdot q$$

$$>\quad0.$$

As the set of signal profiles possibly held by the other agents splits into pairs of the form $\{\hat{\sigma},\hat{\sigma}_{x_l\longleftrightarrow x_r}\}$ this shows that the proposed deviation is not profitable. Deviations from one g-signal to another are ruled out in the same way. It remains to analyze under which circumstances a dove holding some g-signal wants to deviate by reporting some i-signal instead (Part a) and under which circumstances a hawk holding some i-signal wants to deviate by reporting some g-signal (Part b).

a) + b) Assume that agent j is a dove holding signal $s_j=g_l$ for some $l\in\{1,...,m_G\}$ and considers reporting $m=i_r$ for some $r\in\{1,...,m_I\}$. By Equation (4.1), for any profile $\sigma_{-j}\in Piv_{i_r,g_l}\left(q_j\right)$ a truthful report $m=g_l$ will trigger conviction while reporting $m=i_r$ will trigger acquittal. Truthful reporting hence constitutes an equilibrium iff

$$\begin{aligned}0&\leq&Eu\left(m_j=g_l\right)-Eu\left(m_j=i_r\right)\\&=&\sum_{\sigma_{-j}\in Piv_{i_r,g_l}(q_H)}P\left(\sigma_{-j}|s_j=g_l\right)\cdot\beta\left(\sigma_{-j}+\sigma_{g_l}\right)-P\left(\sigma_{-j}\in Piv_{i_r,g_l}\left(q_H\right)|s_j=g_l\right)\cdot q_D\end{aligned}$$

10

which proves Part a). Part b) follows similarly.

c) Consider the case where hawks have critical mass. Let σ_H be a profile such that $\beta(\sigma_H) \geq q_j$ and $\beta(\tilde{\sigma}) < q_j$ for all signal profiles $\tilde{\sigma}$ satisfying $\beta(\tilde{\sigma}) < \beta(\sigma_H)$. Such a profile exists by Assumption 1. We need to show that $\hat{q}_D(q_H) > \beta(\sigma_H)$ in which case σ_H is a conflict profile. Suppose agent j is a dove holding a g_1-signal and considers deviating by reporting $m_j = i_1$. By Part a) it suffices to show that there exist $\sigma_{-j} \in Piv_{i_1,g_1}(q_H)$ such that

$$\beta(\sigma_{-j} + \sigma_{i_1}) < \beta(\sigma_H) < \beta(\sigma_{-j} + \sigma_{g_1}). \tag{A.1}$$

After reshuffling x- and y-entries, we may assume without loss of generality that σ_H satisfies $x_1 \geq ... \geq x_{m_I}$, $y_1 \geq ... \geq y_{m_G}$. First, assume that $y_2 > 0$. Then the profile $\sigma_{-j} = \sigma_H - \sigma_{g_2}$ satisfies (A.1). Similarly, if $x_2 > 0$ then profile $\sigma_{-j} = \sigma_H - \sigma_{i_2}$ satisfies (A.1). So suppose $x_2 = y_2 = 0$. If $y_1 = 0$ then $x_1 = n$ and $\sigma_H = (n, 0, ..., 0)$, so hawks would want to convict irrespective of any information. If $x_1 = 0$ then $y_1 = n$ and $\sigma_H = (0, ..., 0, y_1 = n, 0, ..., 0)$, so hawks would want to convict only given the maximal possible evidence of guilt. Both cases contradict Assumption 1. Finally, assume $x_1 \neq 0 \neq y_1$. Since $n \geq 3$ we must have $x_1 \geq 2$ or $y_1 \geq 2$. In the former case, $\sigma_{-j} = \sigma_H - \sigma_{i_1} - \sigma_{g_1} + \sigma_{i_2}$ satisfies (A.1) while in the latter case $\sigma_{-j} = \sigma_H - \sigma_{i_1} - \sigma_{g_1} + \sigma_{g_2}$ does. This finishes the proof for hawks having critical mass. The proof for the case of doves having critical mass is alike. $\square$

References

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

Austen-Smith, D. and Feddersen, T. J. (2006). Deliberation, preference uncertainty and voting rules, *American Political Science Review* **100**(2): 209–218.

Bhattacharya, S. (2013). Preference monotonicity and information aggregation in elections, *Econometrica* **81**(3): 1229–1247.

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

Dickson, D., Hafer, C. and Landa, D. (2008). Cognition and strategy: A deliberation experiment, *Journal of Politics* **70**(4): 974–989.

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

Gerardi, D., McLean, R. and Postlewaite, A. (2009). Aggregation of expert opinions, *Games and Economic Behavior* **65**(2): 339–371.

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

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

Hummel, P. (2012). Deliberation in large juries with diverse preferences, *Public Choice* **150**(3): 595–608.

Le Quement, M. T. (2013). Communication compatible voting rules, *Theory and Decision* **74**(4): 479–507.

Le Quement, M. T. and Yokeeswaran, V. (2015). Subgroup deliberation and voting, *Social Choice and Welfare* pp. 1–32.

Meirowitz, A. (2007). In defense of exclusionary deliberation: Communication and voting with private beliefs and values, *Journal of Theoretical Politics* **19**(3): 329–360.

Van Weelden, R. (2008). Deliberation rules and voting, *Quarterly Journal of Political Science* **3**(1): 83–88.

Wolinsky, A. (2002). Eliciting information from multiple experts, *Games and Economic Behavior* **41**(1): 141–160.