Persuasion, Transparency, and Commitment

Preliminary draft, comments are welcome

Timothy Feddersen* Ronen Gradwohl†

January 16, 2015

1 Introduction

Organizations ranging from small groups to large societies routinely set up committees who lack formal authority but offer policy advice to those with ultimate authority (Shepsle and Weingast, 1987). Some examples of such committees include academic departments who make recommendations to a dean regarding hiring and promotion; advisory committees for the U.S. Food and Drug Administration (FDA) on applications for new drugs (FDA, 2015); the Intergovernmental Panel on Climate Change that provides policy-relevant assessments on the science related to climate change (IPCC, 2015); the Federal Advisory Council that submits recommendations to the Federal Reserve Board (FAC, 2015); the Investor Advisory Committee that advises the U.S. Securities and Exchange Commission on various financial regulations and initiatives (IAC, 2015).

There is a large body of practice and law relating to the membership and procedures such advisory groups must use—for example, many federal committees are regulated by the Federal Advisory Committee Act of 1972. One common requirement is that advisory committees be transparent, i.e., that the arguments and procedural

*Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA. E-mail: tfed@kellogg.northwestern.edu.

†Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA. E-mail: r-gradwohl@kellogg.northwestern.edu.

steps—such as votes—used to derive the recommendations be accessible to the ultimate decision makers.

A notable example of this requirement lies in the context of the FDA's recent push towards improving the transparency of its operations (FDA, 2014). A document providing guidance on voting procedures, intended for FDA Advisory Committee members and staff, contains the following:

> Transparency and public participation are critical features of the advisory committee process.
>
> Accordingly, to help maximize the integrity, consistency, and utility of advisory committee voting results, FDA recommends that the voting process include the following procedures...the names of the committee members and their respective votes should be read aloud and otherwise made part of the public record shortly after the vote is taken. (FDA, 2008)

A first motivation for the requirement of transparency is, at one level obvious: if decision makers understand the information and processes used to make a recommendation they will be more confident in following the recommendation. However, it is widely understood that if members of a committee care about their personal reputations then imposing a requirement of transparency may alter their strategic landscape (see Levy, 2007a, among others). This may, in turn, counteract the potential benefit of transparency. In this paper we examine a different, but related, problem: if the members of the committee have policy preferences, and if they understand that the committee's recommendation is persuasive, then requiring transparency may make persuasion impossible. As the following simple example illustrates, this also calls the benefit of transparency into question.

Example 1.1 A small academic department consisting of three faculty members must vote on whether to hire a potential candidate. There are two equally-likely states of the world: in state A the candidate is suitable, and in state B she is not. Each faculty member interviews the candidate, and obtains a conditionally-independent signal about the quality of the candidate—each signal $s \in \{A, B\}$ has probability 0.6 in state s. Furthermore, each faculty member prefers hiring as long as he believes the likelihood of state A is at least 0.5. However, regardless of the department's vote, the final decision resides with a dean who is slightly risk averse, and prefers to hire only if she believes the likelihood of state A is at least 0.7.

Suppose that dean asks the committee for advice. The committee meets after each member has obtained his signal, each faculty member votes to hire if and only if he obtained a positive signal. If a majority of the faculty members vote to hire, then the department submits a recommendation to hire to the dean, and otherwise they recommend not to hire. Given that the dean anticipates this process how does the she react? Clearly, on a negative recommendation she will not hire the candidate. On a positive recommendation her posterior belief that the state is A is only 0.648,[1] which is not high enough for her to hire. Thus, here she also refuses to hire the candidate, and so the department is never able to persuade the dean. Note that the process used by the department was opaque: the dean does not observe the tally of votes cast.

It seems that transparency might alleviate this problem. Suppose that instead of obtaining a recommendation from the department, the dean attends the faculty meeting and observes the individual votes. If she believes that each faculty member is voting according to his signal then she will choose to hire only if all three faculty members vote to hire, as in this case her posterior is roughly 0.77. Thus, transparency has rendered the department persuasive! However, there is a fatal flaw in this analysis: the faculty members, aware of the dean in the room and the transparent nature of the interaction, will anticipate that the dean will hire only on a unanimous vote. But in this case, it is no longer an equilibrium for the faculty members to vote according to their signals (Feddersen and Pesendorfer, 1998). If the dean anticipates the department members strategic behavior then even a unanimous vote will be insufficient to persuade her. Accounting for strategic behavior by the committee, transparency does not improve the persuasiveness of the committee.

A second motivation for the requirement of transparency is more subtle. If the committee cannot commit to the process it uses to derive a recommendation, the decision maker may not trust that the recommendation was generated by persuasive process. In this case, it may be possible for transparency to serve as a substitute for commitment by the committee. This is illustrated in the following example.

[1] The derivations for this example appear in Example 3.3.

Example 1.2 Consider the same department and dean from Example 1.1, but suppose that the dean is slightly less risk averse and is willing to hire on a posterior of 0.65. As before, majority voting without transparency will not be persuasive. However, there are other processes that are persuasive. One such process is the following: the two faculty members vote according to their respective signals, and the third faculty member always votes to hire. The department then submits a positive recommendation to the dean upon a unanimous vote to hire. This turns out to be an equilibrium,[2] and in fact persuades the dean: on a positive recommendation, the dean's posterior belief is roughly 0.69, and so she hires the candidate.

But what if the department cannot commit to using this process? The department could claim to use that process, but instead could use the process in which all faculty members vote their signal, and then submit a positive recommendation whenever a majority are in favor of a hire. In fact, the department prefers the majority process to the previous one. Since the recommendations are indistinguishable, the dean can never be sure that the department did not switch to majority. And of course, if she believes that they did switch she would not hire on any recommendation.

Transparency may alleviate this problem. Instead of only submitting a recommendation to the dean, under transparency the dean observes the entire voting profile. It is thus impossible for the department to switch to a voting rule different from unanimity, and so, in a way, the department is "committed" to using unanimity. Hence, transparency serves as a substitute for commitment. Of course, this is not without its problems: as in Example 1.1, imposing transparency changes the incentives of the faculty members, and so possibly also their behavior. It is thus again not clear whether or not transparency will serve its intended goal.

In this paper we develop a model to investigate the impact of the transparency requirement on the performance of advisory committees, and in particular to determine whether the two motivating reasons for transparency discussed in the examples are viable. We consider a committee that must advise a decision maker on a binary choice. In our model nature chooses which of two alternatives is best, i.e., one of two states. Each member of the committee privately observes a signal that is correlated with the state but conditionally independent of the signals possessed by other mem-

[2]It is an asymmetric equilibrium à la McLennan (1998).

bers. The committee uses a decision process that consists of an extensive game, a strategy profile mapping signals into actions in the game, and a disclosure rule—a mapping from the terminal nodes of the game into an arbitrary message space that informs a decision maker. The disclosure rules we consider range from opaque, in which all terminal nodes map into one of two messages and so amount to a binary recommendation, to transparent, in which each terminal node of the game maps into a unique message. The decision maker observes the message disclosed by the decision process and then chooses one of two outcomes.

We ask two questions. First, is a given decision process persuasive—that is, does a rational decision maker who knows that the committee is using the process sometimes change his decision as a consequence of the information disclosed by it? Second, is a given decision process feasible—that is, given the subsequent behavior of the decision maker, is it an equilibrium for every member of the committee to act according to the strategy profile assumed by the process?

It is easy to see that for a given extensive game there are disclosure rules that are more transparent, and ones that are more opaque. We define a transparent version of any opaque decision process R as the process that consists of the same mechanism and strategy profile as R, but in which the disclosure rule not only reveals the policy recommendation of the opaque process but also the entire history of play that produced the recommendation.

In the first part of the paper we address the problem illustrated by Example 1.1—that requiring transparency affects the strategic landscape in such a way as to counteract its possible benefits—and so assume that the mechanism and disclosure rule used by the committee are fixed exogenously. We show that the transparent version of any persuasive opaque decision process is also persuasive, and additionally that there are persuasive transparent decision processes whose opaque counterparts are not persuasive. Thus, from an ex post perspective the decision maker's preference for a transparent decision process is well founded.

However, we show that unless the preferences of the committee and the decision maker are very close, any transparent process that is persuasive is not feasible. In particular, transparency fails to alleviate the problem of the small department of Example 1.1, and in fact here there is no feasible process, opaque or transparent, that is persuasive. In contrast, for sufficiently large committees there are opaque decision processes that are both persuasive and feasible. In these cases, the imposition of transparency actually destroys the persuasiveness of the committee. Thus, except for a narrow range of parameters, the first part of the paper makes a strong case against transparency.

In the second part of the paper we address the problem illustrated by Example 1.2—that without transparency the decision maker may not be sure which decision process the committee is using. We show that transparency does in fact address this problem, and may be a substitute for commitment, but only for a narrow range of parameters. In particular, in Example 1.2 imposing transparency achieves the same effect as the department committing to the feasible unanimous voting process. However, once the preferences of the dean and committee are a bit further apart—specifically, if the dean is more risk averse, and agrees to hire only on a posterior greater than 0.7, then transparency no longer solves this problem. More generally, we show that when the decision maker's and committee's preferences are not sufficiently close then if transparency serves as a successful substitute for commitment to a particular process, then that process cannot be persuasive.

The takeaway from the paper is that in the case of advisory committees in which committee members do not have reputation concerns, transparency significantly undercuts persuasiveness. We stress that although the examples above relate to simple voting processes, our results are much more general, and apply to arbitrary processes that might include, for example, communication amongst the committee members and sequential voting.

The structure of the paper is as follows. Immediately following is a brief literature review, and then in Section 2 we set up the model. In Sections 3 and 4 we provide our central results for the case of an exogenously given decision process—the former on the (in)compatibility of persuasiveness and transparency, and the latter on welfare. This is followed by Section 5, in which we allow the committee to choose, but not credibly commit to, the decision process. Finally, some of the proofs are relegated to the Appendix.

1.1 Literature review

There is a large literature examining transparency in agency relationships—see Prat (2005) for an extensive review. The essential idea is that there is a misalignment of incentives between the principal and the agent, and transparency allows the principal to design incentives for agents in order to align incentives. As Prat observes, it is well known that transparency also comes with costs. In particular, when agents worry about the impact of third parties observing their behavior or about career and reputation concerns more generally, such alignment of incentives may be hindered. This effect arises also in contexts similar to that of our paper, in which a committee of informed experts provides advice to a decision maker. For example, a variety of papers examine incentives for position taking (c.f. Mayhew, 1974) and vote buying, in which policymakers are rewarded by outsiders, such as special-interest groups, for voting a particular way. Such policymakers, in addition to being concerned about the outcome of a vote, are thus also concerned about how their individual votes are perceived. Similarly, papers such as Snyder and Ting (2005), Dal Bó (2007), Felgenhauer and Grüner (2008), and Seidmann (2011) analyze decision-making in committees when members are influenced by outsiders in this way. Additionally, a growing literature (including Sibert, 2003; Fingleton and Raith, 2005; Visser and Swank, 2007; Levy, 2007a,b; Stasavage, 2007; Gersbach and Hahn, 2008) examines the effects of committee members' career concerns on decision making. The main difference between these literatures and the current paper is that members of our committee do not face either reputation or career concerns. There is no threat of action by some third party and no ability of the decision maker to provide any individual incentives to members of the committee.

The basic structure of our model follows the literature on voting and information aggregation studied by Austen-Smith and Banks (1996) and Feddersen and Pesendorfer (1998), among others. In those papers an electorate is composed of individuals with private information and common values who face incentives to vote strategically, i.e., not vote according to their private signal, as a function of the voting rule being used. In this paper we adapt a similar information structure but there is no fixed voting rule that determines the outcome. Instead, the outcome of the committee's activities, such as voting, is reported to a decision maker, who then updates her beliefs about the correct decision and makes a final choice. In this respect our paper is more closely related to the signaling game literature initiated by Crawford and Sobel (1982). Particularly relevant is work by Gillian and Krehbiel (1989) in which a legislature may delegate a decision to a better-informed committee. Unless agenda-setting power is delegated to the committee there is no possibility of information transmission if the committee and decision maker have preferences that are too different. Work by Austen-Smith (1993), Krishna and Morgan (2001) and Battaglini (2002) considers the possibility that multiple experts, e.g., on a committee, may be independently queried by a decision maker, and demonstrate that information transmission is possible even when preferences of experts and the decision maker are substantially different.

On the mechanism design front, work by Gerardi and Yariv (2007) demonstrates that when committees are free to deliberate, i.e., use more complicated decision processes than simply voting, prior to using a single decisive voting rule, then the resulting equilibria include the equilibrium of every voting game. Deliberation makes the final and presumably decisive voting rule redundant, unless the final rule is unanimity. Gerardi and Yariv (2007) do not consider advisory committees or the impact of transparency on deliberations. Gershkov and Szentes (2009) consider an alternative mechanism design problem in which a social planner chooses to query experts in order to make a binary decision. Information acquisition is costly, so experts have an incentive to free-ride. They show that the optimal mechanism sequentially queries agents, who in turn rationally choose to incur the cost to get informed if they are queried.

Perhaps the paper most closely related to ours is Wolinsky (2002). In his paper a set of experts, each of whom observes a private binary signal, can provide binary information to a decision maker (DM). The DM makes a binary choice between H and L, and prefers the former if and only if the number of high signals observed by the experts is above some threshold. The experts prefer the DM choose H if and only if the number of high signals is above a threshold higher than the DM's. Wolinsky shows that when the DM queries each expert separately there is no equilibrium in which DM learns information that causes her to choose H. This is similar to our Theorem 3.4, for the case of decision processes that consist of voting. Wolinsky considers two possible solutions: commitment by the DM and communication amongst the experts. Under commitment the DM may commit to a voting rule, and in that case he shows that the optimal rule is non-monotonic in the number of high signals reported. With communication, Wolinsky assumes that because the experts have common values they will report their high signals if and only if the number of such signals is above their threshold. In this case the DM can do no better than what the experts prefer. Wolinsky comes up with a creative process that allows transmission of more information: put the experts into multiple groups, such that those within each group can communicate (and coordinate their reports) but those in different groups cannot. In Wolinsky (2002) there is no discussion of transparency per se, but it is somewhat implicit—when the DM queries each expert separately, that is equivalent to transparent voting, and when he allows communication he implicitly allows them not to reveal their votes (opacity). However, Wolinsky (2002) does not consider the problem of persuasion from the perspective of the advisory committee. Furthermore, as his focus is on information transmission, he does not study the effect of imposing transparency, neither with nor without the ability of the committee to commit to a process.

2 Model and Definitions

There are two possible, equally likely states of the world, $\Theta = \{A, B\}$, and a decision maker (DM) who must decide between two possible outcomes, $\mathcal{O} = \{A, B\}$. Choosing B always yields a utility of 0, whereas choosing A yields a gain in state A and a loss in state B. The DM is risk averse, and so the loss incurred from a bad decision is greater than the gain from a good decision. Formally, the DM is equipped with a utility function $u_D : \Theta \times \mathcal{O} \mapsto \mathbb{R}$ that satisfies $u_D(A, A) > 0$, $u_D(B, A) < -u_D(A, A)$, and $u_D(\theta, B) = 0$ for both $\theta \in \Theta$.[3] We may assume without loss of generality that $u_D(A, A) = 1$.

Given a belief $\beta = \Pr[\theta = A]$ about the probability that the state is A, the DM's expected utility on choosing outcome o is $U_D(\beta, o) \stackrel{\text{def}}{=} \beta \cdot u_D(A, o) + (1-\beta) \cdot u_D(B, o)$. A rational DM will choose outcome A only if $U_D(\beta, A) \geq 0 = U_D(\beta, B)$. Since $U_D(\beta, A)$

[3]The assumption that $u_D(\theta, B) = 0$ is for simplicity only—all our results hold for any utility function in which the threshold $\beta_D > 1/2$ (see next paragraph).

is increasing in β, there exists a threshold β_D such that the DM will choose A only if $\beta \geq \beta_D$. By our assumption on u_D it holds that $\beta_D > 1/2$. One threshold that will play an important role in this paper is $\overline{\beta}(p) = p^2/(p^2 + (1-p)^2)$, and we omit its dependence on p when clear from context.

Next, there is an advisory committee of n members, $\{1, \ldots, n\}$, each of whom receives a conditionally-independent, identically-distributed signal $s_i \in \{a, b\}$ satisfying

$$\Pr[s_i = a|\theta = A] = \Pr[s_i = b|\theta = B] = p,$$

where $p \in (1/2, 1)$ is the *accuracy* of the signal. Each of the n committee members has a common utility function $u : \Theta \times \mathcal{O} \mapsto \mathbb{R}$. The utility function is identical to the DM's, except that committee members are not risk averse, and so $u(B, A) = -1$.[4] Observe that the committee members prefer outcome A whenever their belief of the state being A is $\beta_C \geq 1/2$.

The order of events is as follows: (1) The committee members obtain their signals; (2) they participate in a mechanism; (3) some information about the path of play is disclosed to the DM; and (4) the DM chooses an outcome.

To describe the mechanism in which committee members participate, as well as the information disclosed, we first recall the following definition of an extensive game with perfect information and simultaneous moves (see for example Osborne and Rubinstein, 1994):

Definition 2.1 (game) *An n-agent* game *is a pair (H, P) where*

- *H is a set of finite* history *sequences such that the empty word $\epsilon \in H$. A history $h \in H$ is* terminal *if $\{a : (h, a) \in H\} = \emptyset$. The set of terminal histories is denoted $Z \subseteq H$.*

- *P is a function that maps nonterminal histories h to sets of agents, namely subsets of $\{1, \ldots, n\}$. For an agent i such that $i \in P(h)$, denote by $A_i(h)$ the set of* actions *available to agent i at history h, such that $\times_{i \in P(h)} A_i(h) = \{a : (h, a) \in H\}$.*

[4]In a few of our examples we will assume that $u(B, A) = -1 \pm \varepsilon$, where $\varepsilon > 0$ is an arbitrarily small number, so as to break ties.

Example 2.2 A particular type of game that will play an important role in this paper is the *voting game*: $P(\epsilon) = \{1, \dots, n\}$, and $H = \{\epsilon\} \cup \{a, b\}^n$. That is, every agent i has an action that is a vote $A_i(\epsilon) = \{a, b\}$, actions are simultaneous and so $A_i(h) = \emptyset$ for all $h \neq \emptyset$, and terminal histories consist of all profiles of votes $\{a, b\}^n$.

A strategy σ_i of agent i in a game (H, P) is a function that, for every h and $i \in P(h)$, maps every (s_i, h) to an element of $A_i(h)$. Denote by $\sigma = (\sigma_1, \dots, \sigma_n)$ a profile of strategies, by $\sigma(s)$ the profile of strategies given signal profile $s = (s_1, \dots, s_n)$, and by $Z(\sigma(s))$ the distribution over terminal histories reached when the profile $\sigma(s)$ is played. In a voting game, the *sincere* voting profile is the one in which $\sigma_i(s_i, \epsilon) = s_i$ for all agents i.

While the committee members play an extensive game, the DM only observes a message about the path of play. This message is specified by the *disclosure rule*: a function $d : Z \mapsto D$, where D is some set of messages. A *mechanism* is then a game (H, P) coupled with a disclosure rule d.

Example 2.3 A *voting mechanism* is a voting game with a disclosure rule $d : \{a, b\}^n \mapsto \{m_A, m_B\}$. For example, the *majority mechanism* is a voting game with the disclosure rule that satisfies $d(z) = m_A$ if and only if the number of a votes in z is at least $n/2$. Similarly, the *unanimity mechanism* is the voting game with the disclosure rule in which $d(z) = m_A$ if and only if $z = (a, \dots, a)$.

Definition 2.4 (process) *A* process *is a tuple (H, P, d, σ) that consists of a mechanism and a strategy profile. A process is* transparent *if $d(h) = h$ for every $h \in Z$, and is* non-transparent *otherwise. A process is* opaque *if the set D in the range of d has cardinality 2. The* transparent version *of a process (H, P, d, σ) is the transparent process (H, P, d^T, σ), and an* opaque version *of that process is an opaque process (H, P, d^O, σ).*

Note that there may be multiple opaque versions of the same process: for example, the set of opaque versions of a voting mechanism consists of all possible voting rules, including majority and unanimity. In contrast, every process has only one transparent version.

After the mechanism is played, the DM obtains a message specified by the disclosure rule. If he is rational, he will update his prior on the state, and will take an action that depends on whether the posterior surpasses his threshold β_D or not. Formally, given a process (H, P, d, σ), denote the decision rule used by the DM on message $m \in D$ as $r(m)$, where $r(m) = A$ if $\Pr[\theta = A \mid m] \geq \beta_D$, and $r(m) = B$ otherwise.

We next define two desirable properties of processes: persuasiveness and feasibility. A process is persuasive if there are some instances in which the DM is influenced by it—that is, if there exist two messages, sent with positive probability, such that a DM who acts rationally will make different decisions after observing these messages. Formally,

Definition 2.5 (persuasiveness) *A process* (H, P, d, σ) *is* persuasive *if there exist two messages* $m, m' \in D$ *such that*

- $\Pr[d(H(\sigma(s))) = m] > 0$ *and* $\Pr[d(H(\sigma(s'))) = m'] > 0$ *for some signal profiles* $s, s' \in \{a, b\}^n$*, and*
- $r(m) \neq r(m')$*.*

A process is feasible if the strategy profile constitutes a Nash equilibrium for the agents conditional on the DM acting rationally. Formally,

Definition 2.6 (feasibility) *A process* (H, P, d, σ) *is* feasible *if for each agent* i*, signal* s_i*, and strategy* σ'_i*,*

$$\mathrm{E}\left[u(\theta, r(d(H(\sigma(s))))) \mid s_i\right] \geq \mathrm{E}\left[u(\theta, r(d(H(\sigma'_i, \sigma_{-i}(s))))) \mid s_i\right],$$

where the expectation is over θ*,* s*, and* σ*.*

3 Persuasiveness and Transparency

Are transparency and persuasiveness compatible? And if yes, is it possible for transparency to sometimes facilitate persuasiveness?

3.1 Without feasibility

When feasibility is not a concern, transparency and persuasiveness are compatible:

Proposition 3.1 *The transparent version of any persuasive process is also persuasive.*

Proof: Fix a process a process (H, P, d, σ), and denote the disclosure rule of its transparent version by $d^T : Z \mapsto D^T$. If the process is persuasive, then there exist two messages $m, m' \in D$ such that $r(m) = A$ and $r(m') = B$. By an averaging argument, this implies that there exist $h, h' \in Z$ such that $d(h) = m$ and $d(h') = m'$ and for which $\Pr[\theta = A \mid h] \geq \beta_D$ and $\Pr[\theta = A \mid h'] < \beta_D$. Thus, in the transparent process $r(d^T(h)) = A$ and $r(d^T(h')) = B$, and so it is persuasive. ■

Sometimes transparency can even facilitate persuasiveness:

Proposition 3.2 *There exist processes that are not persuasive but whose transparent versions are.*

The proof is by example:

Example 3.3 Consider a 3-agent committee and the majority mechanism, together with sincere voting. Suppose also that β_D is greater than the DM's posterior on message m_A—namely,

$$\beta_D > \Pr[\theta = A \mid m_A] = p^3 + 3p^2(1-p).$$

Because the posterior is not high enough, the process is not persuasive. However, if β_D is small enough so that the DM would be convinced if all three agents received the signal $s_i = a$, then the transparent version would be persuasive. Formally, if $\beta_D \leq p^3/(p^3 + (1-p)^3)$ then $r(a, a, a) = A$, and $r(z) = B$ otherwise. Thus, the opaque process is not persuasive while its transparent version is whenever

$$\beta_D \in \left(p^3 + 3p^2(1-p), \frac{p^3}{p^3 + (1-p)^3}\right].$$

3.2 With feasibility

Our main theorem in this section is an impossibility result stating that persuasiveness and transparency are incompatible whenever feasibility is desired.

Theorem 3.4 *There does not exist any feasible, persuasive, and transparent process for any n, $p > 1/2$, and $\beta_D > \overline{\beta}(p)$.*

In particular, note how Example 3.3 fails: in the transparent version, the decision rule used by the DM is unanimity, but the agents are voting sincerely. However, sincere voting is not an equilibrium under unanimity (**?**). So even though the transparent version is persuasive, it is not feasible.

This example helps illustrate the intuition behind Theorem 3.4. Consider any voting mechanism coupled with sincere voting. The transparent version of this process is feasible only if the decision rule $r(\cdot)$ used by the DM to determine the outcome is such that he chooses A whenever half or more of the agents vote a. But on a profile in which exactly half vote a, the DM will not choose A since his posterior will be $\beta = 1/2$, which his below his threshold of $\overline{\beta} > 1/2$. So his decision rule will be one in which a super-majority of a votes is required in order for him to choose A. But in this case, sincere voting is no longer an equilibrium for the agents (**??**). The agents must thus choose a strategy profile in which they sometimes randomize. This, however, adds additional noise to the voting profile, and so the DM would require a greater super-majority of votes for a in order for him to choose A. Theorem 3.4 states that this completely unwinds, and that the only feasible processes are non-persuasive, with the DM always choosing the same outcome (which must thus be B). In fact, the theorem shows that this holds for any mechanism, and not just the simpler voting mechanisms.

Theorem 3.4 is "tight" in the sense that if $\beta_D = \overline{\beta}$ then there does exist a feasible, persuasive, transparent process:

Example 3.5 Fix any n, and consider the voting mechanism coupled with the strategy profile in which agents 1 and 2 vote sincerely and the others always vote a. When $\beta_D = \overline{\beta}$, the DM's rational decision rule in the transparent version of this process is $r(a, \ldots, a) = A$ and $r(z) = B$ otherwise, and so the process is persuasive. Furthermore, it is feasible—this is an asymmetric equilibrium à la McLennan (1998).

In contrast with the incompatibility of persuasiveness and transparency, persuasiveness and opacity are often compatible when committees are large enough. In particular, this is the case in voting mechanisms.

Fix a voting game (H, P), any $\alpha \in (0,1)$, and denote by d^α the disclosure rule for which $d^\alpha(z) = m_A$ whenever $z \in Z$ is a profile with at least $\lfloor \alpha n \rfloor$ votes a, and $d^\alpha(z) = m_B$ otherwise.

Theorem 3.6 *For any β_D and $\alpha \in (0,1)$ there exists n_0 such that for all $n \geq n_0$ there exists a strategy profile σ such that (H, P, d^α, σ) is feasible and persuasive, where (H, P) is the n-agent voting game.*

Proof: FP show full information equivalence. Show that this implies our result. ■

3.2.1 Caveats

Despite the impossibility and possibility results of Theorems 3.4 and 3.6, sometimes transparency can be compatible with persuasiveness, and can actually facilitate persuasiveness, even subject to feasibility. In light of the two theorems, this can only apply under certain settings of the parameters. This is illustrated by the following two examples. The first is an opaque process that is persuasive but not feasible, and the second is an opaque process that is feasible but not persuasive. However, the transparent versions of the processes are both feasible and persuasive.

Example 3.7 Fix some odd $n > 1$, and consider the unanimity mechanism with sincere voting. Also fix $\beta_D \leq p$. Then this process is persuasive but not feasible. In the transparent version $r(z) = A$ if and only if $\lceil n/2 \rceil$ of the players vote a. This is both feasible and persuasive.

Example 3.8 Consider a 2-agent committee with the weak majority mechanism and sincere voting. This is feasible, but if

$$\beta_D > \frac{p^2 + 2p(1-p)}{p^2 + (1-p)^2 + 4p(1-p)}$$

it is not persuasive, as the right-hand-side is the DM's posterior on message m_A. In the transparent version $r(z) = A$ if and only if $z = (a, a)$. This is both feasible and persuasive when $\beta_D \leq \overline{\beta}$, since on this terminal history the DM's posterior is $\overline{\beta}$.

4 Welfare

4.1 Without feasibility

First, without feasibility, transparency always (weakly) increases the DM's welfare.

Proposition 4.1 *Given any opaque process, the DM weakly prefers its transparent version.*

Proof: Fix an opaque process $R = (H, P, d, \sigma)$ with $d : Z \mapsto \{m_A, m_B\}$, and its transparent version $R^T = (H, P, d^T, \sigma)$. In the following, we slightly abuse notation and denote by m_A and m_B, the messages sent by the mechanism, also as the events $(d(H(\sigma(s))) = m_A)$ and $(d(H(\sigma(s))) = m_B)$ (when clear from context). Similarly, we denote by $h \in Z$, a terminal history, also as the event $(d^T(H(\sigma(s))) = h)$.

Define the sets $H_A = \{h \in Z : d(h) = m_A\}$ and $H_B = \{h \in Z : d(h) = m_B\}$. Suppose the process R is persuasive, and so without loss of generality that $r(m_A) = A$ and $r(m_B) = B$. The expected utility of the DM under R is thus

$$U_D^O \stackrel{\text{def}}{=} \Pr[m_A] \cdot U_D(\Pr[\theta = A|m_A], A) + \Pr[m_B] \cdot U_D(\Pr[\theta = A|m_B], B),$$

where

$$U_D(\Pr[\theta = A|m_A], A) = \sum_{h \in H_A} U_D(\Pr[\theta = A|h], A) \cdot \Pr[h|m_A]$$

and

$$U_D(\Pr[\theta = A|m_B], B) = \sum_{h \in H_B} U_D(\Pr[\theta = A|h], B) \cdot \Pr[h|m_B] = 0.$$

The expected utility of the DM under the transparent version R^T is

$$\begin{aligned} U_D^T &\stackrel{\text{def}}{=} \sum_{h \in Z} \Pr[h] \cdot U_D(\Pr[\theta = A|h], r(h)) \\ &= \sum_{h \in Z} \Pr[h] \cdot \max\{U_D(\Pr[\theta = A|h], A), U_D(\Pr[\theta = A|h], B)\}. \end{aligned}$$

Rewriting yields

$$U_D^T = \Pr[m_A] \cdot U_D'(\Pr[\theta = A|m_A]) + \Pr[m_B] \cdot U_D'(\Pr[\theta = B|m_B]),$$

where

$$U_D'(\Pr[\theta = A|m_A]) = \sum_{h \in H_A} U_D(\Pr[\theta = A|h], r(h)) \cdot \Pr[h|m_A]$$

and

$$U'_D(\Pr[\theta = A|m_B]) = \sum_{h \in H_B} U_D(\Pr[\theta = A|h], r(h)) \cdot \Pr[h|m_B].$$

Observe that each term in U'_D is weakly greater than the corresponding term in U_D, since $U_D(\Pr[\theta = A|h], r(h)) \geq U_D(\Pr[\theta = A|h], A)$ and $U_D(\Pr[\theta = A|h], r(h)) \geq U_D(\Pr[\theta = A|h], B)$. Thus, $U_D^T \geq U_D^O$. ■

In fact, without feasibility, transparency can sometimes *strictly* increase the welfare of *both* the DM and the committee.

Proposition 4.2 *There exist opaque processes for which both the DM and the committee members strictly prefer the transparent version.*

Proof: Consider Example 3.3, with the 3-agent majority mechanism. In the example, the opaque version is not persuasive, and so the DM always chooses B and so his utility and that of the committee members is 0. In contrast, in the transparent version the DM chooses A on history (a, a, a) and so his expected utility is

$$\begin{aligned} U_D^T &= \Pr[h = (a,a,a)] \cdot U_D(\Pr[\theta = A|h = (a,a,a)], A) \\ &= \frac{p^3 + (1-p)^3}{2} \cdot U_D\left(\frac{p^3}{p^3 + (1-p)^3}, A\right) \\ &= \frac{p^3 + (1-p)^3}{2} \cdot \left(\frac{p^3 + (1-p)^3}{2} \cdot u_D(A, A) + \left(1 - \frac{p^3 + (1-p)^3}{2}\right) \cdot u_D(B, A)\right) \\ &> 0 \end{aligned}$$

whenever $\beta_D < p^3/(p^3 + (1-p)^3$. The same holds for the committee members, whose threshold is 1/2. ■

Note that although transparency always increases the welfare of the DM, and sometimes also increases that of the committee members, it is also possible for transparency to decrease the welfare of the committee. One example in which this occurs is Example 3.3, when β_D is small enough so the opaque process is persuasive. In this case, the committee prefers the opaque processes, which is majority, to the transparent version, in which the decision rule of the DM implements unanimity.

4.2 With feasibility

With feasibility, it is no longer the case that feasibility increases the welfare of the DM, and certainly not the welfare of both DM and committee.

Proposition 4.3 *Fix $\beta_D > \overline{\beta}$, and let $R^O = (H, P, d^O, \sigma)$ be a feasible opaque process. Then for any feasible transparent process $R^T = (H, P, d^T, \sigma')$ the DM weakly prefers R^O to R^T. For some processes R^O, the DM strictly prefers R^O to all feasible transparent processes R^T.*

Proof: By Theorem 3.4, R^T is not persuasive. If R^O is also not persuasive, the DM is indifferent between the two, since both yield a utility of 0. If R^O is persuasive, then it yields the DM a strictly positive expected utility. He thus prefers R^O to all transparent processes, since the latter is not persuasive. ■

4.2.1 Caveats

If the DM's and committee's preferences are close enough to each other, it is possible for the DM and committee to both strictly prefer the transparent version of a process to its opaque version. For instance, this is the case in Example 3.8, in which the opaque process is not persuasive—and so the utilities of DM and committee are 0—but the transparent version is persuasive—and so both DM and committee obtain strictly positive expected utility.

5 Transparency and (Non-)Commitment

The committee may be able to choose but not commit to a particular mechanism. One might think that transparency serves as a substitute for commitment by forcing the use of other mechanisms, possibly ones that are suboptimal for the committee but better for the DM. Is this true?

5.1 Extended model and definitions

In addition to the n agents and the DM, there is a committee chair (henceforth chair), who has the same utility function as the committee members but does not receive any signal.

The order of events is as follows: (1) The chair chooses a mechanism; (2) the committee members obtain their signals; (3) they participate in the chosen mechanism; (4) some information about the path of play is disclosed to the DM; and (5) the DM makes his decision.

A process is realizable if, fixing the committee's strategy profile, a rational chair does not prefer a different mechanism over the one specified by the process. The mechanisms must satisfy an indistinguishability criterion: For non-transparent processes, it must be the case that the set of possible messages receivable by the DM be a subset of the set of messages in the original mechanism. For transparent processes, it must be the case that the possible realized histories in the mechanism be a subset of the possible realized histories in the original mechanism. First a preliminary definition: A strategy profile σ is well-defined for a process (H, P, d, σ) if σ_i is a strategy for agent i in the game (H, P), for every agent i.

Definition 5.1 (realizability) *A non-transparent (transparent) process $R = (H, P, d, \sigma)$ is* realizable *if for all non-transparent (transparent) processes $R' = (H', Z', A', d', \sigma)$ for which σ is well-defined and with* $\text{range}(d') \subseteq \text{range}(d)$*, a rational chair will not want to deviate to R'. The process R' that the chair prefers most is the* most preferred alternative *to R.*

Claim 5.2 *All transparent processes are realizable.*

Definition 5.3 (feasible realizability) *A non-transparent (transparent) process $R = (H, P, d, \sigma)$ is* feasibly realizable *if for all non-transparent (transparent) processes $R' = (H', Z', A', d', \sigma')$ with* $\text{range}(d') \subseteq \text{range}(d)$*, a rational chair will not want to deviate to R'.*

Note that not all transparent processes are feasibly realizable.

5.2 Without feasibility, transparency is a substitute for commitment

Proposition 5.4 *There exist processes that are not realizable and whose most preferred alternatives are not persuasive, but such that the transparent versions of those processes are both realizable and persuasive.*

Proof: Proof by example: Opaque 3-agent unanimity, with $\beta_D > \overline{\beta}$ and sincere voting profiles. Note that chair will deviate to majority, which would render the process not persuasive. ■

5.3 With feasibility, transparency will usually not be a substitute for commitment

Proposition 5.5 *If $\beta_D > \overline{\beta}$, then every feasibly realizable, transparent process is not persuasive.*

Proof: Feasible realizability of a transparent process $R = (H, P, d, \sigma)$ implies that σ is an optimal strategy profile for (H, P), and so an equilibrium. Thus, R is feasible, and so by Theorem 3.4 not persuasive. ■

5.4 Caveats

There is a narrow set of parameters under which transparency is a substitute for commitment:

Proposition 5.6 *If the DM's and committee's preferences are close enough to each other, then there exist persuasive processes that are not feasibly realizable, but whose transparent versions are feasibly realizable and persuasive.*

Proof: Proof by example: Suppose agents will *not* hire on posterior 1/2. Consider 3-agent unanimity. Make sure that agents prefer majority, whereas DM prefers unanimity with profile in which one agent always votes to hire and others vote sincerely. $\beta_D = \overline{\beta} - \varepsilon$. ■

5.5 Partial transparency

Partial transparency may be a substitute for commitment.

Definition 5.7 (partial transparency) *A process is* (H, P, d, σ) partially transparent *if there exists a nonempty subset of terminal histories $Z(d) \subseteq Z$ such that $d(h) = h$ for every $h \in Z(d)$.*

One example of partial transparency is *agent transparency*. An agent i is transparent if the message disclosed to the DM includes a complete description of agent i's actions. Formally, for a terminal history $h \in Z$ denote by h_i the actions taken by agent i in history h. Then agent i is transparent in a process (H, P, d, σ) if for all $h \in Z$ it holds that $d(h) = (h_i, m_{-i})$, where m_{-i} is a message that does not depend on agent i's actions.

One way to intuitively implement agent transparency in a voting context is with the DM in the room and agent i voting publicly, while other agents vote by secret ballot and only announce the outcome. This way the DM observes the actions of agent i, and only the aggregate of other agents' actions.

A fictional way to implement agent transparency is from Dave Eggers' novel "The Circle", in which individuals and especially politicians are urged to "go transparent"—namely, constantly wear a camera around their necks to capture and broadcast all their actions.

Definition 5.8 (feasible realizability for partial transparency) *A partially transparent process $R = (H, P, d, \sigma)$ is* feasibly realizable *if for all partially transparent processes $R' = (H', Z', A', d', \sigma')$ with* $\text{range}(d') \subseteq \text{range}(d)$ *and $Z(d') = Z(d)$, a rational chair will not want to deviate to R'.*

Proposition 5.9 *There exist persuasive, feasibly realizable, partially transparent processes whose opaque and transparent versions are not feasibly realizable.*

Proof: Same example as above, but with $\beta_D > \overline{\beta}$, and so neither opacity nor transparency are feasibly realizable. However, partial transparency, and in particular agent transparency, may be: Suppose DM always observes actions of agent 1, and hires only if all vote to hire. (Unanimity but with actions of agent 1 transparent.) This is partially transparent, where message space is the set $\{a, b\} \times \{A, B\}$, where first element is the vote of agent 1 and the second element is the decision of the other two. ■

Appendix

A Proof of Theorem 3.4

Theorem A.1 (Theorem 3.4 restated) *There does not exist any feasible, persuasive, and transparent process for any n, $p > 1/2$, and $\beta_D > \overline{\beta}(p)$.*

Fix n and $p > 1/2$, as well as an n-agent feasible, persuasive, and transparent process (H, P, d, σ). We will show that this implies $\beta_D \leq \overline{\beta}(p)$.

A.1 Simple voting mechanisms

We first prove Theorem 3.4 for the special case in which which (H, P) is a voting game, and the DM's decision rule r implements a simple voting rule—namely, $r \equiv r_t$, where $r_t(h) = A$ if and only if t or more agents votes a in history h.

In the following, we slightly abuse notation and denote by $h \in Z$, a terminal history, also the event $(d^T(H(\sigma(s))) = h)$. We also denote by $h_{-i} \in \times_{j \neq i} A_j(\epsilon)$ and $h_i \in A_i(\epsilon)$ the events $(d^T(H(\sigma(s)))_i = h_i)$ and $(d^T(H(\sigma(s)))_i = h_i)$, respectively.

Lemma A.2 *For any player i and $h_{-i} \in \times_{j \neq i} A_j(\epsilon)$ the following holds: If*

$$\Pr\left[\theta = A \mid h_{-i} \cap s_i = b\right] \leq \frac{1}{2}$$

then

$$\Pr\left[\theta = A \mid h_{-i} \cap s_i = a\right] \leq \frac{p^2}{p^2 + (1-p)^2}.$$

The intuition for the lemma is simple. An a signal is equivalent to two a signals and one b signal. So going from the premise of the lemma to the conclusion is analogous to receiving two additional a signals. But two a signals move a prior of 1/2 to a posterior of $p^2/(p^2 + (1-p)^2)$. And of course, one can ignore the conditioning on h_{-i}.

Proof of F: irst, observe that

$$\begin{aligned}
&\Pr\left[s_i = a \cap h_{-i}\right] \\
&\quad = \Pr\left[s_i = a \cap h_{-i} \cap \theta = A\right] + \Pr\left[s_i = a \cap h_{-i} \cap \theta = B\right] \\
&\quad = \Pr\left[s_i = a \mid h_{-i} \cap \theta = A\right] \cdot \Pr\left[h_{-i} \cap \theta = A\right] \\
&\qquad + \Pr\left[s_i = a \mid h_{-i} \cap \theta = B\right] \cdot \Pr\left[h_{-i} \cap \theta = B\right] \\
&\quad = p \cdot \Pr\left[h_{-i} \cap \theta = A\right] + (1-p) \cdot \Pr\left[h_{-i} \cap \theta = B\right]
\end{aligned}$$

since the probability of obtaining signal $s_i = a$ or $s_i = b$ depends only on the state θ, and on h_{-i} also only through θ.

An analogous calculation with b replacing a yields

$$\Pr\left[s_i = b \cap h_{-i}\right] = (1-p) \cdot \Pr\left[h_{-i} \cap \theta = A\right] + p \cdot \Pr\left[h_{-i} \cap \theta = B\right].$$

Next,

$$\begin{aligned}
&\Pr\left[\theta = A \mid h_{-i} \cap s_i = a\right] \\
&\quad = \frac{\Pr\left[s_i = a \mid \theta = A \cap h_{-i}\right] \cdot \Pr\left[\theta = A \cap h_{-i}\right]}{\Pr\left[s_i = a \cap h_{-i}\right]} \\
&\quad = \frac{p \cdot \Pr\left[\theta = A \cap h_{-i}\right]}{\Pr\left[s_i = a \cap h_{-i}\right]} \\
&\quad = \frac{p \cdot \Pr\left[\theta = A \cap h_{-i}\right]}{p \cdot \Pr\left[h_{-i} \cap \theta = A\right] + (1-p) \cdot \Pr\left[h_{-i} \cap \theta = B\right]}. \qquad (1)
\end{aligned}$$

Again, an analogous calculation with b replacing a yields

$$\begin{aligned}
&\Pr\left[\theta = A \mid h_{-i} \cap s_i = b\right] \\
&\quad = \frac{(1-p) \cdot \Pr\left[\theta = A \cap h_{-i}\right]}{(1-p) \cdot \Pr\left[h_{-i} \cap \theta = A\right] + p \cdot \Pr\left[h_{-i} \cap \theta = B\right]}.
\end{aligned}$$

The premise of the lemma,

$$\Pr\left[\theta = A \mid h_{-i} \cap s_i = b\right] \leq \frac{1}{2},$$

thus implies that

$$\frac{(1-p) \cdot \Pr\left[\theta = A \cap h_{-i}\right]}{(1-p) \cdot \Pr\left[h_{-i} \cap \theta = A\right] + p \cdot \Pr\left[h_{-i} \cap \theta = B\right]} \leq \frac{1}{2},$$

and so

$$\Pr\left[\theta = A \cap h_{-i}\right] \leq \frac{p}{1-p} \cdot \Pr\left[\theta = B \cap h_{-i}\right].$$

Plugging this into equation (1) yields

$$\begin{aligned}
&\Pr\left[\theta = A \mid h_{-i} \cap s_i = a\right] \\
&= \frac{p \cdot \Pr\left[\theta = A \cap h_{-i}\right]}{p \cdot \Pr\left[h_{-i} \cap \theta = A\right] + (1-p) \cdot \Pr\left[h_{-i} \cap \theta = B\right]} \\
&\leq \frac{p \cdot \frac{p}{1-p} \cdot \Pr\left[\theta = B \cap h_{-i}\right]}{p \cdot \frac{p}{1-p} \cdot \Pr\left[\theta = B \cap h_{-i}\right] + (1-p) \cdot \Pr\left[h_{-i} \cap \theta = B\right]} \\
&\leq \frac{p^2}{p^2 + (1-p)^2},
\end{aligned}$$

as claimed. ■

Next, we prove the following lemma:

Lemma A.3 *For any $t \in \{1, \ldots, n\}$, if $r \equiv r_t$ then $\beta_D \leq \overline{\beta}(p)$.*

Proof of F: ix t, and assume without loss of generality that in σ, no player has a constant pure strategy in which he always votes for the same alternative regardless of his signal. This is without loss of generality since, given an equilibrium with such players, one can remove them and adjust t and (H, P, d, σ) accordingly. The resulting process would still be feasible, persuasive, and transparent, and r would still implement a simple voting rule.

Observe that for any player i,

$$\Pr\left[\theta = A \mid h_{-i} \cap s_i = b\right] \leq \Pr\left[\theta = A \mid h_{-i} \cap s_i = a\right].$$

Furthermore, this implies that for those h_{-i} in which agent i is pivotal—i.e., ones in which there are exactly $t-1$ other votes a—

$$\Pr\left[\theta = A \mid h_{-i} \cap s_i = b\right] \leq \frac{1}{2}.$$

If the latter were not true, then on both signals a pivotal agent i would strictly prefer to vote for A, and thus in the equilibrium σ he would have a constant strategy.

By Lemma A.2, it thus holds that for each player i

$$\Pr\left[\theta = A \mid h_{-i} \cap s_i = a\right] \leq \frac{p^2}{p^2 + (1-p)^2}.$$

Additionally, observe that

$$\Pr\left[\theta = A \mid h_{-i} \cap s_i = s\right] = \Pr\left[\theta = A \mid h_{-i} \cap s_i = s \cap m_i = m\right]$$

for any $s \in S$ and $m \in M$, since, given the signal s_i, the posterior on θ does not depend on m_i.

Under r_t, the DM prefers A over B whenever t or more players vote a. We will show that this cannot be the case when exactly t vote a. In particular, we will show that conditional on this event, the posterior on $\theta = A$ is at most $\overline{\beta}$.

Consider a history h in which agent i votes a, as well as any $t-1$ other agents. Then the DM's posterior is

$$\begin{aligned}
&\Pr\left[\theta = A \mid h\right] \\
&= \Pr\left[\theta = A \mid h_{-i} \cap m_i = a\right] \\
&= \Pr\left[\theta = A \cap s_i = a \mid h_{-i} \cap m_i = a\right] + \Pr\left[\theta = A \cap s_i = b \mid h_{-i} \cap m_i = a\right] \\
&= \Pr\left[\theta = A \mid h_{-i} \cap m_i = a \cap s_i = a\right] \cdot \Pr\left[s_i = a \mid h_{-i} \cap m_i = a\right] \\
&\quad + \Pr\left[\theta = A \mid h_{-i} \cap m_i = a \cap s_i = b\right] \cdot \Pr\left[s_i = b \mid h_{-i} \cap m_i = a\right] \\
&\leq \frac{p^2}{p^2 + (1-p)^2} \cdot \Pr\left[s_i = a \mid h_{-i} \cap m_i = a\right] + \frac{1}{2} \cdot \Pr\left[s_i = b \mid h_{-i} \cap m_i = a\right] \\
&\leq \frac{p^2}{p^2 + (1-p)^2} = \overline{\beta}.
\end{aligned}$$

Thus, $\Pr\left[\theta = A \mid h\right] \leq \overline{\beta}$. However, if $r \equiv r_t$ then A is an optimal decision on history h, which implies that $\beta_D \leq \overline{\beta}$ as claimed. ■

To prove Theorem 3.4 note that if $\beta_D > \overline{\beta}$ then by Lemma A.3 it holds that $r \not\equiv r_t$ for any $t \in \{1, \ldots, n\}$.

A.2 Normal-form mechanisms

We next prove Theorem 3.4 for normal-form mechanisms—ones in which all histories $h \in A_1(\epsilon) \times \ldots \times A_n(\epsilon)$ are terminal. The structure of the proof is the same as in Section A.1, and in fact Lemma A.2 is true as is. However, Lemma A.3 is not. In particular, Lemma A.3 uses the fact that there exists some agent i and an h_{-i} for which i is pivotal, and such that

$$\Pr\left[\theta = A \mid h_{-i} \cap s_i = b\right] \leq \frac{1}{2}$$

(otherwise agent i would have a constant strategy, which we assumed without loss of generality was not the case). In general normal-form mechanisms this is not longer straightforward, and requires some additional work.

We begin with a definition.

Definition A.4 (monotone decision rule) *A decision rule $r : Z \mapsto \mathcal{O}$ of the DM is* monotone *if:*

- *For each i there is a complete order $\succeq_i$ on elements $A_i(\epsilon)$. For any $h, h' \in Z$ order $h \succeq h'$ if and only if $h_i \succeq_i h'_i$ for all i.*
- *For any $h, h' \in Z$, if $h \succeq h'$ and $r(h') = A$, then $r(h) = A$.*

Consider the following lemma:

Lemma A.5 *In any transparent process, the decision rule r of the DM is monotone.*

Proof: This is a proof outline. For each agent i, the order is such such that $h_i \succeq_i h'_i$ if h_i gives a stronger signal of $s_i = a$ than h'_i. Next, show that if r is not monotone then there are histories $h \succeq h'$ such that $r(h) = B$ and $r(h') = A$, but h yields a higher posterior on $(\theta = A)$ than h'. ■

Next, some notation. Let $\text{piv}_i \subseteq \times_{j \neq i} A_j(\epsilon)$ be the set of pivotal histories for agent i:

$$\text{piv}_i = \{h_{-i} \in \times_{j \neq i} A_j(\epsilon) : \exists h_i, h'_i \in A_i(\epsilon) \text{ s.t. } r(h_i, h_{-i}) = A \text{ but } r(h'_i, h_{-i}) = B\}.$$

Now, to show that there is some $h_{-i} \in \text{piv}_i$ for which

$$\Pr[\theta = A \mid h_{-i} \cap s_i = b] \leq \frac{1}{2},$$

first assume that σ is full support (otherwise just eliminate some actions from the game). Next observe that if there is not h_{-i} as desired, then conditional on being pivotal any of agent i's actions lead to the same probability of getting outcome A (since he always desires this outcome). Furthermore, for any $h_{-i} \in \text{piv}_i$ and any $h_i \neq h'_i$, it must hold that $r(h_i, h_{-i}) = r(h'_i, h_{-i})$ (this holds by monotonicity). However, this implies that i is not pivotal.

The rest of the proof continues as in Section A.1.

A.3 General mechanisms

The case of general mechanisms presents an additional complication to the proof of Theorem 3.4. In normal-form mechanisms, the DM could condition on the realized strategy of the agents, namely h. This is not the case in extensive mechanisms, in which the DM only observes the realized *path of play*. In particular, he cannot distinguish between two different strategies that yield the same path of play.

References

Austen-Smith, D. (1993). Interested experts and policy advice: Multiple referrals under open rule. *Games and Economic Behavior*, **5** 3–43.

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

Battaglini, M. (2002). Multiple referrals and multidimensional cheap talk. *Econometrica*, **70** 1379–1401.

Crawford, V. P. and Sobel, J. (1982). Strategic information transmission. *Econometrica*, **50** 1431–1451.

Dal Bó, E. (2007). Bribing voters. *American Journal of Political Science*, **51** 789–803.

FAC (2015). Federal Advisory Committee. URL `http://www.federalreserve.gov/aboutthefed/fac.htm`.

FDA (2008). Guidance for FDA Advisory Committee members and FDA staff : Voting procedures for advisory committee meetings. URL `http://www.fda.gov/downloads/RegulatoryInformation/Guidances/UCM125641.pdf`.

FDA (2014). FDA Transparency Initiative. URL `http://www.fda.gov/AboutFDA/Transparency/TransparencyInitiative/`.

FDA (2015). URL `http://www.fda.gov/AdvisoryCommittees/`.

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

Felgenhauer, M. and Grüner, H. P. (2008). Committees and special interests. *Journal of Public Economic Theory*, **10** 219–243.

Fingleton, J. and Raith, M. (2005). Career concerns of bargainers. *Journal of Law, Economics, and Organization*, **21** 179–204.

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

Gersbach, H. and Hahn, V. (2008). Should the individual voting records of central bankers be published? *Social Choice and Welfare*, **30** 655–683.

Gershkov, A. and Szentes, B. (2009). Optimal voting schemes with costly information acquisition. *Journal of Economic Theory*, **144** 36–68.

Gillian, T. and Krehbiel, K. (1989). Asymmetric information and legislative rule with heterogenous committees. *American Journal of Political Science*, **33** 459–490.

IAC (2015). Investor Advisory Committee. URL `http://www.sec.gov/spotlight/investor-advisory-committee-2012.shtml`.

IPCC (2015). Intergovernmental panel on climate change. URL `http://www.ipcc.ch/`.

Krishna, V. and Morgan, J. (2001). A model of expertise. *Quarterly Journal of Economics*, **116**.

Levy, G. (2007a). Decision making in committees: Transparency, reputation, and voting rules. *American Economic Review*, **97** 150–168.

Levy, G. (2007b). Decision-making procedures for committees of careerist experts. *American Economic Review*, **97** 306–310.

Mayhew, D. R. (1974). *Congress: The Electoral Connection*. Yale University Press.

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

Osborne, M. J. and Rubinstein, A. (1994). *A course in game theory*. MIT press.

Prat, A. (2005). The wrong kind of transparency. *American Economic Review*, **95** 862–877.

Seidmann, D. J. (2011). A theory of voting patterns and performance in private and public committees. *Social Choice and Welfare*, **36** 49–74.

Shepsle, K. A. and Weingast, B. R. (1987). The institutional foundations of committee power. *The American Political Science Review*, **81** 85–104.

Sibert, A. (2003). Monetary policy committees: Individual and collective reputations. *The Review of Economic Studies*, **70** 649–665.

Snyder, J. M. and Ting, M. M. (2005). Why roll calls? A model of position-taking in legislative voting and elections. *Journal of Law, Economics, and Organization*, **21** 153–178.

Stasavage, D. (2007). Polarization and publicity: Rethinking the benefits of deliberative democracy. *Journal of Politics*, **69** 59–72.

Visser, B. and Swank, O. H. (2007). On committees of experts. *The Quarterly Journal of Economics*, **122** 337–372.

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