DISCLOSURE AND INCENTIVES IN TEAMS*

Paula Onuchic   João Ramos
University of Oxford   USC Marshall

December, 2023
Click here for the most recent version.

Abstract

We consider a team production environment augmented by a stage in which the team decides how to communicate its productive outcome to outside observers. In this context, we characterize equilibrium disclosure of team outcomes when team disclosure choices aggregate individual recommendations through some deliberation procedure. We show that equilibria often involve partial disclosure of the team's outcome and establish a relation between the deliberation procedure and the observer's equilibrium attribution of blame for non-disclosed outcomes ("team failures") across team members. We show that through this blame-attribution channel a team's deliberation procedure determines individuals' incentives to contribute effort to team production. We then characterize deliberation procedures that provide strong effort incentives in different productive environments.

1 Introduction

Productive activities are increasingly conducted in teams. Startups are often founded by entrepreneurial partners, and in established firms new products are proposed and developed mainly

*Paula Onuchic: paula.onuchic@economics.ox.ac.uk. João Ramos: joao.ramos@usc.edu. We are grateful for detailed comments and suggestions from Nageeb Ali, Rohan Dutta, Robert Gibbons, Sam Kapon, Navin Kartik, Elliot Lipnowski, Meg Meyer, Harry Pei, Michael Powell, Debraj Ray, Ludvig Sinander, Caroline Thomas, and Mark Whitmeyer. We also thank the seminar audiences at several institutions, as well as the 2^{nd} Southeast Theory Festival at Nuffield College, 2023 ESSET - Gerzensee, 2023 SAET - Paris, 2023 SITE - Dynamic Games, Contracts and Markets, and 2023 Transatlantic Theory Workshop.

by teams built and empowered within the company.[1] In policy-making or regulatory scenarios, most investigation and evidence-gathering that informs decision-making is done by committees.[2] In all these contexts, team production is typically followed by a disclosure stage in which the team communicates the outcome of its production to outsiders. Entrepreneurial partners decide whether and when to pitch new startups to investors; within-firm teams report their projects' progress in regular meetings with managers; and congressional committees abide by formal rules guiding the publication of reports, gathered evidence, and meeting transcripts.[3]

These communications are collective disclosure decisions that have potentially distinct implications for the individuals in a team. For example, if a development team composed of engineers and marketers puts out a technically impressive but "badly packaged" new product, then such a disclosure is seen as an engineering success but has negative reputational implications for the marketing team members. If, instead, the team chooses to delay the product launch, then a skeptical observer interprets this non-disclosure as a team failure. And the degree to which each co-developer is blamed for this negative collective outcome depends on who the observer sees as responsible for the decision to delay the launch.

This paper proposes a model that incorporates such a disclosure stage into a team-production environment. The main primitive in our model is the deliberation procedure that the team uses to reach a collective disclosure decision. The procedure determines the allocation of *voice rights* across team members — a voice right is defined by Zuckerman (2010) and Freeland and Zuckerman Sivan (2018) as "the right to speak on behalf of an organization." We show that the allocation of such rights impacts how undisclosed team outcomes are perceived by the outside observer: individuals who are more responsible for the team's non-disclosure decision are necessarily regarded with more skepticism. Through this channel, the allocation of voice rights within a team at the disclosure stage determines individuals' incentives to contribute to the team at the production stage. We leverage this insight to design team-deliberation protocols

[1] Tamaseb (2021) documents that 80% of all billion-dollar companies launched since 2005 had two or more founders. Lazear and Shaw (2007) shows that close to 80% of US firms rely on self-managing teams in some capacity. Recent literature also documents the rise of teamwork in scientific research. See, for example, Fortunato et al. (2018) and Jones (2021).

[2] These include both formal government bodies such as congressional committees and minipublics or other mechanisms of citizen participation (described, for example, in Bardhi and Bobkova (2023)).

[3] In the United States, at the start of congress, committees adopt and publish procedural rules that determine, among other things, guidelines for communications with the public. These guidelines vary across committees. For example, in 2017-18, the procedures for the Special Committee on Aging determined that "committee findings and recommendations shall be printed only with the approval of a majority of the committee"; the Committee on Commerce, Science and Transportation resolved that "public hearings of the full committee, or any subcommittee thereof, shall be televised or broadcast only when authorized by the chairman and the ranking minority member of the full committee"; and the Select Committee on Ethics decided that the release of reports to the public be determined by either the chairman or the vice-chairman, who were thus given the authorization to speak on behalf of the committee. Quotes are taken from the published "Authority and Rules of Senate Committees, 2017-18."

that effectively incentivize individual effort provision.

Section 2 introduces the team disclosure environment. A team is made up of two or more members who produce a team outcome, drawn from a distribution known by all team members and by an outside observer.[4] The team outcome conveys to the observer some information about the team members, or about a state relevant to the team members. For instance, the pitch of a new product by a development team conveys to an observer some information about the ability of each team member, who may be motivated by career concerns; and information gathered by a committee indicates the aptness of a new policy that may be supported by some partisan committee members, but not others. After seeing their outcome, the team members decide whether to reveal it to the observer. If the team chooses to disclose, then each team member receives their respective individual value implied by the outcome. If the team chooses not to disclose, then the observer forms a rational *no-disclosure belief* about each individual's value, accounting for the circumstances that may have led the team not to reveal the outcome. Each individual's payoff is equal to the observer's no-disclosure belief about their respective value.

To reach a team disclosure decision, team members' individual disclosure recommendations are aggregated into a team decision via a mapping, which we refer to as the deliberation procedure. The mapping is a given primitive of the team disclosure environment that all team members, as well as the observer, understand. To each possible set of team members that recommends the outcome's disclosure, the deliberation procedure assigns a team disclosure decision. For example, a team's deliberation procedure may be such that every team member can unilaterally decide that the outcome be disclosed, in which case a single disclosure recommendation implies that the team's decision is to disclose. On the other hand, the procedure may require that disclosure decisions be made via consensus so that a single recommendation not to disclose induces the team to choose no disclosure. These and many other natural benchmarks are encompassed by our framework. We assume only that the procedure respects unanimous decisions among the team members and monotonicity so that when disclosure is favored by a larger set of team members, the team discloses its outcome with a higher probability.

Our first result, Theorem 1, describes how the equilibrium set of the team disclosure game depends on the deliberation procedure used by the team. For any deliberation procedure, a full disclosure equilibrium exists in which the observer interprets non-disclosure with *maximal skepticism* about every team member.[5] More interestingly, if the deliberation procedure is such

[4]To study the team-disclosure problem in Section 2, we take the team-outcome distribution as an exogenous model primitive. In Section 4, the outcome distribution is endogenously determined by individual effort contributions to team production.

[5]Formally, we say that no-disclosure beliefs are maximally skeptical about a team member if they indicate that the realized team outcome implied the worst possible value for that team member (in the support of the team outcome distribution).

that not all team members can unilaterally choose disclosure, then the equilibrium set also includes an equilibrium in which some team outcomes are not disclosed. Specifically, the team conceals *team failures*, outcomes that are sufficiently bad news about a sufficiently large set of team members.

This result is in contrast with classic findings from the single-agent evidence disclosure literature following Grossman (1981) and Milgrom (1981), which describe the *unravelling* of equilibria in which some evidence is not disclosed and show that full disclosure is the unique equilibrium outcome. We show instead that, in a team disclosure setting, the unravelling logic is broken because the observer is unable to definitively attribute a team's decision to not disclose an outcome to particular individuals in the team. Equilibrium non-disclosures are met by the observer with skepticism about the team, which does not automatically translate into skepticism about each individual and may not be sufficient to ensure unravelling. Theorem 1 also shows that, in any equilibrium without full disclosure, the observer's no-disclosure beliefs must be maximally skeptical about each team member who can unilaterally choose disclosure, but the observer need not be maximally skeptical about team members who cannot do so.

Our next results — Proposition 1 and its Corollaries 1 and 2 — further characterize the relationship between the deliberation procedure used by the team and the observer's equilibrium no-disclosure beliefs about each individual's value.[6] We show that a change in the deliberation procedure that strengthens the relation between an individual's recommended disclosure action and the team's disclosure decision leads to a decrease in the observer's no-disclosure beliefs about that individual's value. One interpretation of the observer's equilibrium no-disclosure beliefs is as a measure of the observer's perception of each individual's *blame* for a team's failure. In that light, our results show that an increase in an individual's voice rights also increases that individual's equilibrium blame when the team fails.

Theorem 2 completes our characterization of team disclosure equilibria. Under any deliberation procedure, full disclosure is an equilibrium, sometimes supported by off-path beliefs that are not required to satisfy any consistency criterion. For instance, full disclosure equilibrium can always be supported by the observer holding maximally skeptical no-disclosure beliefs about all individuals. In definition 2, we introduce a refinement requiring no-disclosure beliefs, even if off-path, to be justified by the aggregation of individual disclosure recommendations through the deliberation procedure used by the team. The Theorem shows that full disclosure is *consistent with deliberation* if and only if the deliberation procedure is such that disclosing the

[6]These results connect to previous results in the literature on single-agent evidence disclosure games that establish a relation between model primitives, such as sender preferences and the evidence structure, and equilibrium observer skepticism. In a team disclosure context, we introduce and characterize a new individually targeted notion of skepticism, an individual's blame.

team's outcome requires less consensus than concealing it.[7]

Section 4 studies the full team-production and team-disclosure problem. We augment the environment from Section 2 with an initial stage in which team members choose whether to covertly exert costly effort to improve the team's outcome distribution. Importantly, each team member's effort increases not only the value of the team outcome to themselves, but also the value of the team outcome to the other team members. As before, once the team outcome realizes — now drawn from a distribution that depends on the team members' effort profile — the team chooses whether to reveal it to the outside observer. Our main results in this section evaluate the effort incentives provided by different deliberation protocols, through their effect on equilibrium disclosure of team outcomes.

For a given team-disclosure strategy, Lemma 3 shows that an individual's incentive to exert costly effort can be decomposed in two parts: an "individual effort benefit," which compares the individual's own expected outcome value if the individual exerts effort versus if they do not; and a "blame misattribution" component. The blame misattribution component is the novel incentive mechanism introduced by strategic team disclosure. In a full-effort equilibrium, the observer attributes blame for team failures under the assumption that each individual exerted effort. If an individual shirks and the team draws a team failure, then blame for that failure is calculated under the (false) equilibrium premise that all team members exerted effort. Blame may thus be misattributed across the team members, with each individual potentially being over- or under-punished for a group failure. Excessive (misattributed) blame for team failures provides individuals with extra incentives to contribute effort to team production.

Proposition 2 uses this insight to characterize deliberation procedures that induce extra effort incentives through blame misattribution. We show that, if the productive environment is such that effort has low team externalities — in the sense that gains from an individual's effort accrues mostly to themselves — then the deliberation procedure that gives every individual the right to unilaterally disclose the team's outcome provides stronger effort incentives than any other procedure. On the other hand, if effort has high team externalities — that is, effort benefits accrue mostly to an individual's fellow team members and not to themselves — then a protocol that requires disclosure decisions to be reached via consensus dominates those in which disclosure decisions are made unilaterally.[8] Proposition 4 considers environments in which effort strongly improves the correlation between all team members' outcomes (think, for instance, about individuals putting effort towards a common output component). It shows that protocols that require more consensus for disclosure provide more effort incentives, relative to

[7] This notion is formalized in Definition 2.

[8] Our distinction between high and low team externalities environments parallels the distinction between selfish and cooperative investments in a hold-up context, proposed by Che and Hausch (1999).

the unilateral disclosure protocol.

The disclosure equilibria induced by these effort-enhancing protocols can be connected to "corporate cultures" often praised in the business literature. The full-disclosure equilibrium induced by the unilateral-disclosure protocol parallels "radically transparent" organizations, in which individuals are held fully accountable for their contributions to team failures. An article titled "How to Win the Blame Game" in the Harvard Business Review praises transparency and the benefits of a "well-managed blame culture," stating that "when used judiciously (...) blame can prod people to put forth their best efforts." In contrast, the partial-disclosure equilibrium induced by consensus-disclosure protocols resembles a corporate culture in which teams "don't play the blame game," but rather collectively suffer the burden of bad team outcomes — for example, much of the technology world uses a "blameless postmortem" approach to understand the causes of team failures. Also in the Harvard Business Review, the article "When Transparency Backfires, and How to Prevent It" argues that too much transparency can create a blaming culture that "may actually decrease constructive, reciprocal behavior between employees." From the perspective of the result in Proposition 2, both the radically transparent culture and the "no blame game" cultures — induced through the different deliberation protocols that teams may use — can be useful effort-incentivizing tools when applied in the correct productive environment. The former should be employed in low-group-externality teams, while the latter is beneficial in high-team-externality environments.

So far, we have compared deliberation procedures in terms of their effort-incentives provision, but have not characterized "effort-maximizing deliberation procedures." In Section 5, we consider that characterization while restricting attention to binary-outcome environments (in which the support of each individual's value consists of a "high value" and a "low value"). Our results in this binary environment are in line with the general intuition developed in Propositions 2 and 4. We show that effort-maximizing deliberation procedures are such that disclosure requires more consensus in environments with higher effort externalities or in which effort more strongly correlates team members' outcome values.

1.1 Related Literature

Our paper contributes to a large literature on multi-sender communication. Using different communication protocols, seminal contributions by Milgrom and Roberts (1986), Battaglini (2002), and Gentzkow and Kamenica (2016) study models in which multiple senders communicate with a single receiver.[9] All of those papers consider environments in which senders "competi-

[9]Several more recent contributions, including Hagenbach et al. (2014), Hu and Sobel (2019), and Baumann and Dutta (2022), also study models of multi-sender evidence disclosure.

tively" communicate with a receiver by unilaterally sending messages to the same receiver. This competitive communication benchmark corresponds to the unilateral disclosure protocol in our context. Our paper expands on that by considering *communication by a group*, where group members aggregate their individual communication recommendations using some deliberation procedure.

In our model, the team communicates using an evidence disclosure protocol, as in the large literature stemming from Grossman (1981) and Milgrom (1981). We contribute to this literature by studying team disclosure. Among other results, we show that team-disclosure equilibria often feature non-disclosure of some evidence. Our paper is especially close to models with multidimensional evidence, such as Dziuda (2011) and Martini (2018). In particular, Martini (2018) shows that if a single sender separably values the receiver's posterior about each dimension of the state, then partial-disclosure equilibria may exist if the sender's preferences are sufficiently convex. Such equilibria are supported by the fact that, upon seeing no disclosure, the receiver cannot distinguish the dimension in which the sender drew "bad news." Despite the intuitive connection between this characterization and ours, team-disclosure problems are inherently different from individual multi-dimensional disclosure problems, and the former cannot be mapped into instances of the latter through appropriately chosen sender preferences.

The extensive literature on single-agent disclosure games provides a variety of mechanisms that prevent the "unravelling" result from Milgrom (1981).[10] In particular, Dye (1985) observes that equilibria that do not feature full evidence disclosure exist in a single-agent problem when the observer is unsure whether the sender has access to evidence. In the team context, despite senders always having access to evidence, they may be unable to disclose it because other team members may have vetoed it. The observer in our context is, as in Dye (1985), unsure about why the evidence was not disclosed. Our mechanism is also connected to that in Seidmann and Winter (1997) and Giovannoni and Seidmann (2007), who argue that equilibria with some non-disclosure arise when, upon seeing no-disclosure, the observer is unsure whether the sender intended to bias their belief upwards or downwards. In the team context, the observer is similarly unable to attribute the decision to not disclose the outcome to the interests of a particular individual in the team.

Squintani (2020) studies strategic transmission of verifiable information in networks of (perhaps biased) experts and decision makers. He shows that communication to a decision maker through a path of intermediary players—who can sequentially choose to "veto" information transmission—breaks down unless all intermediaries are biased in the same direction relative

[10] See, for example, Dranove and Jin (2010) for a review of both theoretical and empirical explanations of why verifiable information may not be voluntarily disclosed through a process of unravelling.

to the decision maker.[11] One of the deliberation procedures we consider in our team disclosure model is the consensus procedure, in which every team member can veto disclosure. We similarly find that communication between the team and the observer is harmed (although not completely broken down) because deliberation aggregates preferences from multiple agents. In a model of intermediated cheap talk, Ambrus, Azevedo and Kamada (2013) likewise show that intermediaries' competing biases may harm information transmission.

Our paper considers not only disclosure decisions in teams but also the impact of communication decisions on individual effort incentives. Thus, we contribute to a small literature relating disclosure and incentives. Ben-Porath, Dekel and Lipman (2018) show that in a Dye (1985) individual-disclosure environment, partial disclosure equilibria may incentivize the individual to favor risky projects, even at the expense of the project's overall expected value. Matthews and Postlewaite (1985), and more recently Shishkin (2021), Onuchic (2021), and Whitmeyer and Zhang (2022), study the effect of the evidence-disclosure equilibrium on an individual's incentives to acquire evidence. A closely related literature — for example, Austen-Smith and Feddersen (2005), Gerardi and Yariv (2007, 2008), Levy (2007), Visser and Swank (2007), and more recently Name-Correa and Yildirim (2019) and Bardhi and Bobkova (2023) — studies information acquisition and information aggregation in deliberative committees, under various voting and communication protocols, as well as committee compositions. Our paper first departs from that literature in that we study a model of communication by a group, rather than an environment in which an action choice is delegated to the group. More importantly, our paper differs from that literature in its focus: while the deliberative committees literature studies how different protocols fare in terms of information acquisition and aggregation, we characterize disclosure equilibria under various protocols and evaluate their power to incentivize team members to put effort into a team project.

Our work is also related to the literature on incentives provided by career concerns, following Holmström (1999), and specifically to papers on career concerns in teams.[12] More generally, our results on incentive provision relate to the large literature on incentives in teams —

[11] Squintani (2020) uses this observation as a step in his main exercise, in which he studies the design of networks of ideologically differentiated experts and decision makers, with the goal of maximizing the flow of information. He finds that the optimal network is a line, with players ordered by their bias. The exercise of designing a communication network to facilitate information flow may be related to designing deliberation procedures to maximize information disclosure. In contrast, in our paper, we design deliberation procedures to improve team members' effort incentives.

[12] See, for instance, Jeon (1996), Auriol, Friebel and Pechlivanos (2002), Bar-Isaac (2007), Arya and Mittendorf (2011), Chalioti (2016), and Ramos and Sadzik (2023). In particular, Ortega (2003) shows that the power allocation within a team — the distribution of how individual efforts affect team outcomes — affects effort incentives; and Onuchic ⓡ Ray (2023), Ray ⓡ Robson (2018), and Ozerturk and Yildirim (2021) study team production with unequal and endogenous credit attribution to team members.

following, for example, Alchian and Demsetz (1972), Holmström (1982), and Itoh (1991). Our contribution is to study the design of the deliberation procedure determining voice rights within a team; and to show that authority over team communication can be used as a motivational tool. Specifically, we show that it may be gainful to share power among team members, inducing an equilibrium whereby blame for team failures cannot be attributed across team members.[13]

2 Team Disclosure

2.1 Environment

A group $N = \{1, 2, ..., n\}$ of agents makes up a team. The team draws an outcome $\omega = (\omega_1, ..., \omega_n)$ from a distribution μ over a finite outcome space $\Omega \subset \mathbb{R}^n$. Once the outcome ω is realized, the team decides whether to disclose it to an outside observer. For each $\omega \in \Omega$ and $i \in N$, ω_i is the value to individual i of having the observer see outcome ω. This general formulation allows for different interpretations of how individual value is implied by a collective team outcome. In a career concerns interpretation, the team produces a joint outcome that informs the observer about each team member's ability, and ω_i is the value derived by i from that assessment. Alternatively, the outcome may be a piece of evidence that conveys to the observer information about some state of the world relevant to a policy decision. The state of the world and the implied policy decision may be valued differently by each team member i, yielding the individual value described by ω_i.

We make the following assumption about the outcome distribution μ.

Assumption 1. *The outcome distribution μ has a product support—that is, $\Omega = \Omega_1 \times ... \times \Omega_N$—where $\Omega_i \subset \mathbb{R}$ has at least 2 elements for all $i \in N$, and μ has full support over Ω.*

Assumption 1 reflects that team outcomes generate heterogeneous values across the individuals in the team. It implies, for example, that individual outcome values are not perfectly correlated, in which case the team would behave as if it were a single individual (see our discussion in Observation 1 below). We note, however, that Assumption 1 is a very weak restriction, as any distribution can be arbitrarily well approximated by a full-support distribution.

Each team member i, after seeing the outcome ω, makes an individual disclosure recommendation: $x_i(\omega) \in [0, 1]$ indicates the probability that agent i recommends the disclosure of the outcome. With complementary probability $1 - x_i(\omega)$, agent i recommends that the outcome be concealed from the outside observer. Individual disclosure recommendation strategies define

[13] In a single-agent career-concerns environment, Dewatripont, Jewitt and Tirole (1999) show that when the observed outcome is a coarser signal about an individual's abilities, effort incentives may actually be improved.

a distribution over the set of team members who favor the disclosure of each possible outcome $\omega \in \Omega$. Formally, for every subset of team members $X \subseteq N$,

$$\Pi_X(\omega) = \prod_{i \in N} x_i(\omega)^{\mathbb{1}[i \in X]}(1 - x_i(\omega))^{\mathbb{1}[i \notin X]}$$

is the probability that the set of team members who favor the disclosure of outcome ω is X.[14]

The teams' disclosure decision is then made according to some *deliberation procedure*. A deliberation procedure $D : \mathcal{P}(N) \to [0, 1]$ is a mapping that aggregates individual recommendations into a team disclosure decision. If $X \subseteq N$ is the set of team members who recommend disclosure, then $D(X)$ is the probability that the outcome is disclosed. Now considering the distribution of disclosure recommendations, the team's disclosure decision

$$d(\omega) = \sum_{X \subseteq N} \Pi_X(\omega) D(X) \in [0, 1]$$

is the expected probability that the outcome ω is disclosed to the outside observer.

In a real-world scenario, deliberation may be a lengthy process made up of formal rules and communication between team members which somehow aggregates the interests of the group into a team decision.[15] In this model, we interpret our deliberation protocol D as a reduced-form aggregation rule that already accounts for that interplay and informs how individual recommendations map into a team decision. We assume that this protocol agrees with unanimous team decisions, so that if all team members recommend disclosure or all team members recommend non-disclosure, then the unanimous decision is followed. And we require that the probability of disclosure be increasing in the set of people who favor disclosure. Formally:

Assumption 2. *The deliberation process* $D : \mathcal{P}(N) \to [0, 1]$

1. *Respects unanimity:* $D(N) = 1$ *and* $D(\varnothing) = 0$.
2. *Is monotone:* $X \subseteq X'$ *implies* $D(X) \leqslant D(X')$.

For illustration, suppose that the team has only two members, so that $N = \{1, 2\}$. Because $D(\{\varnothing\}) = 0$ and $D(\{1, 2\}) = 1$, the deliberation procedure is fully defined by $D(\{1\}) \in [0, 1]$—the probability that a team discloses an outcome when person 1 recommends its disclosure and person 2 does not—and by $D(\{2\}) \in [0, 1]$, the probability of disclosure when it is

[14] Note that if every team member uses a pure recommendation strategy (so that $x_i(\omega) \in \{0, 1\}$), then $\Pi_X(\omega) = 1$ if X is exactly the set of team members who recommend disclosure, and $\Pi_X(\omega) = 0$ otherwise.

[15] Indeed, previous literature — such as Gerardi and Yariv (2007) — highlights the interplay of formal rules and communication in shaping equilibrium behavior in a deliberative committee.

supported only by team member 2. The set of possible deliberation procedures for a two-person team is accordingly depicted in the two panels in Figure 1.

The figure also highlights some possible features of deliberation procedures. We say that team member i can *unilaterally choose disclosure* if $D(\{i\}) = 1$, so that the team discloses its outcome even if only team member i recommends that decision. We accordingly denote by *unilateral* the deliberation procedure whereby all team members can unilaterally choose disclosure, so that $D(\{i\}) = 1$ for every $i \in N$. In contrast, we say that disclosure decisions are made via *consensus* if $D(X) = 1$ if $X = N$, and $D(X) = 0$ if $X \neq N$, that is, the outcome is disclosed only if every team member favors its disclosure. Both these protocols are highlighted in the left-hand panel of Figure 1. The right-hand panel displays procedures whereby each of the team members can unilaterally choose disclosure. In particular, we highlight the two possible team-leader protocols: team member i is a team leader if $D(X) = 1$ if $i \in X$ and $D(X) = 0$ if $i \notin X$, implying that the team always follows i's recommended action.

Once the team makes its disclosure decision, the outcome is accordingly seen/not seen by the outside observer, who then forms a posterior belief about the outcome that led to that observation. If ω is disclosed, then the observer perfectly understands it and team member i's value is equal to the realized ω_i, for each $i \in N$. If instead ω is not disclosed, then i's value is equal to the observer's inference about their value. Specifically, i's payoff equals the observer's mean posterior about ω_i, given by

$$\omega_i^{ND} \equiv \mathbb{E}\left[\omega_i | \text{no disclosure}\right] = \frac{\sum_{\Omega} \omega_i (1 - d(\omega))\mu(\omega)}{\sum_{\Omega} (1 - d(\omega))\mu(\omega)}, \tag{1}$$

for each $i \in N$, if $\sum_{\Omega}(1 - d(\omega))\mu(\omega) > 0$. Note that if no disclosure is an off-path (measure zero) event, then the observer's mean posterior is indeterminate. Throughout the paper, we refer to ω_i^{ND} as the observer's *no-disclosure belief* about team member i.

2.2 Equilibrium

Definition 1 (Equilibrium)**.** *Given a deliberation procedure D, individual recommendations x_i for each $i \in N$ and no-disclosure posteriors ω_i^{ND} for each $i \in N$ constitute an equilibrium if*

1. *Individual disclosure recommendations are as if pivotal:*

   For each $i \in N$ and $\omega \in \Omega$, $\omega_i > \omega_i^{ND} \Rightarrow x_i(\omega) = 1$ and $\omega_i < \omega_i^{ND} \Rightarrow x_i(\omega) = 0$.

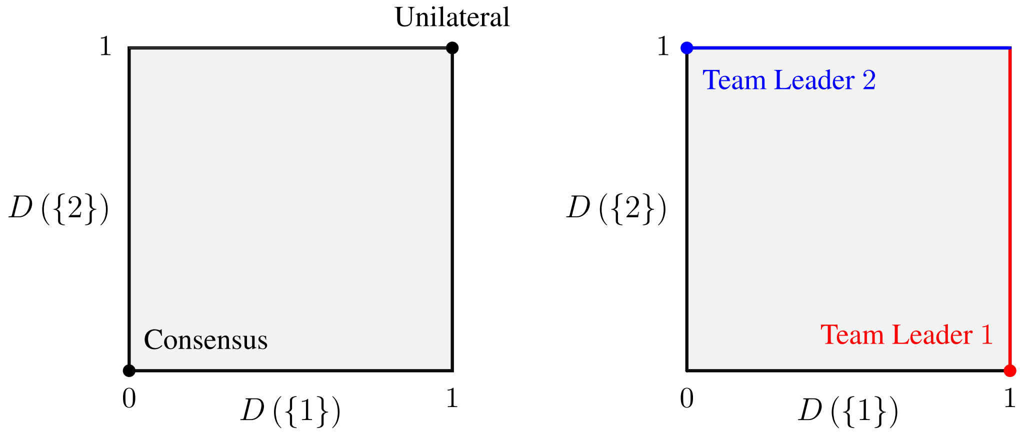


Figure 1: The set of deliberation procedures for two-person teams. The left-hand panel highlights the *unilateral* procedure, in which $D(\{1\}) = D(\{2\}) = 1$ and the *consensus* procedure, in which $D(\{1\}) = D(\{2\}) = 0$. The right-hand panel highlights in red the procedures such that team member 1 can unilaterally choose disclosure, and in blue those in which team member 2 can unilaterally disclose.

2. *Individual disclosure recommendations are determined by own outcome values:*

$$\textit{For each } i \in N \textit{ and } \omega, \hat{\omega} \in \Omega,\ \omega_i = \hat{\omega}_i \Rightarrow x_i(\omega) = x_i(\hat{\omega}).$$

3. *No-disclosure posteriors are Bayes-consistent: for each $i \in N$, ω_i^{ND} satisfies (1).*

The equilibrium notion described above is a Perfect Bayesian Equilibrium in which all team members and the outside observer understand the deliberation process. Additionally, we require that individuals make disclosure recommendations as if they are pivotal to the team's decision. With this requirement, we refine out equilibria in which individuals position themselves for/against disclosure solely because they believe themselves not to be pivotal, and the equilibrium strategies indeed support that their recommendations are not pivotal. Condition 2 requires team members to condition their individual disclosure recommendations only on their own value of the team outcome. Given condition 1, this requirement binds only when an outcome realization is such that $\omega_i = \omega_i^{ND}$ for some team member i; in that case, in which i is indifferent between disclosure or non-disclosure, we impose that i does not correlate their disclosure recommendation with the outcome's value to other team members.[16] Like condition 1, condition 2 imposes a refinement on the equilibrium set. Finally, condition 3 requires

[16] All results stated in section 3 also hold if equilibrium is defined without the requirement of Condition 2. This condition is used for results in sections 4 and 5 that require a characterization of the entire equilibrium set in the team disclosure game. Further note that Condition 2 would be moot if outcome values were distributed continuously, so that realizations in which $\omega_i = \omega_i^{ND}$ for some team member i would constitute a set with zero measure.

Bayes-consistency for beliefs reached on the equilibrium path.

We know from the previous literature on disclosure with verifiable information that when disclosure decisions are made by a single individual, the unique equilibrium has full disclosure — so that all outcomes are disclosed to the observer. The key insight supporting this result is that, if the observer knows that an individual holds some evidence of their outcome, then the non-disclosure of that evidence makes the observer skeptical. The observer's skepticism then generates an *unraveling* of any equilibrium with (partial) non-disclosure. We first remark that in our environment, if all team members have perfectly correlated outcomes, then the team-disclosure game is equivalent to a disclosure problem for a single individual (regardless of the deliberation procedure).

Observation 1. *Suppose that μ is such that outcomes are perfectly correlated across team members. Then for any deliberation protocol D, the unique equilibrium outcome is full disclosure.*

Observation 1 highlights the differences between individual- and team-disclosure problems. In contrast to individual-disclosure problems, our results show that when disclosure decisions are made by teams, equilibria often involve partial non-disclosure. We will see that the usual unraveling argument fails if the outside observer cannot fully attribute a non-disclosure decision to a specific team member.

3 Equilibrium Team Disclosure

Given deliberation procedure D, some team members have complete "voice rights" and can unilaterally choose disclosure on behalf of the team. Theorem 1 shows that equilibrium team disclosure distinguishes between these team members, to whom the observer fully attributes the team's decision to not disclose an outcome, and those members who cannot unilaterally choose disclosure.

We say that an equilibrium has full disclosure if the observer can always perfectly infer the realized outcome $\omega \in \Omega$ on path; or, equivalently, if there is at most one $\omega \in \Omega$ such that $d(\omega) < 1$.

Theorem 1. *Given a deliberation procedure, the following is true about the equilibrium set:*

1. *A full-disclosure equilibrium exists, with*

$$\omega_i^{ND} = \min(\Omega_i) \textit{ for every } i \in N.$$

2. *If i is a team member who can unilaterally choose disclosure, then*

$$\omega_i^{ND} = \min(\Omega_i) \text{ in every equilibrium without full disclosure.}$$

3. *Conversely, if $I \subseteq N$ is the set of team members who cannot unilaterally choose disclosure, then there exists an equilibrium without full disclosure where*

$$\omega_i^{ND} > \min(\Omega_i) \text{ for every } i \in I.$$

A full proof of Theorem 1 is available in Appendix A. The first statement in the Theorem argues that regardless of the deliberation procedure D with which the team makes disclosure decisions, a full-disclosure equilibrium always exists. To see this, suppose that the observer's no-disclosure beliefs satisfy $\omega_i^{ND} = \min(\Omega_i)$ for every team member $i \in N$. That is, upon seeing no disclosure, the observer is *maximally skeptical* about all team members and believes that the realized outcome surely corresponds to the worst possible realization for all of them. In that case, every team member is willing to recommend the disclosure of all outcomes — for no outcome yields a strictly worse payoff than the observer's no-disclosure belief — and, consequently, the team's decision to disclose all outcomes is unanimous. In turn, because all outcomes are disclosed, no-disclosure happens only off the equilibrium path and therefore the observer's beliefs are consistent with Bayes updating.

More interestingly, the Theorem further describes the equilibrium set. Specifically, it states that if a team member i can unilaterally choose disclosure, then the observer must be maximally skeptical about their outcome in any team-disclosure equilibrium. But a converse also holds: there is an equilibrium in which the observer is *not* maximally skeptical about any team member who cannot unilaterally disclose the team's outcome.

For an illustration, refer to Figure 2. In each panel, the figure depicts the space of possible team outcomes in a two-person team, with values for agent 1 plotted on the x-axis and values for agent 2 plotted on the y-axis. Conjecture an equilibrium in which the observer is not maximally skeptical about either team member, so that $\omega_1^{ND} > \min(\Omega_1)$ and $\omega_2^{ND} > \min(\Omega_2)$. Given this conjecture, each team member recommends the disclosure of an outcome if and only if their realized value is larger than the conjectured no-disclosure beliefs about their value. In both panels, the red-shaded area depicts team outcomes that agent 1 recommends to conceal and the blue-shaded area represents those that agent 2 recommends to conceal. These individual recommendations are then aggregated according to the team's deliberation procedure. The left-hand side panel supposes that team member 1 is the team leader, and the one on the right-hand

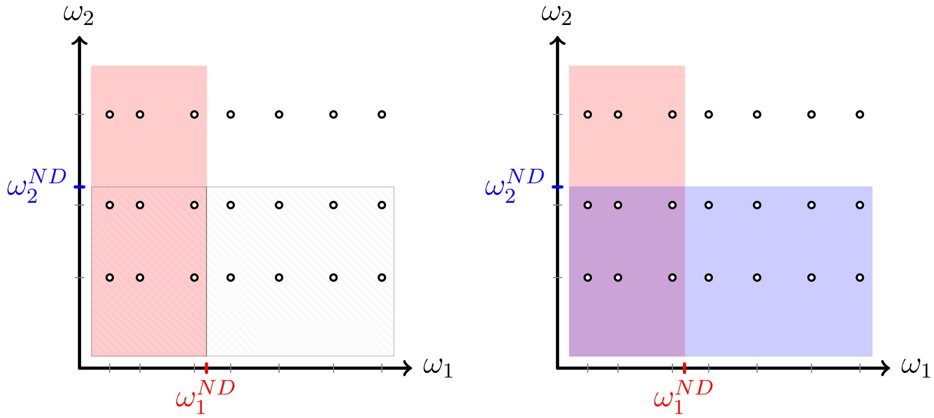


Figure 2: Both panels depict *candidate* team-disclosure equilibria in a team with $n = 2$. The left-hand panel supposes a deliberation protocol in which team member 1 is the team leader. The right-hand panel supposes a protocol whereby disclosure decisions are made via consensus.

side supposes that the deliberation procedure is such that disclosure is chosen via consensus.

Under the former deliberations procedure, the team's decision follows precisely team member 1's recommendations, and therefore the team does not disclose outcomes in the red-shaded area, but discloses all other outcomes. The no-disclosure recommendation in the blue-shaded area is not followed by the team, as represented by the dashed pattern in the figure. Now note that all outcomes in the red-shaded no-disclosure area are such that ω_1 is smaller than the originally conjectured no-disclosure belief ω_1^{ND}. Therefore it cannot be that the conjectured belief is Bayes-consistent; and consequently, there is no equilibrium in which $\omega_1^{ND} > \min(\Omega_1)$, as described in statement 2 of the Theorem. This argument is precisely describes that, if a team member i can unilaterally choose disclosure, then any equilibrium in which the observer is not fully skeptical about i "unravels."

Instead, if disclosure is chosen via consensus, so that neither team member can unilaterally choose disclosure, then both the red- and the blue-shaded areas are not disclosed by the team, for at least one of the team members recommends that each of those outcomes be concealed. In that case, there are some outcome realizations that are "good news" for team member 1 ($\omega_1 > \omega_1^{ND}$) which are not disclosed because team member 2 favors their concealment; and likewise the disclosure of some "good news" about individual 2 is blocked by team member 1. A consequence is that the Bayes-consistent update made by the observer upon seeing no-disclosure is not necessarily lower than the initially conjectured no-disclosure beliefs. Indeed, statement 3 in the Theorem shows that there is such an initially conjectured pair of no-disclosure beliefs—not maximally skeptical about either team member—that satisfies Bayes-consistency. More broadly, the Theorem shows that when team members cannot unilaterally choose to disclose the team's outcome, an equilibrium in which some outcome realizations are not disclosed is supported by the fact that the observer is not able to attribute responsibility for the non-disclosure of the team's outcome to any particular team member.

In such equilibria without full disclosure, the team conceals outcomes that are "bad news" about a set of people in the team who can collectively block the outcome's disclosure — this is the team-equivalent of "sanitization strategies," in the language of Song (1994). Accordingly, upon seeing that an outcome is not disclosed, the observer becomes skeptical about the team and interprets non-disclosure as a *team failure*. The observer, understanding the deliberation procedure that led to the team's decision, rationally translates their "aggregate skepticism" about the team into "targeted skepticism" about individual team members. In other words, the deliberation procedure determines the extent to which the observer perceives each team member as being to blame for a team's failure. Section 3.1 formally introduces a notion of individual *blame* and characterizes its relationship to the deliberation procedure used by the team. Section 3.2 complements our characterization of the team-disclosure equilibrium set by discussing when the full-disclosure equilibrium can be deemed "inconsistent" with the deliberation procedure.

3.1 Blame as Targeted Skepticism

We see the vector of the observer's no-disclosure beliefs ω^{ND} as describing each individual's equilibrium level of *blame* for a team failure — remember that the observer interprets "non disclosure" as a team failure. In our exercise in this section, we will refer to decreases in ω_i^{ND} as increases in team member i's blame. Note that our measure is not normalized by the total aggregate blame directed at the team (as for example $\omega_i^{ND} / \sum_{j \in N} \omega_j^{ND}$).[17] Our choice reflects the fact that, in a Bayesian learning model like ours, blame is not a "zero-sum game." Agents are not just "sharing a pie," but rather each one is signaling their outcome realizations to the observer. Thus, it is possible for all team members to be fully blamed for a group failure, or for none of them to be blamed.

In this exercise, we will focus on strict equilibria. In our environment, a strict equilibrium is one in which for any realized outcome, each team member has strict preferences over disclosing or not disclosing that outcome. That is, the observer's no-disclosure belief about each individual does not coincide with any possible realization in their (finite) realization set, $\omega_i^{ND} \notin \Omega_i$ for any $i \in N$. Consequently, a strict equilibrium can be fully described by the no-disclosure belief vector ω^{ND}. For each deliberation procedure D and each outcome distribution μ, we then let $\mathcal{E}_\mu^D \subset \text{co}(\Omega)$ be the set of strict equilibria of the disclosure game.

[17] Our chosen measure of i's blame, ω_i^{ND}, is also not normalized by the distribution of i's outcomes, but our results remain unchanged if we use the measure $\omega_i^{ND} / [\max(\Omega_i) - \min(\Omega_i)]$ instead.

Proposition 1 considers the effect of marginal changes in the deliberation procedure D on the equilibrium blame vector ω^{ND} around a particular strict equilibrium. Lemmas 1 and 2 ensure that this is a well-defined exercise. Remember that a deliberation procedure is described by 2^n numbers in $[0,1]$, specifying for each subset $I \subseteq N$ of team members supporting disclosure a probability of disclosure $D(I) \in [0,1]$. Our assumptions require that $D(\varnothing) = 0$, $D(N) = 1$, and $I \subseteq I' \Rightarrow D(I) \leqslant D(I')$. Consequently, the relevant space of deliberation procedures is a compact subset of $[0,1]^{2^n-2}$.

Lemma 1. *For every full-support outcome distribution μ, there is an open set of deliberation procedures D such that a strict equilibrium exists; that is, $\mathcal{E}_\mu^D \neq \varnothing$. Additionally, for every deliberation procedure with $D(\{i\}) < 1$ for every $i \in N$, there is an open set of outcome distributions μ such that $\mathcal{E}_\mu^D \neq \varnothing$.*

Lemma 2. *Fix a deliberation procedure D and a distribution μ such that $\mathcal{E}_\mu^D \neq \varnothing$. Consider a particular strict equilibrium $\varepsilon \in \mathcal{E}_\mu^D$. There exists $\bar{\delta} > 0$ such that in any δ-neighborhood of D, with $\delta < \bar{\delta}$, there is a unique continuous selection E of the strict-equilibrium correspondence $\mathcal{E}_\mu^D$ such that $E(D) = \varepsilon$.*

The proofs of both lemmas are in Appendix A. Lemma 1 shows that strict equilibria often exist in our team-disclosure model, and Lemma 2 ensures that the notion of marginal changes in ω^{ND} due to marginal changes in D around a particular strict equilibrium is well defined. Because D is a multidimensional object, the effect of marginal changes in D on ω^{ND} depends on the direction of these marginal changes. Proposition 1 characterizes directions of changes in the deliberation procedure such that a team member i's blame increases or decreases.

Proposition 1. *Fix a deliberation procedure D and a distribution μ such that $\mathcal{E}_\mu^D \neq \varnothing$, and a strict equilibrium $\varepsilon \in \mathcal{E}_\mu^D$. Consider marginal changes in the deliberation procedure D and their effect on the observer's no-disclosure belief about team member i, ω_i^{ND}. If the marginal change in D satisfies*

$$\min\left\{\frac{dD(I)}{1-D(I)} : i \in I \subseteq N\right\} \geqslant \max\left\{\frac{dD(I)}{1-D(I)} : i \notin I \subseteq N\right\}, \tag{2}$$

then $d\omega_i^{ND}/dD \leqslant 0$. Conversely, $d\omega_i^{ND}/dD \geqslant 0$ if

$$\min\left\{\frac{dD(I)}{1-D(I)} : i \notin I\right\} \geqslant \max\left\{\frac{dD(I)}{1-D(I)} : i \in I\right\}. \tag{3}$$

The proof of Proposition 1 is presented in Appendix A. Intuitively, condition (2) requires that the deliberation procedure move in a direction that increases the probability of disclosure when sets of team members that include team member i recommend disclosure relatively more than when sets of team members that do not include team member i recommend disclosure. In that case, the observer's equilibrium no-disclosure belief about team member i's outcome must decrease. An interpretation is that condition (2) ensures that team member i's voice rights — the right to speak on behalf of the team — increase; and, as a consequence, i's blame for team failures also increases. Two corollaries of Proposition 1 stated below highlight two different sets of directional changes in deliberation that ensure an increase in i's voice rights.

Corollary 1 considers changes in the deliberation procedure that increase the probability of disclosure after every possible set of team members recommends disclosure, but does so in a proportional way. Because the probability of disclosure after either all team members or no team members recommend disclosure is fixed by assumption, the direction of change requires that those remain constant. We say this direction of change makes the deliberation protocol *more unilateral*, because it corresponds to a convex combination between the original deliberation protocol and the unilateral disclosure protocol — see the left panel of Figure 3 for an illustration.

Corollary 1 (to Proposition 1)**.** *We say the deliberation protocol D becomes more unilateral if*

$$\frac{dD(I)}{1 - D(I)} = \frac{dD(I')}{1 - D(I')} \geqslant 0 \text{ for every } I, I' \subseteq N \text{ with } I, I' \notin \{\varnothing, N\}.$$

If D becomes more unilateral, then $d\omega_i^{ND}/dD \leqslant 0$ for all $i \in N$.

When the deliberation procedure becomes more unilateral, the observer understands that all team members have an increased opportunity to enforce the disclosure of a given outcome. Therefore, the observer interprets the team's choice to not disclose an outcome as worse news about each individual team member. In other words, an "aggregate increase" in all team members' voice rights leads to a corresponding "aggregate increase" in their blame for team failures. Corollary 2 shows that a *relative* increase in a particular team member's voice rights — making their disclosure recommendations more pivotal — leads to an increase in their own blame.

Corollary 2 (to Proposition 1)**.** *We say team member i becomes more pivotal if for every $I \subseteq N$*

$$dD(I) > 0 \Rightarrow i \in I \text{ and } dD(I) < 0 \Rightarrow i \notin I.$$

If i becomes more pivotal, then $d\omega_i^{ND}/dD \leqslant 0$.

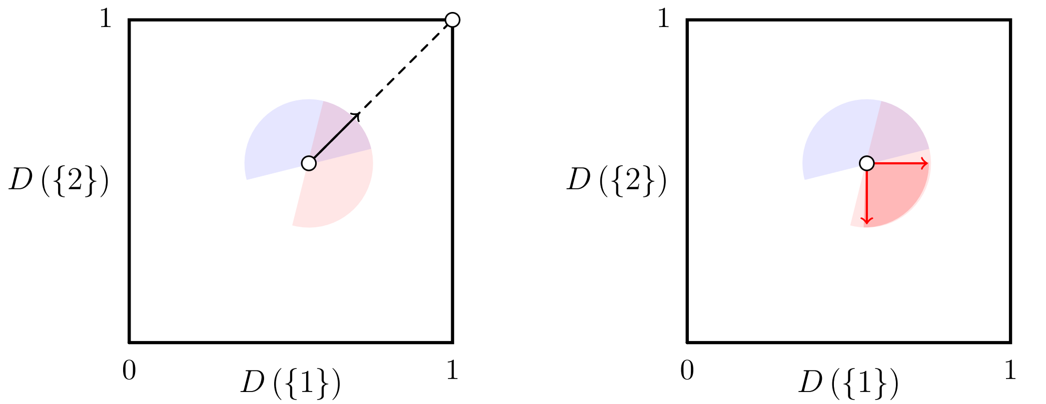


Figure 3: Both panels show the effect of changes in the deliberation procedure in a two-person team on the observer's no-disclosure belief about each team member. Red-shaded areas indicate directions of changes in D that decrease ω_1^{ND} (increase team member 1's blame) and blue-shaded areas indicate directions that increase team member 2's blame. The left panel illustrates in black the direction that makes the protocol more unilateral (see Corollary 1), and the right panel illustrates in darker red the directions that increase team member 1's pivotality (see Corollary 2).

The right panel in Figure 3 shows the directions of change in the deliberation procedure that increase a team member's pivotality in a two-person team. In two-person teams, an increase in team member 1's pivotality implies a decrease in team member 2's pivotality, but this is not necessarily true in larger teams. However, for any team size, Corollary 2 shows that relative increases in i's voice rights — due to either an increase in the probability of disclosure after i recommends it or to decreases in the probability of disclosure when it is recommended only by other team members — ensures an increase in i's blame.

3.2 Is Full-Disclosure Consistent with Deliberation?

Theorem 1 shows that equilibria without full disclosure, when present, coexist with a full-disclosure equilibrium. By definition, in a full-disclosure equilibrium the event of "no disclosure" either does not happen at all on the path of play or happens only after one realization $\omega \in \Omega$ (so that the observer can perfectly infer the outcome realization even after seeing no disclosure). If no disclosure happens only off-path, then the equilibrium notion in Definition 1 does not impose any requirements on the observer's vector of no-disclosure beliefs ω^{ND}. Therefore, a vector of beliefs that sustains full disclosure as the team's strategy is vacuously consistent with equilibrium.

In this section, we inspect the plausibility of the off-path no-disclosure beliefs that support full disclosure as part of the equilibrium set. Specifically, we wish to evaluate whether these beliefs are consistent with the aggregation of individual recommendations via the team's deliberation procedure. Our notion of consistency — in the spirit of structural consistency from Kreps and Wilson (1982) — requires that the observer's vector of no-disclosure beliefs be reached via Bayesian updating for some conjectured (not necessarily optimal) team disclosure strategy that aggregates individual recommendations via the given deliberation procedure and reaches the "no-disclosure" information set with positive probability.

Definition 2. *No-disclosure beliefs* $\omega^{ND} = (\omega_1^{ND}, ..., \omega_N^{ND})$ *are* ***consistent with deliberation*** *procedure* D *if there exists some team disclosure strategy* d *with* $d(\omega) < 1$ *for some* $\omega \in \Omega$*, and a vector of individual recommendations* x *such that*

1. *Individual disclosure recommendations are determined by own outcome values:*

$$\text{For each } i \in N \text{ and } \omega, \hat{\omega} \in \Omega,\ \omega_i = \hat{\omega}_i \Rightarrow x_i(\omega) = x_i(\hat{\omega}).$$

2. *The team's disclosure decision aggregates the individual disclosure strategies* x*:*

$$d(\omega) = \sum_{X \subseteq N} \Pi_X(\omega) D(X) \text{ for every } \omega \in \Omega.$$

3. *No-disclosure posteriors are Bayes-consistent: for each* $i \in N$*,* ω_i^{ND} *satisfies (1).*

Theorem 2 states a necessary and sufficient condition on a team's deliberation procedure for the existence of a full-disclosure equilibrium that is supported by consistent beliefs. This condition requires the deliberation protocol to be such that reaching the decision to disclose an outcome does not require more consensus among team members than reaching the decision to conceal an outcome with positive probability. Formally, we say that *disclosure requires more consensus than concealing* if for every subgroup $I \subseteq N$ such that $D(I) = 1$ and $D(N \setminus I) < 1$, there exists a smaller subgroup $J \subset I$ such that $D(N \setminus J) < 1$ and $D(J) < 1$. In this requirement, I is a group of team members who can jointly implement disclosure, regardless of other team members' recommendations (because $D(I) = 1$); and group I can also jointly block disclosure (with some probability), as $D(N \setminus I) < 1$. If the requirement is satisfied, then for every such group I, there exists a smaller subgroup of team members who can jointly block disclosure, but cannot jointly implement disclosure. We then conclude that disclosure requires the consensus of the larger group I, but that non-disclosure requires only the consensus of the smaller subgroup J. For example, in a team with $n = 2$, disclosure requires more consensus than concealing if and only if the decision to disclose an outcome with probability 1 must be reached via consensus between the two team members. or equivalently, if and only if neither team member can unilaterally choose disclosure; that is, $D(\{1\}) < 1$ and $D(\{2\}) < 1$.

Theorem 2. *A full-disclosure equilibrium that is consistent with the deliberation procedure D exists if and only if D is such that disclosure does not require more consensus than concealing.*

To understand the result, suppose that there are only two team members and that disclosure requires more consensus than concealing, so that $D(\{1\}) < 1$ and $D(\{2\}) < 1$. Conjecture a full-disclosure equilibrium supported by a vector of no-disclosure beliefs ω^{ND}. It must be that $\omega_1^{ND} = \min(\Omega_1)$ and $\omega_2^{ND} = \min(\Omega_2)$, for otherwise one of the team members would strictly prefer to not disclose realizations where they draw their worst possible outcome. Because of the assumption on the deliberation procedure, they would be able to unilaterally impose such non-disclosure with positive probability, thereby contradicting the initial assumption that the equilibrium has full disclosure.

Now we wish to craft a pair of individual disclosure strategies $\hat{x}$ to be used to "justify" the beliefs $\omega^{ND} = (\min(\Omega_1), \min(\Omega_2))$. These strategies must imply that some realization $\hat{\omega}$ is not disclosed with positive probability, and therefore it must be that either $\hat{x}_1(\hat{\omega}_1, \omega_2) < 1$ for all $\omega_2 \in \Omega_2$ or $\hat{x}_2(\omega_1, \hat{\omega}_2) < 1$ for all $\omega_1 \in \Omega_1$. If the former is true, then all realizations $\omega_2 \in \Omega_2$ are concealed with positive probability, which implies that the no-disclosure posterior ω_2^{ND} consistent with $\hat{x}$ is strictly larger than $\min(\Omega_2)$. If the latter is true, then $\omega_1^{ND} > \min(\Omega_1)$. Combining these two cases, we conclude that the off-path beliefs necessary to sustain full disclosure cannot be justified by *any* disclosure strategy that is consistent with the deliberation procedure. With some work shown in the Appendix, this argument generalizes to teams with more than two members, so long as the deliberation procedure is such that disclosing requires more consensus than concealing.[18]

For the other direction of Theorem 2, consider again a team with two individuals, and now suppose that disclosure does not require more consensus than concealing; without loss, let $D(\{1\}) = 1$. Given this deliberation procedure, there exists a full-disclosure equilibrium where $\omega_1^{ND} = \min(\Omega_1)$ and $\omega_2^{ND} > \min(\Omega_2)$. Furthermore, these no-disclosure beliefs can be justified by the following individual disclosure strategies: $\hat{x}_1(\omega) = 0$ if $\omega_1 = \min(\Omega_1)$ and $\hat{x}_1(\omega) = 1$ otherwise; and $\hat{x}_2(\omega) = 0$ for all $\omega \in \Omega$. Again, in the Appendix, we show that this construction can be generalized to teams with $n > 2$ team members if, under the team's deliberation procedure, disclosure does not require more consensus than concealing.

Together, Theorems 1 and 2 illustrate that full disclosure is harder to support in an equilibrium the higher the degree of consensus required for a team to choose disclosure. The first result

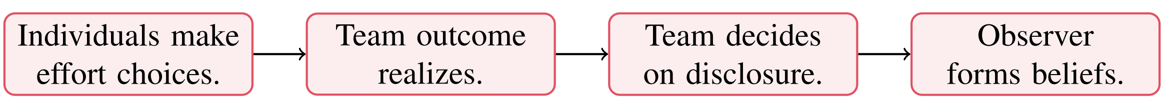
[18]More precisely, in any full-disclosure equilibrium there must be a subgroup $I \subseteq N$ of the team, who can together choose disclosure (that is, $D(I) = 1$) and such that $\omega_i^{ND} = \min(\Omega_i)$ for all $i \in I$. We show that it is impossible to construct a strategy profile $\hat{x}$ that justifies these off-path beliefs. To argue this point, we use the fact that there is a subset of team members $J \subset I$ that can together ensure that an outcome is not disclosed with some probability, that is, $D(J) < 1$.

shows that unless disclosure is very easy — in the sense that it can be chosen unilaterally by all team members — full disclosure is not the unique equilibrium outcome. Theorem 2 strengthens the observation by showing that if disclosure requires more consensus than concealing, then not only do equilibria without full disclosure exist, but they are the only equilibria that survive our proposed refinement. For further illustration of the conditions in Theorems 1 and 2, consider anonymous deliberation procedures, in which the team's disclosure decision depends only on the *number* of team members who recommend disclosure. Within that class, disclosure can be chosen unilaterally by all team members if $D(I) = 1$ if and only if $|I| \geqslant 1$; and disclosure requires more consensus than concealing if $D(I) = 1$ if and only if $|I| \geqslant n/2$.

Corollary 3 (to Theorems 1 and 2)**.** *Suppose D is an anonymous deliberation procedure, with $D(I) = 1$ if and only if $|I| \geqslant k$. Full disclosure is the unique equilibrium outcome if and only if $k = 1$; and full disclosure is consistent with deliberation if and only if $k \leqslant n/2$.*

4 Deliberation and Incentives

So far in this paper, we have studied a problem whereby a team chooses how to communicate about their group outcomes with outside observers, with a fixed outcome distribution μ. We now explore the effects of team disclosure on individual incentives to contribute to team production in the first place. To do so, we add to the team's problem a production stage that takes place before the disclosure stage.



Formally, at the initial stage, each agent unilaterally and covertly makes an effort decision: team member $i \in N$ chooses $e_i \in \{0, 1\}$, bearing the individual cost $c_i > 0$ if they choose to put effort into the team project ($e_i = 1$) and no cost otherwise. Individual effort choices are collected into the team's effort vector $e = (e_1, ..., e_n)$. Once individuals make effort decisions, a team outcome is drawn from the distribution $\mu(\cdot; e)$, which now depends on the effort vector e chosen by the team. After the team outcome realizes, the disclosure stage ensues as before: all team members see the realized outcome ω and make disclosure recommendations. Disclosure is decided by the aggregation of individual recommendations through the deliberation procedure D. The observer sees the disclosed/not-disclosed outcome, according to the team's decision, but never sees the team members' effort choices — that is the sense in which effort decisions are covert.

Assumption 3 imposes that the support of outcomes is invariant to the chosen vector of efforts and that the outcome distribution increases in the team's effort.

Assumption 3. *For each* $e \in \{0,1\}^n$*,* $\mu(\cdot; e)$ *has full support over* $\Omega = \Omega_1 \times \ldots \times \Omega_n$*, where* $\Omega_i \subset \mathbb{R}$ *has at least* 2 *elements for all* $i \in N$*. Moreover, effort is productive, so that*[19]

$$e \geqslant e' \Rightarrow \mu(\cdot; e) \succsim_{FOS} \mu(\cdot; e').$$

In our extended game, an equilibrium is defined by an equilibrium of the team-disclosure game (as in Definition 1) and individual rationality at the effort-choice stage, given the team-disclosure equilibrium.

Throughout the game, the deliberation procedure is fixed at D, and it is commonly understood by all team members and by the observer. A natural interpretation of the fixed disclosure protocol D is that, even prior to the effort stage, team members collectively pick a deliberation procedure — that is, they agree on a governance structure to use to aggregate individual preferences in future communication decisions. With such an interpretation in mind, our exercise in this section describes the types of governance structures a team should implement if the objective is to incentivize individuals to contribute effort to team production.

This "deliberation design" problem assumes that the team can commit to a protocol to aggregate individual disclosure recommendations into a team disclosure decision, but that it cannot commit to a rule that directly specifies a disclosure decision for each possible team outcome. Our understanding is that, in many applied contexts, contracting on individual recommendations is easier than contracting directly on team production outcomes, as these may not be immediately observable or measurable. For example, in environments in which individuals have career concerns, the value of an outcome to each team member is given by the observer's perception of their respective type, which is not an easily measurable object.

4.1 Team Disclosure and Effort Incentives

There are many possible criteria with which to evaluate the effort incentives provided by different deliberation processes. We evaluate effort incentives provided by different deliberation procedures by comparing their implied sets of cost vectors for which a full-effort equilibrium

[19] The notation $\succsim_{FOS}$ indicates (multivariate) first order stochastic dominance. We say that a random vector X dominates a random vector Y in the first order stochastic if $\mathbb{P}(X \in U) \geqslant \mathbb{P}(Y \in U)$ for every upper set $U \in \mathbb{R}^n$. Equivalently, random vector X dominates random vector Y in the first order stochastic if $\mathbb{E}[\varphi(X)] \geqslant \mathbb{E}[\varphi(Y)]$ for all increasing functions φ for which the expectations exist. See Shaked and Shanthikumar (2007).

exists, in which all team members exert costly effort.[20] Although we use this specific criterion, the mechanisms highlighted in the current analysis are more general, and our results could be adapted to other criteria, such as the equilibrium implementation of efficient effort and comparisons in terms of the overall set of efforts implementable in equilibrium by each procedure.

Lemma 3 establishes the basis for this analysis, characterizing the relation between disclosure strategies used at the team-disclosure stage and team members' incentives to exert costly effort. Let $c \in \mathbb{R}^n_{++}$ be the vector of effort costs for the team. For a team-disclosure strategy $d : \Omega \to [0,1]$, we say that d implements full effort if, given that the team uses strategy d in the disclosure subgame after any individual effort choices, it is optimal for each individual $i \in N$ to choose $e_i = 1$ at the effort stage. For any subgroup $I \subset N$, we use notation e_I to indicate an effort vector such that individuals $i \in I$ exert effort and individuals $i \in N \setminus I$ do not.

Lemma 3. *A team-disclosure strategy $d : \Omega \to [0,1]$ implements full effort for a given cost vector $c \in \mathbb{R}^N$ if and only if, for every $i \in N$,*[21]

$$\underbrace{\mathbb{E}\left(\omega_i|e_N\right) - \mathbb{E}\left(\omega_i|e_{N\setminus i}\right)}_{\textit{Individual Effort Benefits}} + \mathbb{P}\left(ND|e_{N\setminus i}\right) \underbrace{\left[\mathbb{E}\left(\omega_i|ND;e_{N\setminus i}\right) - \mathbb{E}\left(\omega_i|ND;e_N\right)\right]}_{\textit{Misattributed Blame}} \geqslant c_i. \quad (4)$$

The expression in (4) clarifies how the selective disclosure of the team's outcomes can be used to incentivize team members to put in effort beyond their baseline "full-disclosure" incentives. On the left-hand side of (4), the first highlighted term corresponds to team member i's direct individual benefits from exerting effort. The term compares individual i's expected outcome value when they choose $e_i = 1$ versus $e_i = 0$ while maintaining the assumption that all other team members exert effort. Under full disclosure, that is exactly the benefit of exerting effort for individual i.

The second term corresponds to the extra incentives to exert effort that are provided by selective non-disclosure. Specifically, the second term accounts for how the observer misattributes blame to team member i when they do not exert effort and a team failure ensues. More

[20] Our results in Section 2 show that there are often multiple equilibria in the team-disclosure game. Accordingly, there are often multiple equilibrium effort vectors that are implementable in the larger game under the same deliberation procedure. Our criterion therefore requires that full effort be implementable in some, but not necessarily all, such equilibria.

[21] The following rewriting of (4) expresses the relation with the team disclosure strategy d more directly:

$$c_i \leqslant \sum_{\Omega} \omega_i \mu(\omega; e_N) - \sum_{\Omega} \omega_i \mu(\omega; e_{N\setminus i})$$

$$- \sum_{\Omega} (1 - d(\omega))\mu(\omega; e_{N\setminus i}) \left[\frac{\sum_{\Omega} \omega_i (1 - d(\omega))\mu(\omega; e_N)}{\sum_{\Omega} (1 - d(\omega))\mu(\omega; e_N)} - \frac{\sum_{\Omega} \omega_i (1 - d(\omega))\mu(\omega; e_{N\setminus i})}{\sum_{\Omega} (1 - d(\omega))\mu(\omega; e_{N\setminus i})} \right].$$

formally, in a full-effort equilibrium, the observer expects all team members to contribute effort to the team project, so that $e = e_N$. But suppose person i deviates and shirks, so the true effort vector is $e = e_{N\setminus i}$. And given this deviation, suppose that the drawn outcome is such that the team chooses not to disclose it. Because i's effort is covert, the observer still calculates i's "blame" under the presumption that all team members exerted effort — and therefore i's value is $\mathbb{E}(\omega_i|ND;e_N)$, as opposed to the "correct" blame assessment $\mathbb{E}\left(\omega_i|ND;e_{N\setminus i}\right)$.[22]

The misattributed blame term is positive — and therefore selective non-disclosure via d provides stronger effort incentives than full-disclosure — if, for each team member i, the observer's equilibrium blame attribution to them is harsher than the correct assessment given a deviation by i to shirking. Intuitively, this is the case when the team's disclosure decisions are more correlated with i's outcome under full effort than under the alternative effort vector $e_{N\setminus i}$.[23]

The decomposition of individual effort incentives in Lemma 3 parallels incentive decompositions for different governance structures in hold-up models. For example, Grossman and Hart (1986) show that individual incentives to invest in a relationship are determined first by the direct benefits of that investment and second by the effect of that investment on actions later chosen by whoever has control over the relationship asset. Moreover, the efficient allocation of control over the relationship asset is the one that best aligns this second term with efficient effort incentives.[24] Similarly in our model, effort incentives are determined partly by the direct effect of effort on individual outcomes and partly by the effect of effort on the team's disclosure decisions in the communication stage (and the implied blame assignment following those decisions). And effort-maximizing allocations of voice-rights — effort-maximizing deliberation protocols — are those that maximize this latter term.

4.2 Comparing Deliberation Procedures

To evaluate the effort incentives provided by a particular deliberation procedure D, we assess whether full effort can be implemented by team-disclosure strategies induced in equilibrium

[22] To calculate these conditional expectations, we maintain the team-disclosure strategy d unchanged as a function of the realized outcome ω, and vary the outcome distribution with i's effort choice. In any equilibrium, disclosure strategies in the disclosure stage must not depend on the effort choice in the initial stage, because each agent chooses effort covertly.

[23] Indeed, the following rewriting of the left-hand side of (4) expresses i's effort gains directly in terms of the improvement of the covariance between i's outcome and disclosure:

$$\left[1-\mathbb{P}\left(ND|e_{N\setminus i}\right)\right]\left[\mathbb{E}\left(\omega_i|e_N\right)-\mathbb{E}\left(\omega_i|e_{N\setminus i}\right)\right]+\frac{\mathbb{P}\left(ND|e_{N\setminus i}\right)}{\mathbb{P}\left(ND|e_N\right)}Cov\left(\omega_i,d|e_N\right)-Cov\left(\omega_i,d|e_{N\setminus i}\right). \tag{5}$$

Please see the Appendix, where we derive this expression from (4).

[24] See also Gibbons (2005), section 2.2, for a more detailed description.

when the team makes disclosure decisions using D. For a given procedure D, we let $FE(D) \subset \mathbb{R}^n_{++}$ be its corresponding full-effort set.

Definition 3 (Full-Effort Set)**.** *A cost vector $c \in \mathbb{R}^n_{++}$ is in the full effort set of the deliberation procedure D — $c \in FE(D)$ — if*

1. *There is a team disclosure strategy d that implements full effort for cost vector c;*
2. *Given deliberation procedure D and outcome distribution $\mu(\cdot; e_N)$, team disclosure strategy d is an equilibrium of the team disclosure game.*

This definition imposes only that a team disclosure strategy d that implements full effort be an equilibrium of the disclosure subgame after all team members exert effort — that is, under the outcome distribution $\mu(\cdot; e_N)$. But note that, if that is the case, then there exists an equilibrium of the full extended game in which all team members exert effort at the production stage and, at the disclosure state, team strategy d is played after all possible individual effort choices. (We state this in Observation 2 below.) This observation holds because effort is covert, and therefore the observer's no-disclosure beliefs are calculated based on *equilibrium* individual effort choices and do not change after an unobserved individual effort deviation. As such, each individual is willing to use the same disclosure recommendation strategy after a deviation as they do on the equilibrium path.

Observation 2. *If the cost vector is c and the deliberation procedure is D, with $c \in FE(D)$, then there exists an equilibrium of the production plus disclosure game in which every team member exerts effort.*

We use the full-effort sets of deliberation procedures to compare them in terms of the provision of effort incentives. We say that a deliberation procedure D *dominates* a deliberation procedure D' if $FE(D') \subseteq FE(D)$. Or equivalently, D dominates D' if for every cost vector c such that full effort can be implemented in equilibrium under procedure D', full effort can also be implemented in equilibrium under procedure D.

4.3 Effort Environments and Effective Deliberation

We now use the characterization in Lemma 3 to study which types of deliberation procedures effectively incentivize effort. Note that, while Assumption 3 guarantees that effort is productive, a key observation is that there are different ways in which an individual's effort can improve the outcome distribution. Our analysis contrasts an effort environment in which a team member's effort improves the distribution of their own outcome values with a "high externality" environment, in which an individual's effort improves the distribution of outcome values for other team members. In these distinct environments, different types of deliberation procedures should be used to effectively incentivize effort.

We say effort is *purely self-improving* if, for every $i \in N$ and every $I \subseteq N$,[25]

$$\mu_{N\setminus i}(\cdot; e_I) = \mu_{N\setminus i}(\cdot; e_{I\setminus i}), \text{ and } \mu_i(\cdot|\omega_{N\setminus i}; e_I) \succ_{FOS} \mu_i(\cdot|\omega_{N\setminus i}; e_{I\setminus i}) \text{ for all } \omega_{N\setminus i} \in \Omega_{N\setminus i}$$

The notation $\mu_{N\setminus i}(\cdot; e_I)$ indicates the joint distribution of outcomes of all team members except for team member i, when the team's effort is e_I. In turn, $\mu_i(\cdot|\omega_{N\setminus i}; e_I)$ indicates the outcome distribution for team member i conditional on outcome realization $\omega_{N\setminus i}$ for all other team members, given the team effort e_I. Accordingly, we say that effort is self-improving if, for every team member $i \in N$, their own effort leaves the outcome distribution of other team members unchanged, but improves their own outcome distribution, conditional on others' outcome realization. A purely self-improving environment describes a low-team-externality situation, in which all the benefits from exerting effort accrue directly to the individual who puts in that effort, and not to their fellow teammates.

In contrast, we capture a high-team-externality environment — in which an individual's effort benefits accrue not to themselves directly, but to the other team members — as an environment in which effort is *purely team-improving*. Formally, effort is purely self-improving if for every $i \in N$ and every $I \subset N$,

$$\mu_{N\setminus i}(\cdot|\omega_i; e_I) \succ_{FOS} \mu_{N\setminus i}(\cdot|\omega_i; e_{I\setminus i}) \text{ for all } \omega_i \in \Omega_i, \text{ and } \mu_i(\cdot; e_I) = \mu_i(\cdot; e_{I\setminus i}).$$

Both these definitions describe effort environments in which the team-outcome distribution may involve outcome correlation across team members. If we consider the special case in which outcomes are independent across team members, then self-improving effort corresponds to a situation in which i's outcome distribution increases in the first order stochastic if i exerts effort, and j's outcome distribution remains unchanged for all $j \neq i$. Analogously, team-improving effort is such that j's outcome distribution increases with i's effort, for $j \neq i$, but i's distribution is unchanged. Our distinction between high- and low-team-externality environments parallels the distinction between selfish and cooperative investments in a hold-up context proposed by Che and Hausch (1999).

[25]The notation $\succ_{FOS}$ indicates *strict* (multivariate) first-order stochastic dominance. We say that a random vector X strictly dominates a random vector Y, both defined over Ω, in the first order stochastic if $\mathbb{P}(X \in U) \geqslant \mathbb{P}(Y \in U)$ for every upper set $U \in \Omega$, and strictly so for some such U.

Proposition 2 ranks deliberation procedures in both of these effort environments. The proposition uses the following previously introduced terms: the *unilateral disclosure* protocol is the procedure in which every team member can unilaterally choose disclosure; and the *consensus disclosure* protocol is such that every team member can unilaterally veto disclosure.[26]

Proposition 2. *1. If effort is purely self-improving, then the unilateral disclosure protocol dominates any other deliberation procedure.*

2. If effort is purely team-improving, then consensus disclosure strictly dominates every deliberation procedure in which some team member can unilaterally choose disclosure.

A full proof of Proposition 2 is in the Appendix. When effort is self-improving, we show that all of i's gains from effort are captured when the team fully discloses their outcomes, which is the (unique) equilibrium team disclosure attained when the deliberation process allows any team member to unilaterally choose disclosure. Intuitively, because an agent's effort affects only their own outcome, non-disclosure can only harm effort incentives by concealing some of the effort gains from the observer. As a consequence, the unilateral disclosure protocol — by inducing an equilibrium with full disclosure — maximizes the team's "full effort cost set." Note that, as per Theorem 1, full disclosure is an equilibrium for any deliberation procedure; therefore, the maximal full-effort cost set can be attained regardless of the deliberation process. In that sense, the unilateral disclosure procedure is sufficient for maximizing the full-effort set. By Theorem 2, we can refine the equilibrium set induced by different deliberation protocols — specifically, we can refine out the full-disclosure equilibrium when "disclosing requires more consensus than concealing." If we accordingly define our dominance criterion accounting for this refinement, we can then establish that the unilateral disclosure protocol strictly dominates deliberation protocols in which disclosing requires more consensus than concealing.

Suppose instead that effort is purely team improving. The proposition argues that in such a high-team-externality effort environment, the equilibrium team-disclosure strategy implemented by consensus disclosure produces larger effort incentives than equilibrium disclosure induced under more unilateral deliberation protocols. Consider the consensus disclosure procedure, and remember that an equilibrium exists in which each team member favors disclosure if and only if their own-outcome draw is good enough. When team member i puts in effort, they improve the odds that all other team members will draw an outcome for which they favor disclosure; thereby improving the odds that i's disclosure recommendation is pivotal to the

[26] Proposition 2 states results for purely self-improving and purely team-improving environments, but the respective statements also hold (by continuity) for environments close enough to either of these extreme cases. Likewise, Proposition 2 holds for comparisons between almost-unilateral and almost-consensus deliberation procedures.

overall team decision. And consequently team member i is "more to blame" for team failures under full effort than in the deviation where i does not exert effort. In other words, under the consensus disclosure deliberation protocol, each team member has incentives to improve the outcomes of their partners, so as to avoid situations in which others veto the disclosure of their own good outcome realizations. Through this mechanism, each individual partly internalizes the externality that their effort imposes on fellow team members.

Proposition 3 below extends Proposition 2 to environments in which effort is not fully self-improving or fully team-improving. We show that as the effort environment becomes "more self-improving," the unilateral disclosure procedure becomes more dominant relative to other procedures. Analogously, as the environment becomes "more team-improving," the extra effort incentives provided by the consensus procedure increase, relative to the unilateral procedure.

Proposition 3. *Let* $\{\hat{\mu}(\cdot; e)\}$ *and* $\{\mu(\cdot; e)\}$ *be two effort environments satisfying* $\hat{\mu}(\cdot; e_N) = \mu(\cdot; e_N)$*, and suppose that* $\{\hat{\mu}(\cdot; e)\}$ *is such that effort is purely self-improving. For each* $\alpha \in [0, 1]$*, let* $\{\mu^\alpha(\cdot; e)\}$ *be an effort environment such that for each* $e \in \{0, 1\}^n$*,*

$$\mu^\alpha(\cdot; e) = \alpha\hat{\mu}(\cdot; e) + (1 - \alpha)\mu(\cdot; e),$$

so that α *measures the degree to which effort is self-improving. There exists* $\bar{\alpha} \in [0, 1]$ *such that the unilateral disclosure protocol dominates any other deliberation procedure in effort environment* $\{\mu^\alpha(\cdot; e)\}$ *if and only if* $\alpha \geqslant \bar{\alpha}$*.*

Suppose instead that $\{\hat{\mu}(\cdot; e)\}$ *is such that effort is purely team-improving, so that* α *measures the degree to which effort is team-improving. Then there exists* $\bar{\alpha} \in [0, 1]$ *such that consensus disclosure strictly dominates the unilateral disclosure protocol in effort environment* $\{\mu^\alpha(\cdot; e)\}$ *if and only if* $\alpha > \bar{\alpha}$*.*

Beyond the classification of effort environments in terms of the externalities that team members impose on their partners, productive settings also differ in terms of how team members' efforts contribute to the correlation between individual outcomes. Proposition 4 below considers effort environments in which effort increases the correlation between team members' outcomes. We can interpret these as investments in a common component of team production. The proposition states that if each team member's effort sufficiently improves the correlation between all team members outcomes, then the unilateral disclosure procedure is dominated by all other procedures. To show that, we momentarily assume that the support of outcomes does not differ across agents, so that $\Omega = \Omega_i^n$ for some $\Omega_i \subset \mathbb{R}$; and we say that a distribution ν over Ω has perfect correlation across team members' outcomes if it has full support on the locus

$\omega_1 = \ldots = \omega_n$.[27]

Proposition 4. *Suppose that μ and ν are two distributions over $\Omega = \Omega_i^n$, where μ has full support and ν has perfect correlation across team members' outcomes; and suppose that $\nu \succsim_{FOS} \mu \succsim \mu(\cdot; e_{N\setminus i})$ for every $i \in N$. Consider varying the correlation in $\mu(\cdot; e_N)$ by letting, for $\epsilon \in (0, 1)$,*

$$\mu(\cdot; e_N) = (1 - \epsilon)\mu + \epsilon\nu.$$

Let D be the unilateral disclosure protocol and D' be a deliberation protocol in which no team member can unilaterally choose disclosure. There exists some $\bar{\epsilon} \in (0, 1)$ such that if $\epsilon > \bar{\epsilon}$, D' strictly dominates D.

4.4 Deliberation Procedures and Corporate Culture

One way to interpret deliberation procedures in real-world team production environments is as the "corporate culture" of a team. O'Reilly and Chatman (1996) define corporate culture as "a set of norms and values that are widely shared and strongly held throughout the organization." In our environment, the norm that is being upheld in the team is the one guiding how the team aggregates individual disclosure recommendations into the team's disclosure decision.

The unilateral disclosure procedure induces an equilibrium in which teams, after every possible team outcome realization, reveal to the observer which exact realization occurred. Therefore, both after team successes and after team failures, each team member's share of blame for that outcome is clarified. These equilibria parallel the idea of "radically transparent" corporate cultures, in which individuals are held fully accountable for their contributions to their team's successes and failures. Business sources often praise the effort incentives provided by radically transparent cultures. In that context, the article "How to Win the Blame Game," in the Harvard Business Review posits that "when used judiciously (...) blame can prod people to put forth their best efforts."

In contrast, the consensus disclosure protocol induces equilibria in which the team suffers the burden of team failures collectively. When an outcome realization happens that is deemed a team failure, the team decides not to disclose it to the outside observer, who then spreads the blame for this outcome across all team members. This dynamic resembles corporate cultures in which teams are committed to "not play the blame game." The Harvard Business Review article "When Transparency Backfires, and How to Prevent It," acknowledges the benefits of such cul-

[27] These assumptions are made for notational convenience. Proposition 4 holds if the support of outcomes differs across agents, and under the weaker assumption that ν is supported on the locus $\omega_j = \varphi_{ij}(\omega_i)$ for some strictly increasing function φ_{ij} for all $i, j \in N$.

tures, in comparison with more transparent teams: "too much transparency can create a blaming culture that may actually decrease constructive, reciprocal behavior between employees."

From the perspective of our Proposition 2, both types of corporate cultures may be effective in incentivizing individuals to contribute effort to their teams, each proving to be best suited to a particular type of effort environment. This result points to a possible empirical exercise that attempts to assess whether indeed "accountability," "transparency," and "blame" cultures are less present in effort environments with higher team externalities. We view this exercise as beyond the scope of the current theoretical paper, but note the difficulties in measuring both the relevant component of corporate culture and the degree to which a team's activity has high or low team externalities.[28]

5 Disclosure and Incentives in Binary Environments

We have ranked different deliberation procedures in terms of their effort-incentives provision. We now characterize team disclosure and its relation to effort incentives in an environment in which individual outcome values are binary. That is, for each $i \in N$, $\Omega_i = \{\omega_i^\ell, \omega_{,i}^h\}$, with $\omega_i^\ell < \omega_i^h$. use the further tractability implied by binary environments, the results in this section provide characterizations of deliberation procedures that *maximize* effort incentives. Our results confirm the importance of the forces highlighted in Section 4. Proposition 6 shows that the effort-maximizing deliberation procedure requires more consensus for disclosure when effort is "more team-improving," and when effort more strongly correlates the team members' outcome values. The same pattern is confirmed in a numerical exercise in Section 5.3.

5.1 Symmetric Binary Environments

In a fully general binary environment, an effort-maximizing deliberation procedure may not exist, because the ordering we impose on the space of deliberation procedures is not complete. For example, when compared to procedure 2, procedure 1 may improve effort incentives for some team members but decrease them for other individuals. In that case, neither procedure dominates the other. To ensure the existence of effort-maximizing procedures, we consider only symmetric environments:

[28] The empirical literature that aims to measure corporate culture — for example Guiso, Sapienza and Zingales (2015) — highlight and measure five "core corporate values." They are innovation, integrity, quality, respect, and teamwork. Li, Mai, Shen and Yan (2021) develop a culture dictionary for each of these core values; from their documentation, we can see that accountability and transparency fall under the umbrella of integrity. However, other components of integrity do not seem to correlate with the mechanism highlighted in our paper.

Assumption 4 (Symmetry)**.** *A binary-outcomes environment is symmetric if*

(i) The deliberation procedure is symmetric: for all $X, X' \subseteq N$*,*

$$|X| = |X'| \Rightarrow D(X) = D(X').$$

(ii) Agents' outcomes share the same binary support: $\Omega_i = \{\omega^\ell, \omega^h\}$ *for every* $i \in N$*.*

(iii) If all agents exert effort, the outcome distribution is symmetric: for every $\omega \in \Omega$*,*

$$|\{i \in N : \omega_i = \omega^h\}| = |\{i \in N : \omega'_i = \omega^h\}| \Rightarrow \mu(\omega; e_N) = \mu(\omega'; e_N).$$

(iv) Agents' efforts affect the outcome distribution symmetrically: for every $i, j \in N$ *and* $\omega, \omega' \in \Omega$ *with* $\omega_i = \omega'_j$*,* $\omega_j = \omega'_i$*, and* $\omega_k = \omega'_k$ *for all* $k \neq i, j$*,*

$$\mu(\omega; e_{N\setminus i}) = \mu(\omega'; e_{N\setminus j}).$$

The symmetry assumption requires the deliberation procedure to be symmetric, as well as the outcome distribution and the effort environment. Conditions (iii) and (iv) are sufficient symmetry requirements for our evaluation of effort incentives in terms of the implementation of full-effort equilibria. If all team members exert effort, then condition (iii) requires the probability of an outcome ω to depend only on the number of agents to whom ω is a high outcome. In this way, the probability of that outcome does not depend on the identity of the agents to whom ω is a high or a low outcome. Condition (iv), in turn, imposes that the impact of an agent's effort decision on their own outcome — as well as the impact of their effort on other team members' outcomes — is the same across all team members. Note that we do not impose that the outcome distribution when all but one team member exerts effort be itself symmetric, because we allow an agent's effort to affect their own and other team members' outcomes differently.

Proposition 5 characterizes the equilibrium set under our symmetry assumption: there exists at most one equilibrium without full disclosure. The key observation in the proof of Proposition 5 is that, in a symmetric environment, in an equilibrium in which the observer is maximally skeptical about some team member $i \in N$, the observer must also be maximally skeptical about every other team member $j \in N$. Consequently, there are only two possible "equilibrium types:" one with full disclosure, in which $\omega_i^{ND} = \omega^\ell$ for every $i \in N$, and one without full disclosure, in which $\omega_i^{ND} = \omega_j^{ND} > \omega^\ell$ for every $i, j \in N$.

Proposition 5. *Suppose that Assumption 4 holds. In each team disclosure subgame:*

1. *A full-disclosure equilibrium exists in which, for each $i \in N$, x_i depends only on ω_i.*

2. *If $D(\{i\}) = 1$, every team disclosure equilibrium involves full disclosure.*

3. *If $D(\{i\}) \neq 1$, there exists a unique equilibrium without full disclosure in which, for each $i \in N$, x_i depends only on ω_i. In it, each team member recommends disclosure if and only if they draw a high outcome:*

$$x_i(\omega) = \begin{cases} 1, \textit{ if } \omega_i = \omega_{h,i}, \\ 0, \textit{ if } \omega_i = \omega_{\ell,i}. \end{cases}$$

Lemma 4 uses the characterization in Proposition 5 to show that, under the symmetry assumption (Assumption 4), the notion of an effort-maximizing symmetric deliberation procedure is well-defined.

Lemma 4. *Suppose that Assumption 4 holds. There exists a symmetric deliberation procedure D that maximizes effort incentives among symmetric deliberation procedures.*

A full proof of Lemma 4 is in the Appendix. First, Proposition 5 implies that, to evaluate the effort incentives provided by a symmetric procedure D, it suffices to consider the incentives provided by the unique equilibrium without full disclosure of the disclosure subgame. Because both D and the full-effort distribution $\mu(\cdot; e_N)$ are symmetric, we know then that the set $FE(D)$is equal to $(0, \bar{c}(D)]^n$ for some $\bar{c}(D) \in \mathbb{R}_+$. And so D maximizes effort incentives if $\bar{c}(D) \geqslant \bar{c}(D')$ for all symmetric deliberation procedures D'.

By Lemma 3 in Section 4, we then know that $\bar{c}(D)$ is determined by the extra effort-incentives provided by "blame misattribution" in the equilibrium without full disclosure under deliberation procedure D, if one exists, and $\bar{c}(D) = 0$ if D induces full disclosure as the only equilibria of the disclosure game. The maximum cost $\bar{c}(D)$ is thus the unique value of c that satisfies equation (4) with equality. Lemma 4 then follows from the fact that $\bar{c}(D)$ is a continuous function of D, and that the space of symmetric deliberation procedures is compact.

5.2 Effort-Maximizing Deliberation with Two Team Members

If a team is made up of two individuals, there are four possible team outcomes: $(\omega_\ell, \omega_\ell)$, (ω_ℓ, ω_h), (ω_h, ω_ℓ), and (ω_h, ω_h). For a given team member i, consider two distributions over these four outcomes: the distribution $\mu(\cdot; e_N)$ implied if both team members exert effort, and the distribution $\mu(\cdot; e_{N\setminus i})$ induced if only team member $-i$ exerts effort. The following features of each of these distributions are important for our analysis:

$$\rho = \frac{\mu\left[(\omega_\ell, \omega_\ell); e_{N\setminus i}\right]}{\mu\left[(\omega_h, \omega_\ell); e_{N\setminus i}\right] + \mu\left[(\omega_\ell, \omega_h); e_{N\setminus i}\right]} \text{ and } \bar{\rho} = \frac{\mu\left[(\omega_\ell, \omega_\ell); e_N\right]}{\mu\left[(\omega_h, \omega_\ell); e_N\right] + \mu\left[(\omega_\ell, \omega_h); e_N\right]}$$

measure the correlation between the two team members' low outcomes, when one or both agents exert effort, respectively. Specifically, if at least one agent has a low outcome, then ρ and $\bar{\rho}$ equal the ratio between the probability that both agents had a low outcome relative to the probability that exactly one of them did. (Because of our symmetry assumption, ρ is independent of the choice of $i \in \{1, 2\}$.) The terms

$$\sigma = \frac{\mu\left[(\omega_i = \omega_h, \omega_{-i} = \omega_\ell); e_{N\setminus i}\right]}{\mu\left[(\omega_h, \omega_\ell); e_{N\setminus i}\right] + \mu\left[(\omega_\ell, \omega_h); e_{N\setminus i}\right]} \text{ and } \bar{\sigma} = \frac{\mu\left[(\omega_i = \omega_h, \omega_{-i} = \omega_\ell); e_N\right]}{\mu\left[(\omega_h, \omega_\ell); e_N\right] + \mu\left[(\omega_\ell, \omega_h); e_N\right]}$$

measure the probability that i has a high outcome, conditional on exactly one team member having a high outcome — again calculated if only team member $-i$ or both team members exert effort, respectively. If $\sigma < 1/2$, then team member $-i$ has a higher expected outcome than team member i when i does not exert effort; the opposite holds if $\sigma > 1/2$. Our symmetry assumption implies that $\bar{\sigma} = 1/2$. Therefore $\sigma < 1/2$, or equivalently $\bar{\sigma} - \sigma > 0$, indicates that by exerting effort team member i can balance the outcome distribution in their own favor; conversely, $\bar{\sigma} - \sigma < 0$ indicates that i's effort favors their partner $-i$. In sum,

- $\Delta_\sigma = \bar{\sigma} - \sigma$ measures the degree to which effort is self-improving.
- $\Delta_\rho = \bar{\rho} - \rho$ measures the degree to which effort correlates team members' outcome values.

Proposition 6 relates the effort-maximizing deliberation procedure to these features of the effort environment. Because there are only two individuals in a team, a symmetric deliberation procedure is described by a single parameter $D(1) \equiv D(\{1\}) = D(\{2\})$. Proposition 6 shows that the effort-maximizing value of $D(1)$ is increasing in Δ_σ and decreasing in Δ_ρ.

Proposition 6. *The effort-maximizing level of $D(1)$, D^*, is fully determined by $(\rho, \bar{\rho}, \sigma, \bar{\sigma})$. Fix $\bar{\rho}$ and $\bar{\sigma}$, and let ρ and σ vary.*

1. *If $\Delta_\rho = \bar{\rho} - \rho < 0$, $D^* \in \{0, 1\}$ and $D^* = 1$ if and only if*

$$\frac{\sigma}{\bar{\sigma}} \leqslant \frac{\rho + 1}{\bar{\rho} + 1}.$$

D^* is therefore non-decreasing in Δ_σ and non-increasing in Δ_ρ.

2. *If $\Delta_\rho > 0$, then D^* is a continuous non-decreasing function of Δ_σ, and a continuous non-increasing function of Δ_ρ.*

The proof of Proposition 6 is in the Appendix. This proposition complements the results in Section 4 by showing that, in a binary-outcome environment, the effort-maximizing deliberation procedure is "more unilateral" when effort is more self-improving or when effort correlates individuals' outcomes to a lesser extent.

5.3 Effort-Maximizing Deliberation in Larger Teams

We now consider, in a numerical exercise, effort-maximizing deliberation in teams with more than two members. To that end, we specify outcome distributions as follows. First, suppose that all team members exert effort. Then with probability $\bar{\rho} \in (0, 1)$, all team members receive the same outcome, so that either $\omega = (\omega_\ell, ..., \omega_\ell)$ (with probability h_T) or $\omega = (\omega_h, ..., \omega_h)$ (with probability $1 - h_T$). With complementary probability $1 - \rho$, each team member $i \in N$ draws their own outcome $\omega_i \in \{\omega_\ell, \omega_h\}$ independently; the probability of a high outcome for each individual is $\bar{h}$. If instead individual i deviates to no effort, the probability that all team members receive the same outcome is ρ, the probability that i receives an independent high outcome is h_i, and the probability that a team member $j \neq i$ receives an independent high outcome is h_j.

We use $\Delta_\sigma = (\bar{h} - h_i)/(\bar{h} - h_j)$ as a measure of how self-improving i's effort is; and $\Delta_\rho = \bar{\rho} - \rho$ as a measure of how much i's effort increases the correlation in individuals' outcomes. Further, we assume that the team must use a symmetric and deterministic deliberation procedure: $D(X) = 1$ if $|X| \geqslant K$ and $D(X) = 0$ if $|X| < K$, for some $1 \leqslant K \leqslant N$. The value K is therefore the number of individual recommendations required for an outcome to be disclosed. If $K = N$, then disclosure must be a decision made by consensus, and $K = 1$ corresponds to the unilateral disclosure protocol. We wish to assess K^*, the degree of consensus required for disclosure in the effort-maximizing deliberation procedure, and how it varies with Δ_ρ and Δ_σ.

In the Appendix, we calculate the "misattributed blame" component of individual effort incentives under different deliberation procedures and state a proposition ranking effort incentives provided by different consensus levels K. We use these expressions in our numerical exercise to determine the effort-maximizing required consensus K^*.

Figure 4 displays the numerical results, which are in line with the results shown for teams with $n = 2$. The parameters used in the simulation are specified in the figure, but the results

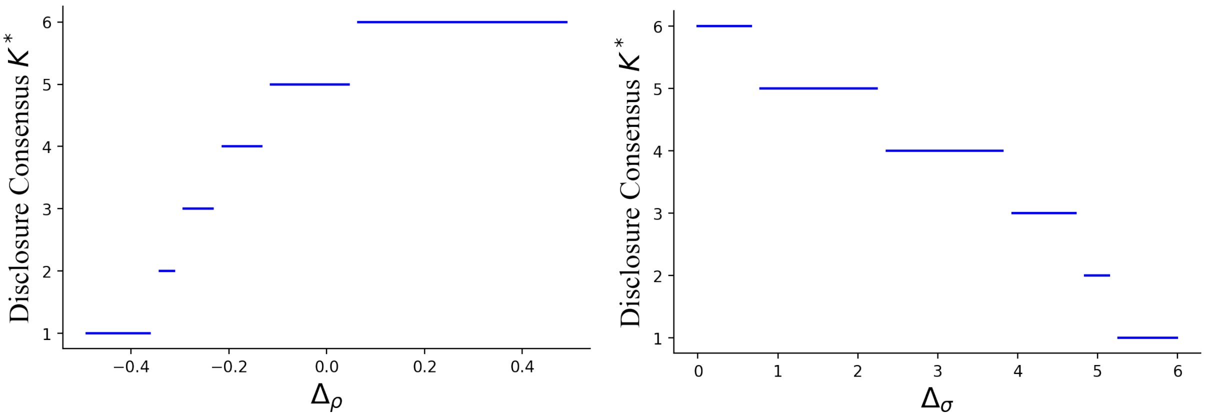


Figure 4: Effort-maximizing degree of consensus required for disclosure as a function of Δ_ρ and Δ_σ. In the left panel, $\bar{\rho} = .5$ and changes in ρ create the variation in Δ_ρ. Other parameters are fixed at $h_T = h_i = h_j = .5$ and $\bar{h} = .6$. In the right panel, $\bar{h} = .6$ and $h_j = .5$, and changes in h_i create the variation in Δ_σ. Other parameters are fixed at $h_T = .5$ and $\rho = \bar{\rho} = .5$. In both panels, the number of team members is set to $n = 10$.

are robust to various parameter specifications. We see that in effort environments in which i's effort more strongly correlates individuals' outcomes, it is best (in terms of effort incentives) to require higher degrees of consensus in order to disclose the team's outcome. And in more "self-improving" effort environments, it is best to require lower degrees of consensus for the team to choose to disclose an outcome.

References

Alchian, Armen A and Harold Demsetz (1972) "Production, Information Costs, and Economic Organization," *American Economic Review*, 62 (5), 777–795.

Ambrus, Attila, Eduardo M Azevedo, and Yuichiro Kamada (2013) "Hierarchical Cheap Talk," *Theoretical Economics*, 8 (1), 233–261.

Arya, Anil and Brian Mittendorf (2011) "The Benefits of Aggregate Performance Metrics in the Presence of Career Concerns," *Management Science*, 57 (8), 1424–1437.

Auriol, Emmanuelle, Guido Friebel, and Lambros Pechlivanos (2002) "Career Concerns in Teams," *Journal of Labor Economics*, 20 (2), 289–307.

Austen-Smith, David and Timothy Feddersen (2005) "Deliberation and Voting Rules," in *Social choice and Strategic Decisions: Essays in Honor of Jeffrey S. Banks*, 269–316.

Bar-Isaac, Heski (2007) "Something to Prove: Reputation in Teams," *The RAND Journal of Economics*, 38 (2), 495–511.

Bardhi, Arjada and Nina Bobkova (2023) "Local Evidence and Diversity in Minipublics," *Journal of Political Economy*, 131 (9), 2451–2508.

Battaglini, Marco (2002) "Multiple Referrals and Multidimensional Cheap Talk," *Econometrica*, 70 (4), 1379–1401.

Baumann, Leonie and Rohan Dutta (2022) "Strategic Evidence Disclosure in Networks and Equilibrium Discrimination," working paper.

Ben-Porath, Elchanan, Eddie Dekel, and Barton L Lipman (2018) "Disclosure and Choice," *Review of Economic Studies*, 85 (3), 1471–1501.

Chalioti, Evangelia (2016) "Team Production, Endogenous Learning about Abilities and Career Concerns," *European Economic Review*, 85, 229–244.

Che, Yeon-Koo and Donald B Hausch (1999) "Cooperative Investments and the Value of Contracting," *American Economic Review*, 89 (1), 125–147.

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole (1999) "The Economics of Career Concerns, Part I: Comparing Information Structures," *Review of Economic Studies*, 66 (1), 183–198.

Dranove, David and Ginger Zhe Jin (2010) "Quality Disclosure and Certification: Theory and Practice," *Journal of Economic Literature*, 48 (4), 935–963.

Dye, Ronald A (1985) "Disclosure of Nonproprietary Information," *Journal of Accounting Research*, 123–145.

Dziuda, Wioletta (2011) "Strategic Argumentation," *Journal of Economic Theory*, 146 (4), 1362–1397.

Fortunato, Santo, Carl T Bergstrom, Katy Börner et al. (2018) "Science of Science," *Science*, 359 (6379), eaao0185.

Freeland, Robert F and Ezra W Zuckerman Sivan (2018) "The Problems and Promise of Hierarchy: Voice Rights and the Firm," *Sociological Science*, 5, 143–181.

Gentzkow, Matthew and Emir Kamenica (2016) "Competition in Persuasion," *Review of Economic Studies*, 84 (1), 300–322.

Gerardi, Dino and Leeat Yariv (2007) "Deliberative Voting," *Journal of Economic theory*, 134 (1), 317–338.

——— (2008) "Information Acquisition in Committees," *Games and Economic Behavior*, 62 (2), 436–459.

Gibbons, Robert (2005) "Four Formal (izable) Theories of the Firm?" *Journal of Economic Behavior & Organization*, 58 (2), 200–245.

Giovannoni, Francesco and Daniel J Seidmann (2007) "Secrecy, Two-Sided bias and the Value of Evidence," *Games and economic behavior*, 59 (2), 296–315.

Grossman, Sanford J (1981) "The Informational Role of Warranties and Private Disclosure about Product Quality," *Journal of Law and Economics*, 24 (3), 461–483.

Grossman, Sanford J and Oliver D Hart (1986) "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94 (4), 691–719.

Guiso, Luigi, Paola Sapienza, and Luigi Zingales (2015) "The Value of Corporate Culture," *Journal of Financial Economics*, 117 (1), 60–76.

Hagenbach, Jeanne, Frédéric Koessler, and Eduardo Perez-Richet (2014) "Certifiable Pre-Play Communication: Full Disclosure," *Econometrica*, 82 (3), 1093–1131.

Holmström, Bengt (1982) "Moral Hazard in Teams," *The Bell Journal of Economics*, 324–340.

——— (1999) "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies*, 66 (1), 169–182.

Hu, Peicong and Joel Sobel (2019) "Simultaneous Versus Sequential Disclosure," working paper.

Itoh, Hideshi (1991) "Incentives to Help in Multi-Agent Situations," *Econometrica: Journal of the Econometric Society*, 611–636.

Jeon, Seonghoon (1996) "Moral Hazard and Reputational Concerns in Teams: Implications for Organizational Choice," *International Journal of Industrial Organization*, 14 (3), 297–315.

Jones, Benjamin F (2021) "The Rise of Research Teams: Benefits and Costs in Economics," *Journal of Economic Perspectives*, 35 (2), 191–216.

Kreps, David M and Robert Wilson (1982) "Sequential Equilibria," *Econometrica*, 863–894.

Lazear, Edward P and Kathryn L Shaw (2007) "Personnel Economics: The Economist's View of Human Resources," *Journal of Economic Perspectives*, 21 (4), 91–114.

Levy, Gilat (2007) "Decision Making in Committees: Transparency, Reputation, and Voting Rules," *American Economic Review*, 97 (1), 150–168.

Li, Kai, Feng Mai, Rui Shen, and Xinyan Yan (2021) "Measuring Corporate Culture Using Machine Learning," *Review of Financial Studies*, 34 (7), 3265–3315.

Martini, Giorgio (2018) "Multidimensional Disclosure."

Matthews, Steven and Andrew Postlewaite (1985) "Quality Testing and Disclosure," *The RAND Journal of Economics*, 328–340.

Milgrom, Paul (1981) "Good News and Bad News: Representation Theorems and Applications," *The Bell Journal of Economics*, 380–391.

Milgrom, Paul and John Roberts (1986) "Relying on the Information of Interested Parties," *The RAND Journal of Economics*, 18–32.

Name-Correa, Alvaro J and Huseyin Yildirim (2019) "Social Pressure, Transparency, and Voting in Committees," *Journal of Economic Theory*, 184, 104943.

Onuchic, Paula (2021) "Advisors with Hidden Motives," working paper.

Onuchic, Paula ⓡ Debraj Ray (2023) "Signaling and Discrimination in Collaborative Projects," *American Economic Review*, 113 (1), 210–252.

O'Reilly, Charles A and Jennifer A Chatman (1996) "Culture as Social Control: Corporations, Cults, and Commitment.," *Research in Organizational Behavior*, 18, 157–200.

Ortega, Jaime (2003) "Power in the Firm and Managerial Career Concerns," *Journal of Economics & Management Strategy*, 12 (1), 1–29.

Ozerturk, Saltuk and Huseyin Yildirim (2021) "Credit Attribution and Collaborative Work," *Journal of Economic Theory*, 195, 105264.

Ramos, Joao and Tomasz Sadzik (2023) "Partnership with Persistence," working paper.

Ray, Debraj ⓡ Arthur Robson (2018) "Certified Random: A New Order for Coauthorship," *American Economic Review*, 108 (2), 489–520.

Seidmann, Daniel J and Eyal Winter (1997) "Strategic Information Transmission with Verifiable Messages," *Econometrica*, 163–169.

Shaked, Moshe and J George Shanthikumar (2007) *Stochastic Orders*: Springer.

Shishkin, Denis (2021) "Evidence Acquisition and Voluntary Disclosure," working paper.

Song, Shin Hyun (1994) "The Burden of Proof in a Game of Persuasion," *Journal of Economic Theory*, 64 (1), 253–264.

Squintani, Francesco (2020) "Information Transmission in Political Networks," working paper.

Tamaseb, Ali (2021) *Super Founders: What Data Reveals about Billion-Dollar Startups*: PublicAffairs.

Visser, Bauke and Otto H Swank (2007) "On Committees of Experts," *Quarterly Journal of Economics*, 122 (1), 337–372.

Whitmeyer, Mark and Kun Zhang (2022) "Costly Evidence and Discretionary Disclosure," working paper.

Zuckerman, Ezra W (2010) "Chapter 16 Speaking with One Voice: a "Stanford School" Approach to Organizational Hierarchy," in *Stanford's Organization Theory Renaissance, 1970–2000*, 289–307: Emerald Group Publishing Limited.

A Proofs

A.1 Proof of Theorem 1

We prove the three statements in the Theorem separately.

A.1.1 Proof of Statement 1

It is easy to see that a full-disclosure equilibrium always exists, where $x_i(\omega) = 1$ for all $\omega \in \Omega$ and all $i \in N$, and $\omega_i^{ND} = \min\{\omega_i : \omega \in \Omega\}$. Given this vector of no-disclosure beliefs, "always disclose" is individual's as-if pivotal optimal behavior. The vector of no-disclosure beliefs is Bayes-consistent, because no-disclosure does not happen on-path. □

A.1.2 Proof of Statement 2

Let i be a team member who can unilaterally disclose. Suppose by contradiction that a partial-disclosure equilibrium exists in which $\omega_i^{ND} > \min(\Omega_i)$. Then person i's as-if pivotal disclosure recommendations must satisfy $x_i(\omega) = 1$ whenever $\omega_i > \omega_i^{ND}$; and because i can unilaterally choose disclosure, all such outcome realizations are disclosed. Consequently, all outcomes ω that are not disclosed with some probability must satisfy $\omega_i \leqslant \omega_i^{ND}$. Also note that if an outcome ω is not disclosed with some probability, then the outcome $\hat{\omega}$ with $\hat{\omega}_j = \omega_j$ for all $j \neq i$ and $\hat{\omega}_i = \min(\Omega_i)$ must also be concealed with equal or larger probability. These two observations imply that $\mathbb{E}[\omega_i | \text{ no disclosure}]$ is strictly smaller than the initially conjectured ω_i^{ND}, which contradicts that the initial conjecture was indeed an equilibrium.

Consequently, in all partial-disclosure equilibria, we must have $\omega_i^{ND} = \min(\Omega_i)$ for every team member i who can unilaterally choose disclosure. □

A.1.3 Proof of Statement 3

Define a map $\Phi : co(\Omega) \rightrightarrows co(\Omega)$, as follows:
For each $\bar{\omega} \in co(\Omega)$, $\hat{\omega} \in \Phi(\bar{\omega})$ if and only if there exists a vector x of individual disclosure recommendation strategies satisfying $x_i(\omega) = 0$ if $\omega_i = \min(\Omega_i)$, and

$$\omega_i > \bar{\omega}_i \Rightarrow x_i(\omega) = 1 \text{ and } \omega_i < \bar{\omega}_i \Rightarrow x_i(\omega) = 0,$$

and such that

$$\hat{\omega}_i = \frac{\sum_\Omega \omega_i(1 - d(\omega))\mu(\omega)}{\sum_\Omega (1 - d(\omega))\mu(\omega)}, \text{ where } d(\omega) = \sum_{X \subseteq N} \Pi_X(\omega) D(X) \text{ for every } \omega \in \Omega.$$

In words, Φ maps each "candidate vector" of equilibrium no-disclosure posteriors into a vector of "individually rational" no-disclosure posteriors which is consistent with the starting candidate vector. These "individually rational" posteriors are those consistent with agents' as-if pivotal optimal behavior given the candidate vector of no-disclosure beliefs. We allow individuals to use any mixed strategy if their realized outcome equals their candidate no-disclosure posterior, with the exception that individuals always recommend to not disclose if their worst possible outcome realizes. For any no-disclosure belief, it is weakly optimal for an agent to recommend the non-disclosure of their worst possible outcome; this restriction is therefore consistent with our equilibrium definition. Because we only wish to prove that an equilibrium exists, it is enough to show that an equilibrium exists in which agents use strategies in this restricted class.

First note that $\Phi(\bar{\omega})$ is non-empty for every $\bar{\omega} \in co(\Omega)$, because no-disclosure happens on path for all the described strategies — at the very least, all agents recommend non-disclosure when $\omega = (\min(\Omega_1), ..., \min(\Omega_N))$, and the team chooses no disclosure by consensus. Now observe that, because the construction of Φ allows individuals to use any mixed strategy when their realized outcome equals their candidate no-disclosure posterior, then $\Phi(\bar{\omega})$ is a closed set for all $\bar{\omega} \in co(\Omega)$; and Φ is upper-hemicontinuous. Therefore, Φ has a closed graph, and by the Kakutani fixed point theorem, Φ has a fixed point in $co(\Omega)$. It is easy to see that a fixed point of Φ defines an equilibrium of the team-disclosure game.

Now let $I \subseteq N$ be the set of team members who cannot unilaterally choose disclosure. We will argue that there must be a fixed point w of Φ with $w_i > \min(\Omega_i)$ for all $i \in I$. To that end, let $w \in \Phi(w)$ be a fixed point of Φ. Then it must be that there is a vector of individual disclosure recommendation strategies x satisfying $x_i(\omega) = 0$ if $\omega_i = \min(\Omega_i)$ such that for every $i \in N$,

$$w_i = \frac{\sum_\Omega \omega_i (1 - d(\omega))\mu(\omega)}{\sum_\Omega (1 - d(\omega))\mu(\omega)}, \text{ where } d(\omega) = \sum_{X \subseteq N} \Pi_X(\omega) D(X) \text{ for every } \omega \in \Omega.$$

Now take $i \in I$; it must be that all realizations ω with $\omega_j = \min(\Omega_j)$ for every $j \neq i$ are not disclosed — regardless of the realization of ω_i — because i cannot unilaterally choose disclosure. Consequently, every possible realized outcome for individual i is not disclosed with positive probability, and therefore $w_i > \min(\Omega_i)$. This fixed-point of Φ thus defines a partial-disclosure equilibrium in which $\omega_i^{ND} > \min(\Omega_i)$ for every $i \in I$. □

□

A.2 Proof of Lemma 1

For each team member $i \in N$, we index the possible outcomes for player i—that is, for $k \in \{1, ..., |\Omega_i|\}$, we denote by ω_i^k the k^{th} lowest value in Ω_i.

For the first claim of the lemma, fix a full-support outcome distribution μ. Consider the class of deliberation procedures $\mathcal{D}_\epsilon$ such that for some $\epsilon \in (0, 1)$, $\epsilon \leqslant D(I) < 1$ for every $I \notin \{\varnothing, N\}$. We want to show that if ϵ is large enough, there exists a strict equilibrium for any deliberation procedure in that class. Consider individual disclosure recommendation strategies such that for each i, $x_i(\omega) = 0$ if $\omega_i = \omega_i^1$ and $x_i(\omega) = 1$ otherwise. These strategies constitute a strict equilibrium if and only if, for each i, the implied no-disclosure beliefs satisfy $\omega_i^{ND} \in (\omega_i^1, \omega_i^2)$.

We now show that, if players follow such strategies and $\epsilon < 1$ high enough, then for any deliberation procedure in the introduced class, indeed $\omega_i^{ND} \in (\omega_i^1, \omega_i^2)$. First, note that because $\epsilon < 1$, some outcomes that are not the worst for player i are not disclosed, and hence for any $\epsilon \in (0, 1)$ and deliberation procedure $D \in \mathcal{D}_\epsilon$, $\omega_i^{ND} > \omega_i^1$. Next, because i recommends disclosure of all outcomes with $\omega_i > \omega_i^1$, it must be that $d(\omega) \geqslant \epsilon$ for all such ω. In contrast, if $\omega_i = \omega_i^1$ for all i, then $d(\omega) = 0$. Consequently, if $\epsilon \in (0, 1)$ is sufficiently large then we must have that for any deliberation procedure $D \in \mathcal{D}_\epsilon$, $\omega_i^{ND} < \omega_i^2$.

For the second claim of the lemma, fix a deliberation procedure D with $D(\{i\}) < 1$ for every i. Consider the class of outcome distributions $\mathcal{U}_\epsilon$ satisfying $\mu(\omega) > 0$ for all $\omega \in \Omega$ and $\mu(\omega) < \epsilon$ for all $\omega \neq (\min(\Omega_1), ..., \min(\Omega_n))$. We want to show that if ϵ is small enough then the individual strategies $x_i(\omega) = 0$ if $\omega_i = \omega_i^1$ and $x_i(\omega) = 1$ otherwise, for each i, constitute a strict equilibrium for any outcome distribution in the class. As before, this strategy is a strict equilibrium if, and only if, for each i, the implied no-disclosure beliefs satisfy $\omega_i^{ND} \in (\omega_i^1, \omega_i^2)$.

We now show that, if players follow the given strategies and $\epsilon > 0$ is small enough, then for any outcome distribution $\mu \in \mathcal{U}_\epsilon$, indeed $\omega_i^{ND} \in (\omega_i^1, \omega_i^2)$. First, although i recommends disclosure of all outcomes with $\omega_i > \omega_i^1$, some such outcomes are not disclosed (because $D(\{i\}) < 1$), and hence $\omega_i^{ND} > \omega_i^1$. Second, the outcome $\omega = (\omega_1^1, ...\omega_n^1)$ is not disclosed with certainty, because every team member recommends its non-disclosure. Further, as $\epsilon \to 0$, the probability of outcomes with $\omega_i > \omega_i^1$ also approaches 0. As a consequence, as $\epsilon \to 0$, it must be that $\max_{\mu \in \mathcal{U}_\epsilon} \omega_i^{ND}(\mu) \to \omega_i^1$. And indeed, for a small enough ϵ, we have $\omega_i^{ND} \in (\omega_i^1, \omega_i^2)$ for each $i \in N$ for any outcome distribution in $\mathcal{U}_\epsilon$.

□

A.3 Proof of Lemma 2

At the fixed strict equilibrium, we have

$$\omega_i^{ND} = \frac{\sum_\Omega \omega_i (1-d(\omega))\mu(\omega)}{\sum_\Omega (1-d(\omega))\mu(\omega)}, \text{ where } d(\omega) = \sum_{X\subseteq N} \Pi_X(\omega) D(X) \text{ for every } \omega \in \Omega,$$

where $\Pi_X(\cdot)$ is determined by the equilibrium individual recommendation strategies. Consider a different deliberation procedure $\hat{D}$, and let for each $i \in N$,

$$\hat{\omega}_i^{ND} = \frac{\sum_\Omega \omega_i (1-\hat{d}(\omega))\mu(\omega)}{\sum_\Omega (1-\hat{d}(\omega))\mu(\omega)}, \text{ with } \hat{d}(\omega) = \sum_{X\subseteq N} \Pi_X(\omega) \hat{D}(X) \text{ for every } \omega \in \Omega, \tag{6}$$

where $\hat{d}$ is calculated under the same original equilibrium individual recommendation strategies, but the new deliberation procedure $\hat{D}$. It is easy to see that for every $\epsilon > 0$, there is some $\delta > 0$ such that $e(D, \hat{D}) < \delta$ implies $e(\omega^{ND}, \hat{\omega}^{ND}) < \epsilon$, where e indicates the Euclidian distance.

For each team member $i \in N$, for $k \in \{1, ..., |\Omega_i|\}$, we denote by ω_i^k the k^{th} lowest value in Ω_i. Then take $\hat{D}$ with $e(D, \hat{D})$ small enough so that for every $i \in N$, $\omega_i^k < \omega_i^{ND} < \omega_i^{k+1}$ implies $\omega_i^k < \hat{\omega}_i^{ND} < \omega_i^{k+1}$. Therefore the equilibrium disclosure recommendation strategies given ω^{ND} are also equilibrium recommendation strategies given $\hat{\omega}^{ND}$; and so $\hat{\omega}^{ND}$ is a strict equilibrium under $\hat{D}$, $\hat{\omega}^{ND} \in \mathcal{E}^\mu_{\hat{D}}$. And so for a small enough neighborhood of D, there exists a continuous selection E of the strict-equilibrium correspondence such that $E(D) = \varepsilon$.

Now take $\bar{\delta}$ such that $e(D, \hat{D}) < \bar{\delta}$ implies $e(E(D), E(\hat{D})) < \min_{i,k} \frac{|\omega_i^{ND} - \omega_i^k|}{2} = \bar{\epsilon}$. Now fix $\delta < \bar{\delta}$ and let E' be a selection of the strict-equilibrium correspondence that is continuous in a δ-neighborhood of D with $E'(D) = \omega^{ND}$. We'll use the following claim.

Claim 1. *If $\hat{D}$ is such that $\hat{\omega}^{ND} = E'(\hat{D})$ satisfies*

$$\omega_i^k < \hat{\omega}_i^{ND} < \omega_i^{k+1} \textit{ for each } i \in N, \tag{7}$$

where ω_i^k and ω_i^{k+1} are such that $\omega_i^k < \omega_i^{ND} < \omega_i^{k+1}$ — then $E'(\hat{D}) = E(\hat{D})$.

Proof of Claim. Suppose $\omega_i^k < \hat{\omega}_i^{ND} < \omega_i^{k+1}$ for each i — where ω_i^k and ω_i^{k+1} are such that $\omega_i^k < \omega_i^{ND} < \omega_i^{k+1}$. Then it must be that in the equilibrium $E'(\hat{D})$, every team member uses the same recommendation strategy as in the original equilibrium $\omega^{ND} \in \mathcal{E}^\mu_D$. And consequently $\hat{\omega}^{ND}$ must satisfy (6), and thus $E'(\hat{D}) = E(\hat{D})$. □

Let $\mathcal{D}$ be the set of deliberation procedures for which (7) holds. If there is some D' with $e(D, D') < \delta$ such that (7) does not hold, then it must be that E' is discontinuous at the boundary of $\mathcal{D}$ — because $E'(\hat{D}) = E(\hat{D})$, by Claim 1, and $e(E(D), E(\hat{D})) < \bar{\epsilon}$ for every $\hat{D} \in \mathcal{D}$. This contradicts the assumption that E' is continuous. Therefore, it must be that $\mathcal{D}$ is the entire δ-neighborhood of D over which E' is defined; and we conclude that $E' = E$.

□

A.4 Proof of Proposition 1

Fixing a starting strict equilibrium with no-disclosure beliefs ω^{ND}, we can partition each team member's outcome realizations Ω_i into low realizations with $\omega_i < \omega_i^{ND}$ and high realizations with $\omega_i > \omega_i^{ND}$. Accordingly, for each team outcome realization $\omega \in \Omega$, we can define the set of team members for which this realization was high: $H(\omega) = \{i \in N : \omega_i > \omega_i^{ND}\}$. Remember that the distribution of outcomes is μ. With a slight abuse of notation, for any given set $I \subseteq N$, we let $\mu\left(H(\omega) = I\right)$ be the probability that an outcome ω realizes which is a high realization for exactly team members I.

Lemma 5. *Fix a starting deliberation procedure D and a strict equilibrium ω^{ND}. Let $dD = (dD(I))_{I \subseteq N}$ be a marginal change to the deliberation procedure. Then we have, for each $i \in N$,*

$$d\omega_i^{ND} = \sum_{I \subseteq N} \frac{\mu(H(\omega) = I)}{\sum_{I' \subseteq N} \mu(H(\omega) = I')(1 - D(I'))} \left[\omega_i^{ND} - \mathbb{E}\left(\omega_i | H(\omega) = I\right)\right] dD(I). \tag{8}$$

Proof of Lemma. We can write ω_i^{ND} as

$$\omega_i^{ND} = \frac{\sum_{I \subseteq N} \mu(H(\omega) = I)(1 - D(I))\mathbb{E}(\omega_i | H(\omega) = I)}{\sum_{I \subseteq N} \mu(H(\omega) = I)(1 - D(I))},$$

where note that we do not have to consider agent mixed-strategies because the equilibrium is strict. Now note that small variations in the protocol D only change individual disclosure strategies for zero-measure sets of outcome realizations — because the original equilibrium is strict. Therefore the change in ω_i^{ND} can be computed only as its "direct effect," as follows.

$$\begin{aligned} d\omega_i^{ND} = &\sum_{I \subseteq N} \Bigg[\frac{-\mu(H(\omega) = I)\mathbb{E}(\omega_i | H(\omega) = I)}{\sum_{I' \subseteq N} \mu(H(\omega) = I')(1 - D(I'))} \\ &+ \mu(H(\omega) = I) \frac{\sum_{I' \subseteq N} \mu(H(\omega) = I')(1 - D(I'))\mathbb{E}(\omega_i | H(\omega) = I')}{\left[\sum_{I' \subseteq N} \mu(H(\omega) = I')(1 - D(I'))\right]^2}\Bigg] dD(I) = \\ = &\sum_{I \subseteq N} \frac{\mu(H(\omega) = I)}{\sum_{I' \subseteq N} \mu(H(\omega) = I')(1 - D(I'))} \left[\omega_i^{ND} - \mathbb{E}\left(\omega_i | H(\omega) = I\right)\right] dD(I). \end{aligned}$$

□

Back to the proof of the proposition. Suppose the first condition, condition (2), holds. Let $m = \min\left\{\frac{dD(I)}{1-D(I)} : i \in I\right\}$ and $M = \max\left\{\frac{dD(I)}{1-D(I)} : i \notin I\right\}$, so that $m \geqslant M$. Then, using equation (8), we have

$$d\omega_i^{ND} = \sum_{I\subseteq N} \frac{\mu(H(\omega)=I)(1-D(I))}{\sum_{I'\subseteq N}\mu(H(\omega)=I')(1-D(I'))}\left[\omega_i^{ND} - \mathbb{E}\left(\omega_i|H(\omega)=I\right)\right]\frac{dD(I)}{(1-D(I))}$$

$$\leqslant m\left[\sum_{i\in I\subseteq N} \frac{\mu(H(\omega)=I)(1-D(I))}{\sum_{I'\subseteq N}\mu(H(\omega)=I')(1-D(I'))}\left(\omega_i^{ND} - \mathbb{E}\left(\omega_i|H(\omega)=I\right)\right)\right]$$
$$+ M\left[\sum_{i\notin I\subseteq N} \frac{\mu(H(\omega)=I)(1-D(I))}{\sum_{I'\subseteq N}\mu(H(\omega)=I')(1-D(I'))}\left(\omega_i^{ND} - \mathbb{E}\left(\omega_i|H(\omega)=I\right)\right)\right]$$

$$\leqslant m\left[\sum_{I\subseteq N} \frac{\mu(H(\omega)=I)(1-D(I))}{\sum_{I'\subseteq N}\mu(H(\omega)=I')(1-D(I'))}\left(\omega_i^{ND} - \mathbb{E}\left(\omega_i|H(\omega)=I\right)\right)\right]$$

$$= m\left[\omega_i^{ND} - \sum_{I\subseteq N}\frac{\mu(H(\omega)=I)(1-D(I))\mathbb{E}\left(\omega_i|H(\omega)=I\right)}{\sum_{I'\subseteq N}\mu(H(\omega)=I')(1-D(I'))}\right] = 0.$$

The inequalities follow from (2) and the fact that $\omega_i^{ND} \leqslant \mathbb{E}(\omega_i|H(\omega) = I)$ if $i \in I$ and $\omega_i^{ND} \geqslant \mathbb{E}(\omega_i|H(\omega) = I)$ if $i \notin I$. The last equality follows from the definition of ω_i^{ND}. Following analogous steps, it is easy to see that if condition (3) holds, then $d\omega_i^{ND} \geqslant 0$.

□

A.5 Proof of Theorem 2

The proof uses the following auxiliary lemma:

Lemma 6. *D is such that disclosing requires more consensus than concealing if and only if for all $I \subseteq N$ such that $D(I) = 1$ and $D(N \setminus I) < 1$, there exists $J \subset I$ such that $D(N \setminus J) < 1$.*

Proof of Lemma 6. One direction (⇒) is trivial given the definition of "disclosing requires more consensus than concealing." Consider the other direction (⇐). Suppose D is such that for all $I \subseteq N$ such that $D(I) = 1$ and $D(N \setminus I) < 1$, there exists $J \subset I$ such that $D(N \setminus J) < 1$. Fix one such I, and take a particular $J \subset I$ such that $D(N \setminus J) < 1$.

If $D(J) < 1$, then the proof is done. Suppose instead that $D(J) = 1$. Then there must be some $K \subset J$ such that $D(N \setminus K) < 1$. If $D(K) < 1$, then the proof is done. If not, we can repeat this procedure until we find such a subset of I, L, such that $D(N \setminus L) < 1$ and $D(L) < 1$. The repetition of the procedure must return such a subset L of I because (i) the set of team members is finite, (ii) the procedure D is not the unilateral disclosure procedure,[29] and therefore for any individual i, $D(\{i\}) < 1$. □

We can now prove the theorem in two parts:

Part 1 (⇒) . If disclosing does not require more consensus than concealing, then there exists a full disclosure equilibrium that is consistent with deliberation.

Suppose D is such that disclosing *does not* require more consensus than concealing. By Lemma 6, we know that there exists some subgroup $I \subset N$ such that $D(I) = 1$, $D(N \setminus I) < 1$ and, for all $J \subset I$, $D(N \setminus J) = 1$. Consider a candidate full-disclosure equilibrium where $x_i(\omega) = 1$ for all $\omega \in \Omega$ and all $i \in I$ (we will specify the other agents' individual disclosure strategies later). These disclosure strategies aggregated according to the given deliberation protocol guarantee that all evidence is disclosed. Moreover, in this candidate equilibrium, we conjecture that the (off-path) no-disclosure beliefs are $\omega_i^{ND} = \min(\Omega_i)$ for each $i \in I$.

We want to build another vector of individual disclosure strategies to be used to "justify" these off-path beliefs. We do so as follows: for every $i \in I$, let $\hat{x}_i(\omega) = 0$ if $\omega_i = \min(\Omega_i)$ and $\hat{x}_i(\omega) = 1$ otherwise. And for every $j \in N \setminus I$, $\hat{x}_j(\omega) = 1$ for all $\omega \in \Omega$. Given the deliberation protocol, the team disclosure strategy implied by $\hat{x}$ satisfies $d(\omega) = D(\hat{x}(\omega)) = 0$ if and only if $\omega_i = \min(\Omega_i)$ for all $i \in I$. And therefore Bayes updating implies that $\omega_i^{ND} = \min(\Omega_i)$ for all $i \in I$. To complete the construction of the equilibrium, for every $j \in N \setminus I$, let ω_j^{ND} be the Bayes-consistent no-disclosure beliefs implied by $\hat{x}$. And for every $j \in N \setminus I$, let their equilibrium individual disclosure strategy be $x_j(\omega) = 1$ if $\omega_j \geqslant \omega_j^{ND}$ and $x_j(\omega) = 0$ otherwise.

Part 2 (⇐) . If disclosure requires more consensus than concealing, then there is no full-disclosure equilibrium that is consistent with deliberation.

Let D be such that disclosure requires more consensus than concealing. And suppose a vector x of individual disclosure strategies and a vector ω^{ND} of no-disclosure posteriors constitute a full-disclosure equilibrium. Let $I \subset N$ be the largest subgroup of team members such that

$$\omega_i^{ND} = \min(\Omega_i) \text{ for all } i \in I.$$

First note that the set I is non-empty, and $D(I) = 1$. To see this, suppose towards a contradiction that $D(I) < 1$ (which would vacuously hold if I were empty). Then every member of subgroup

[29] The unilateral procedure does not satisfy our starting assumption that for all $I \subseteq N$ such that $D(I) = 1$ and $D(N \setminus I) < 1$, there exists $J \subset I$ such that $D(N \setminus J) < 1$.

$N \setminus I$ strictly prefers to not disclose all realizations ω where $\omega_i = \min(\Omega_i)$ for every $i \in N \setminus I$. And moreover, because $D(I) < 1$, the subgroup $N \setminus I$ is able to block the disclosure of such realizations with positive probability. This contradicts the assumption that the starting equilibrium has full-disclosure.

Now take a vector of individual disclosure strategies $\hat{x}$ to be used as a candidate to "justify" the off-path no-disclosure beliefs ω^{ND}. Take some $\hat{\omega} \in \Omega$ with $\hat{\omega}_i = \min(\Omega_i)$ for every $i \in I$, and such that $d(\hat{\omega}) = D(\hat{x}(\hat{\omega})) < 1$ — such a $\hat{\omega}$ must exist if $\hat{x}$ is to justify the conjectured no-disclosure beliefs. Let I' be the set of team members such that $\hat{x}_i(\hat{\omega}) < 1$ for $i \in I'$. We consider two cases.

Case 1. Suppose there is some $i^* \in I \setminus I'$; that is, there is some $i^* \in I$ such that $\hat{x}_{i^*}(\hat{\omega}) = 1$. Then there must exist some $\hat{\omega}'$ with $\hat{\omega}'_i = \hat{\omega}_i$ for all $i \in N \setminus \{i^*\}$ and $\hat{\omega}'_{i^*} \neq \hat{\omega}_{i^*}$ such that $d(\hat{\omega}') \leqslant d(\hat{\omega}) < 1$ (because each individual strategy depends only on their own realized outcome). But note that, because $\hat{\omega}'_{i^*} \neq \hat{\omega}_{i^*}$, then it must be that $\hat{\omega}'_{i^*} > \min(\Omega_{i^*})$; and therefore the no-disclosure posterior about i^*'s outcome implied by $\hat{x}$ cannot be $\min(\Omega_{i^*})$.

Case 2. Suppose instead that $I \subseteq I'$ (and therefore $I \setminus I' = \varnothing$).

In this case, it must be that $D(I') = 1$ — because $D(I) = 1$ and the deliberation procedure is monotonic — and $D(N \setminus I') < 1$ by construction, since we assumed that $d(\hat{\omega}) < 1$. And therefore, because D is such that disclosure requires more consensus than concealing, there exists some $I'' \subset I'$ such that $D(N \setminus I'') < 1$. If $I \setminus I'' = \varnothing$, then I'' itself must have a subset I''' such that $D(N \setminus I''') < 1$. By iterating this process, we note that there is some $J \subset I'$ such that $D(N \setminus J) < 1$ and $I \setminus J \neq \varnothing$.

Then take some $i^* \in I \setminus J$. There exists some $\hat{\omega}'$ with $\hat{\omega}'_i = \hat{\omega}_i$ for all $i \in N \setminus \{i^*\}$ and $\hat{\omega}'_{i^*} = \hat{\omega}_{i^*}$ such that $d(\hat{\omega}') \leqslant d(\hat{\omega}) < 1$ (because each individual strategy depends only on their own realized outcome). But then it must be that $\hat{\omega}'_{i^*} > \min(\Omega_{i^*})$; and therefore the no-disclosure posterior about i^*'s outcome implied by $\hat{x}$ cannot be $\min(\Omega_{i^*})$.

Combining cases 1 and 2, we conclude that there is no vector of individual disclosure rules $\hat{x}$, with each individual strategy depending only on their own realized outcome, that can "justify" the conjectured no-disclosure posteriors as consistent with the deliberation protocol. And this is true for any conjectured full-disclosure equilibrium. Consequently, there is no full-disclosure equilibrium that is consistent with deliberation. □

A.6 Proof of Lemma 3

Fix a vector of effort costs $c \in \mathbb{R}^n_{++}$. Suppose team member i anticipates that every other team member will choose $e_j = 1$ (for $j \neq i$), and that the team disclosure strategy will be d. Then i's payoff from choosing effort $e_i = 1$ is

$$\sum_{\Omega} \omega_i d(\omega)\mu(\omega; e_N) + \sum_{\Omega}(1 - d(\omega))\omega_i^{ND}\mu(\omega; e_N) - c_i$$

$$= \sum_{\Omega} \omega_i d(\omega)\mu(\omega; e_N) + \sum_{\Omega}(1 - d(\omega)) \left[\frac{\sum_{\Omega} \omega_i(1 - d(\omega))\mu(\omega; e_N)}{\sum_{\Omega}(1 - d(\omega))\mu(\omega; e_N)}\right] \mu(\omega; e_N) - c_i \tag{9}$$

Now adding and subtracting $\sum_{\Omega} \omega_i(1 - d(\omega))\mu(\omega; e_N)$, (9) equals

$$\sum_{\Omega} \omega_i\mu(\omega; e_N) + \sum_{\Omega}(1 - d(\omega)) \left[\frac{\sum_{\Omega} \omega_i(1 - d(\omega))\mu(\omega; e_N)}{\sum_{\Omega}(1 - d(\omega))\mu(\omega; e_N)} - \omega_i\right] \mu(\omega; e_N) - c_i$$

$$= \sum_{\Omega} \omega_i\mu(\omega; e_N) - c_i,$$

where the equality uses the fact that $\sum_{\Omega}(1 - d(\omega)) \left[\frac{\sum_{\Omega} \omega_i(1-d(\omega))\mu(\omega;e_N)}{\sum_{\Omega}(1-d(\omega))\mu(\omega;e_N)} - \omega_i\right] \mu(\omega; e_N) = 0$.

And i's payoff from choosing effort $e_i = 0$ is

$$\sum_{\Omega} \omega_i\mu(\omega; e_{N\setminus i}) + \sum_{\Omega}(1 - d(\omega)) \left[\frac{\sum_{\Omega} \omega_i(1 - d(\omega))\mu(\omega; e_N)}{\sum_{\Omega}(1 - d(\omega))dF(\omega; e_N)} - \omega_i\right] \mu(\omega; e_{N\setminus i}),$$

where note that in the second term the distribution of outcomes is affected by i's effort choice, but the value of ω_i^{ND} is still calculated under the presumption that $e_i = 1$, for the deviation to $e_i = 0$ is not seen by the observer. Therefore, there is an equilibrium of the effort-choice stage where every team member exerts effort if and only if for every $i \in N$,

$$\sum_{\Omega} \omega_i \left[\mu(\omega; e_N) - \mu(\omega; e_{N\setminus i}))\right] + \sum_{\Omega}(1 - d(\omega)) \left[\omega_i - \frac{\sum_{\Omega} \omega_i(1 - d(\omega))\mu(\omega; e_N)}{\sum_{\Omega}(1 - d(\omega))\mu(\omega; e_N)}\right] \mu(\omega; e_{N\setminus i})$$
$$\geqslant c_i.$$

Or equivalently if and only if

$$-\sum_{\Omega}(1 - d(\omega))\mu(\omega; e_{N\setminus i}) \left[\frac{\sum_{\Omega} \omega_i(1 - d(\omega))\mu(\omega; e_N)}{\sum_{\Omega}(1 - d(\omega))\mu(\omega; e_N)} - \frac{\sum_{\Omega} \omega_i(1 - d(\omega))\mu(\omega; e_{N\setminus i})}{\sum_{\Omega}(1 - d(\omega))\mu(\omega; e_{N\setminus i})}\right]$$

$$+\sum_{\Omega}\omega_i\left[\mu(\omega;e_N)-\mu(\omega;e_{N\setminus i}))\right]\geqslant c_i, \text{ for every } i\in N.$$

□

A.7 Rewriting Equation (4) as (5)

The left-hand side of equation (4) is $\mathbb{E}\left(\omega_i|e_N\right)-\mathbb{E}\left(\omega_i|e_{N\setminus i}\right)-\mathbb{P}\left(ND|e_{N\setminus i}\right)[\mathbb{E}\left(\omega_i|ND;e_N\right)-\mathbb{E}\left(\omega_i|ND;e_{N\setminus i}\right)]$. Or equivalently,

$$\sum_{\Omega}\omega_i\left[\mu(\omega;e_N)-\mu(\omega;e_{N\setminus i})\right]+\sum_{\Omega}\omega_i(1-d(\omega))\mu(\omega;e_{N\setminus i})$$
$$-\frac{\sum_{\Omega}(1-d(\omega))\mu(\omega;e_{N\setminus i})}{\sum_{\Omega}(1-d(\omega))\mu(\omega;e_N)}\sum_{\Omega}\omega_i(1-d(\omega))\mu(\omega;e_N)$$

$$=\sum_{\Omega}d(\omega)\mu(\omega;e_{N\setminus i})\left[\sum_{\Omega}\omega_i\mu(\omega;e_N)-\sum_{\Omega}\omega_i\mu(\omega;e_{N\setminus i})\right]$$
$$+\sum_{\Omega}(1-d(\omega))\mu(\omega;e_{N\setminus i})\sum_{\Omega}\omega_i\mu(\omega;e_N)-\sum_{\Omega}(1-d(\omega))\mu(\omega;e_{N\setminus i})\sum_{\Omega}\omega_i\mu(\omega;e_{N\setminus i})$$
$$+\sum_{\Omega}\omega_i(1-d(\omega))\mu(\omega;e_{N\setminus i})-\frac{\sum_{\Omega}(1-d(\omega))\mu(\omega;e_{N\setminus i})}{\sum_{\Omega}(1-d(\omega))\mu(\omega;e_N)}\sum_{\Omega}\omega_i(1-d(\omega))\mu(\omega;e_N)$$

$$=\sum_{\Omega}d(\omega)\mu(\omega;e_{N\setminus i})\left[\sum_{\Omega}\omega_i\mu(\omega;e_N)-\sum_{\Omega}\omega_i\mu(\omega;e_{N\setminus i})\right]$$
$$+\sum_{\Omega}\omega_i(1-d(\omega))\mu(\omega;e_{N\setminus i})-\sum_{\Omega}(1-d(\omega))\mu(\omega;e_{N\setminus i})\sum_{\Omega}\omega_i\mu(\omega;e_{N\setminus i})$$
$$+\frac{\sum_{\Omega}(1-d(\omega))\mu(\omega;e_{N\setminus i})}{\sum_{\Omega}(1-d(\omega))\mu(\omega;e_N)}\left[\sum_{\Omega}(1-d(\omega))\mu(\omega;e_N)\sum_{\Omega}\omega_i\mu(\omega;e_N)-\sum_{\Omega}\omega_i(1-d(\omega))\mu(\omega;e_N)\right]$$

$$=\left[1-\mathbb{P}\left(ND|e_{N\setminus i}\right)\right]\left[\mathbb{E}\left(\omega_i|e_N\right)-\mathbb{E}\left(\omega_i|e_{N\setminus i}\right)\right]+Cov\left(\omega_i,(1-d)|e_{N\setminus i}\right)$$
$$-\frac{\mathbb{P}\left(ND|e_{N\setminus i}\right)}{\mathbb{P}\left(ND|e_N\right)}Cov\left(\omega_i,(1-d)|e_N\right)$$

$$=\left[1-\mathbb{P}\left(ND|e_{N\setminus i}\right)\right]\left[\mathbb{E}\left(\omega_i|e_N\right)-\mathbb{E}\left(\omega_i|e_{N\setminus i}\right)\right]$$

$$+\frac{\mathbb{P}\left(ND|e_{N\setminus i}\right)}{\mathbb{P}\left(ND|e_N\right)}Cov\left(\omega_i,d|e_N\right)-Cov\left(\omega_i,d|e_{N\setminus i}\right),$$

which is the expression in (5).

A.8 Proof of Proposition 2

A.8.1 Proof of Statement 1

The proof of statement 1 uses the following lemma.

Lemma 7. *Suppose effort is purely self-improving, and consider a team disclosure strategy d which is an equilibrium of the disclosure subgame given full effort. For all $i \in N$:*

- *If $\omega_i^{ND} = \min(\Omega_i)$, then*

$$\mathbb{E}\left(\omega_i | ND; e_{N\setminus i}\right) = \mathbb{E}\left(\omega_i | ND; e_N\right) = \omega_i^{ND} = \min(\Omega_i).$$

- *If $\omega_i^{ND} > \min(\Omega_i)$, then*

$$\mathbb{E}\left(\omega_i | ND; e_{N\setminus i}\right) < \mathbb{E}\left(\omega_i | ND; e_N\right).$$

Proof of Lemma. Fix team member i, and let $\nu_{\omega_{-i}}(\cdot) = \mu_i(\cdot|\omega_{-i}; e_N)$, $\hat{\nu}_{\omega_{-i}}(\cdot) = \mu_i(\cdot|\omega_{-i}; e_{N\setminus i})$ for all $\omega_{-i} \in \Omega_{N\setminus i}$ and $\eta(\cdot) = \mu_{N\setminus i}(\cdot; e_N)$, (which implies also that $\eta(\cdot) = \mu_{N\setminus i}(\cdot; e_{N\setminus i})$). Then we have

$$\mathbb{E}\left(\omega_i | ND; e_{N\setminus i}\right) = \frac{\sum_{\Omega_{N\setminus i}} \sum_{\Omega_i} \omega_i (1 - d(\omega_i, \omega_{-i}))\hat{\nu}_{\omega_{-i}}(\omega_i)\eta(\omega_{-i})}{\sum_{\Omega_{N\setminus i}} \sum_{\Omega_i} (1 - d(\omega_i, \omega_{-i}))\hat{\nu}_{\omega_{-i}}(\omega_i)\eta(\omega_{-i})}. \tag{10}$$

$$\mathbb{E}\left(\omega_i | ND; e_N\right) = \frac{\sum_{\Omega_{N\setminus i}} \sum_{\Omega_i} \omega_i (1 - d(\omega_i, \omega_{-i}))\nu_{\omega_{-i}}(\omega_i)\eta(\omega_{-i})}{\sum_{\Omega_{N\setminus i}} \sum_{\Omega_i} (1 - d(\omega_i, \omega_{-i}))\nu_{\omega_{-i}}(\omega_i)\eta(\omega_{-i})}. \tag{11}$$

If d is such that the expectation in (11) equals $\min(\Omega_i)$, then it must be that the expectation in (10) equals $\min(\Omega_i)$ as well. This proves the first statement in the proposition.

Suppose instead that d is such that $\mathbb{E}\left(\omega_i | ND; e_N\right) = \omega_i^{ND} > \min(\Omega_i)$. Because d is the aggregation of individual disclosure recommendations where, for each i, x_i depends only on ω_i, we can re-express d as follows. For each outcome $\omega_{-i} \in \Omega_{N\setminus i}$, there exist $\alpha(\omega_{-i}) \in [0, 1]$ and $\beta(\omega_{-i}) \in [0, 1]$ such that if $\omega = (\omega_i, \omega_{-i})$, then $(1 - d(\omega)) = \alpha(\omega_{-i}) + (1 - x_i(\omega_i))\beta(\omega_{-i})$.[30] (Because $\omega_i^{ND} > \min(\Omega_i)$, we also know that $\alpha(\omega_{-i}) > 0$ for some $\omega_{-i} \in \Omega_{N\setminus i}$.)

[30] To see this, remember that $d(\omega) = \sum_{X\subseteq N} \Pi_X(\omega) D(X)$, where for each $X \in N$, $\Pi_X(\omega) = \prod_{j\in N} x_j(\omega)^{\mathbb{1}[j\in X]}(1 - x_j(\omega))^{\mathbb{1}[j\notin X]}$. Let $\Pi_X^{-i}(\omega) = \prod_{j\in N\setminus i} x_j(\omega)^{\mathbb{1}[j\in X]}(1 - x_j(\omega))^{\mathbb{1}[j\notin X]}$, and note that $\Pi_X^{-i}(\omega)$ depends only on ω_{-i}. We can then rewrite d as

$$d(\omega) = \sum_{X\subseteq N} \Pi_X^{-i}(\omega_{-i}) D(X \setminus \{i\}) + \sum_{X\subseteq N\setminus i} x_i(\omega_i)\Pi_X^{-i}(\omega_{-i}) \left[D(X \cup \{i\}) - D(X)\right].$$

For convenience, assume $\omega_i^{ND} \notin \Omega_i$, so that i's disclosure recommendation strategy is fully described by $x_i(\omega_i) = 0$ if $\omega_i < \omega_i^{ND}$ and $x_i(\omega_i) = 1$ if $\omega_i > \omega_i^{ND}$. (The proof is analogous if $\omega_i^{ND} \in \Omega_i$.) Using this, we can rewrite (10) as

$$\mathbb{E}\left(\omega_i | ND; e_{N\setminus i}\right) = \frac{\sum_{\Omega_{N\setminus i}} \left[\alpha(\omega_{-i}) \sum_{\Omega_i} \omega_i \hat{\nu}_{\omega_{-i}}(\omega_i) + \beta(\omega_{-i}) \sum_{\omega_i < \omega_i^{ND}} \omega_i \hat{\nu}_{\omega_{-i}}(\omega_i)\right] \eta(\omega_{-i})}{\sum_{\Omega_{N\setminus i}} \left[\alpha(\omega_{-i}) \sum_{\Omega_i} \hat{\nu}_{\omega_{-i}}(\omega_i) + \beta(\omega_{-i}) \sum_{\omega_i < \omega_i^{ND}} \hat{\nu}_{\omega_{-i}}(\omega_i)\right] \eta(\omega_{-i})},$$

Let $\hat{\mathcal{V}}_{\omega_{-i}}$ be the cdf implied by $\hat{\nu}_{\omega_{-i}}$, so that for each $\omega_i \in \Omega_i$, $\hat{\mathcal{V}}_{\omega_{-i}}(\omega_i) = \sum_{\omega_i \leqslant \omega_i} \hat{\nu}_{\omega_{-i}}(\omega_i)$. Also index the possible outcomes for team member i with $k \in \{1, ..., |\Omega_i|\}$, so that ω_i^k is the k^{th} lowest value in Ω_i; and take $K \in \{1, ..., |\Omega_i|\}$ so that ω_i^K is the largest value in Ω_i satisfying $\omega_i^K < \omega_i^{ND}$. Finally, to facilitate notation, let $\omega_i^0 = 0$. Use this towards the following rewriting:

$$\mathbb{E}\left(\omega_i | ND; e_{N\setminus i}\right) = \tag{12}$$

$$= \frac{\sum_{\Omega_{N\setminus i}} \alpha(\omega_{-i}) \sum_{j=1}^{|\Omega_i|} \left(1 - \hat{\mathcal{V}}_{\omega_{-i}}(\omega_i^{j-1})\right) \left(\omega_i^j - \omega_i^{j-1}\right) \eta(\omega_{-i})}{\sum_{\Omega_{N\setminus i}} \left[\alpha(\omega_{-i}) + \beta(\omega_{-i}) \hat{\mathcal{V}}_{\omega_{-i}}(\omega_i^K)\right] \eta(\omega_{-i})}$$

$$+ \frac{\sum_{\Omega_{N\setminus i}} \beta(\omega_{-i}) \sum_{j=1}^{K} \left(\hat{\mathcal{V}}_{\omega_{-i}}(\omega_i^K) - \hat{\mathcal{V}}_{\omega_{-i}}(\omega_i^{j-1})\right) \left(\omega_i^j - \omega_i^{j-1}\right) \eta(\omega_{-i})}{\sum_{\Omega_{N\setminus i}} \left[\alpha(\omega_{-i}) + \beta(\omega_{-i}) \hat{\mathcal{V}}_{\omega_{-i}}(\omega_i^K)\right] \eta(\omega_{-i})}.$$

Suppose by contradiction that $\mathbb{E}\left(\omega_i | ND; e_{N\setminus i}\right) \geqslant \mathbb{E}\left(\omega_i | ND; e_N\right) = \omega_i^{ND}$. In that case, we can show that the expression in (12) is decreasing in $\hat{\mathcal{V}}_{\omega_{-i}}(\omega_i^K)$.[31] Because effort is purely self-improving, we have $\hat{\mathcal{V}}_{\omega_{-i}}(\omega_i^K) \geqslant \mathcal{V}_{\omega_{-i}}(\omega_i^K)$, which implies the first inequality in

$$\mathbb{E}\left(\omega_i | ND; e_{N\setminus i}\right) \leqslant \frac{\sum_{\Omega_{N\setminus i}} \alpha(\omega_{-i}) \sum_{j=1}^{|\Omega_i|} \left(1 - \hat{\mathcal{V}}_{\omega_{-i}}(\omega_i^{j-1})\right) \left(\omega_i^j - \omega_i^{j-1}\right) \eta(\omega_{-i})}{\sum_{\Omega_{N\setminus i}} \left[\alpha(\omega_{-i}) + \beta(\omega_{-i}) \mathcal{V}_{\omega_{-i}}(\omega_i^K)\right] \eta(\omega_{-i})}$$

$$+ \frac{\sum_{\Omega_{N\setminus i}} \beta(\omega_{-i}) \sum_{j=1}^{K} \left(\mathcal{V}_{\omega_{-i}}(\omega_i^K) - \hat{\mathcal{V}}_{\omega_{-i}}(\omega_i^{j-1})\right) \left(\omega_i^j - \omega_i^{j-1}\right) \eta(\omega_{-i})}{\sum_{\Omega_{N\setminus i}} \left[\alpha(\omega_{-i}) + \beta(\omega_{-i}) \mathcal{V}_{\omega_{-i}}(\omega_i^K)\right] \eta(\omega_{-i})}$$

$$\Rightarrow 1 - d(\omega) = \left[1 - \sum_{X \subseteq N} \Pi_X^{-i}(\omega_{-i}) D(X)\right] + (1 - x_i(\omega_i)) \left[\sum_{X \subseteq N\setminus i} \Pi_X^{-i}(\omega_{-i}) \left[D(X \cup \{i\}) - D(X)\right]\right],$$

so that $\alpha(\omega_{-i})$ is the first term in brackets and $\beta(\omega_{-i})$ is the second term in brackets.

[31] To see this, note that the derivative of the expression in (12) with respect to $\hat{\mathcal{V}}_{\omega_{-i}}(\omega_i^K)$ is proportional to $\left[\omega_i^K - \mathbb{E}\left(\omega_i | ND; e_{N\setminus i}\right)\right] \leqslant \left[\omega_i^K - \omega_i^{ND}\right] < 0$, where the last equality holds by the definition of ω_i^K.

$$< \frac{\sum_{\Omega_{N\setminus i}} \alpha(\omega_{-i}) \sum_{j=1}^{|\Omega_i|} \left(1 - \mathcal{V}_{\omega_{-i}}(\omega_i^{j-1})\right) \left(\omega_i^j - \omega_i^{j-1}\right) \eta(\omega_{-i})}{\sum_{\Omega_{N\setminus i}} \left[\alpha(\omega_{-i}) + \beta(\omega_{-i})\mathcal{V}_{\omega_{-i}}(\omega_i^K)\right] \eta(\omega_{-i})}$$

$$+ \frac{\sum_{\Omega_{N\setminus i}} \beta(\omega_{-i}) \sum_{j=1}^{K} \left(\mathcal{V}_{\omega_{-i}}(\omega_i^K) - \mathcal{V}_{\omega_{-i}}(\omega_i^{j-1})\right) \left(\omega_i^j - \omega_i^{j-1}\right) \eta(\omega_{-i})}{\sum_{\Omega_{N\setminus i}} \left[\alpha(\omega_{-i}) + \beta(\omega_{-i})\mathcal{V}_{\omega_{-i}}(\omega_i^K)\right] \eta(\omega_{-i})}$$

$$= \mathbb{E}\left(\omega_i | ND; e_N\right). \tag{13}$$

The second inequality follows from the fact that $\hat{\mathcal{V}}_{\omega_{-i}}(\omega_i^j) \geqslant \mathcal{V}_{\omega_{-i}}(\omega_i^j)$ for each $j \in \{1, ..., |\Omega_i|\}$ (strictly so for some such j), and the fact that $\alpha(\omega_{-i}) > 0$ for some $\omega_{-i} \in \Omega_{N\setminus i}$. The statement in (13) then contradicts the assumption that $\mathbb{E}\left(\omega_i|ND; e_{N\setminus i}\right) \geqslant \mathbb{E}\left(\omega_i|ND; e_N\right)$; thereby proving the second statement in the lemma. □

Using Lemma 7 we therefore know that for any full-effort equilibrium disclosure rule d with $d(\omega) < 1$ for some $\omega \in \Omega$, the second term in the left-hand side of equation (4) — in Lemma 3 — is negative. And therefore, by Lemma 3, we know that if c is such that full effort is implementable by disclosure strategy d, then full effort can also be implemented by the full-disclosure strategy. And consequently $FE(D) \subset FE(D')$, where D' is the unilateral disclosure protocol and D is any other deliberation procedure. □

A.8.2 Proof of Statement 2

First note that for any team disclosure rule d and any effort vector e, we can write

$$\mathbb{E}\left(\omega_i|ND; e\right) = \frac{\sum_{\Omega_i} \omega_i(1 - d_i(\omega_i))\mu_i(\omega_i; e)}{\sum_{\Omega_i}(1 - d_i(\omega_i))\mu_i(\omega_i; e)}, \tag{14}$$

where $\mu_i(\cdot; e)$ is the marginal distribution of ω_i given e, and

$$d_i(\omega_i) = \sum_{w \in \Omega_{N\setminus i}} d(\omega_i, w)\mu_{N\setminus i}(w|\omega_i; e)$$

is the overall probability that a ω_i would be disclosed (integrating over all the possible outcome realizations for other team members, given that ω_i happens).

Now consider the consensus disclosure deliberation procedure. By Theorem 1, there exists an equilibrium in which $\omega_i^{ND} > \min(\Omega_i)$ for every $i \in N$. For ease of exposition, assume that equilibrium is strict, so that for each $i \in N$, $\omega_i^{ND} \notin \Omega_i$. We can thus express i's individual disclosure strategy without loss as $x_i(\omega) = 1$ if $\omega_i > \omega_i^{ND}$ and $x_i(\omega) = 0$ otherwise. The proof works analogously if $\omega_i^{ND} \in \Omega_i$ for any $i \in N$.

Because the disclosure must be chosen by consensus, we have that for each $i \in N$ and a given effort vector e,

$$d_i(\omega_i; e) = \begin{cases} 0, \text{ if } \omega_i \leqslant \omega_i^{ND} \\ \mathbb{P}\left(\omega_j > \omega_j^{ND} \text{ for all } j \neq i | \omega_i; e\right), \text{ if } \omega_i > \omega_i^{ND}. \end{cases}$$

And because effort is purely team improving, for all $i \in N$,

$$d_i(\omega_i; e_N) = d_i(\omega_i; e_{N\setminus i}) = 0, \text{ if } \omega_i \leqslant \omega_i^{ND} \text{ and } d_i(\omega_i; e_N) > d_i(\omega_i; e_{N\setminus i}), \text{ if } \omega_i > \omega_i^{ND},$$

where this inequality is due to $\omega_j > \omega_j^{ND}$ for all j being an upper set of $\Omega_{N\setminus i}$. Combining this with the expression in (14), and dropping the dependence of $\mu_i(\cdot; e)$ on effort (since team improving effort does not affect an agent's own marginal outcome distribution), we thus have

$$\mathbb{E}\left(\omega_i | ND; e_{N\setminus i}\right) = \frac{\sum_{\Omega_i} \omega_i (1 - d_i(\omega_i; e_{N\setminus i}))\mu_i(\omega_i)}{\sum_{\Omega_i} (1 - d_i(\omega_i; e_{N\setminus i}))\mu_i(\omega_i)}$$

$$= \frac{\sum_{\omega_i \leqslant \omega_i^{ND}} \omega_i \mu_i(\omega_i) + \sum_{\omega_i > \omega_i^{ND}} \omega_i (1 - d_i(\omega_i; e_{N\setminus i}))\mu_i(\omega_i)}{\sum_{\omega_i \leqslant \omega_i^{ND}} \mu_i(\omega_i) + \sum_{\omega_i > \omega_i^{ND}} (1 - d_i(\omega_i; e_{N\setminus i}))\mu_i(\omega_i)}$$

$$> \frac{\sum_{\omega_i \leqslant \omega_i^{ND}} \omega_i \mu_i(\omega_i) + \sum_{\omega_i > \omega_i^{ND}} \omega_i (1 - d_i(\omega_i; e_N))\mu_i(\omega_i)}{\sum_{\omega_i \leqslant \omega_i^{ND}} \mu_i(\omega_i) + \sum_{\omega_i > \omega_i^{ND}} (1 - d_i(\omega_i; e_N))\mu_i(\omega_i)} = \omega_i^{ND} = \mathbb{E}\left(\omega_i | ND; e_N\right).$$

And therefore in any full-effort equilibrium, for each $i \in N$, the second term on the left-hand side of (4) is strictly positive. And consequently the full-effort equilibrium set under the consensus disclosure protocol (let's denote it D) is non-empty, so $FE(D) \neq \varnothing$.

In contrast, for any deliberation procedure D' in which some team member can unilaterally choose disclosure, it must be that $FE(D') = \varnothing$. To see, let i be a team member who can unilaterally choose disclosure. By Theorem 1, in any equilibrium we must have $d(\omega) = 1$ for all ω with $\omega_i > \min(\Omega)$; and therefore in any such equilibrium, $\mathbb{E}(\omega_i | ND, e_N) = \mathbb{E}(\omega_i | ND, e_{N\setminus i})$. Moreover, because effort is purely team improving $\mathbb{E}(\omega_i | e_N) = \mathbb{E}(\omega_i | e_{N\setminus i})$. And so there are no direct or indirect benefits to i from exerting effort. Because effort is costly, there cannot be an equilibrium in which i exerts effort. Consequently, for any $c \in \mathbb{R}^n_{++}$, $FE(D') = \varnothing$.

And so we trivially have $FE(D') \subset FE(D)$, which concludes the proof of the statement. □

A.9 Proof of Proposition 3

A.9.1 Proof of Statement 1

Let $\{\hat{\mu}(\cdot;e)\}$ and $\{\mu(\cdot;e)\}$ be two effort environments satisfying $\hat{\mu}(\cdot;e_N) = \mu(\cdot;e_N)$, and suppose $\{\hat{\mu}(\cdot;e)\}$ is such that effort is purely self-improving. For each $\alpha \in [0,1]$, let $\{\mu^\alpha(\cdot;e)\}$ be an effort environment such that for each $e \in \{0,1\}^n$,

$$\mu^\alpha(\cdot;e) = \alpha\hat{\mu}(\cdot;e) + (1-\alpha)\mu(\cdot;e),$$

so that α measures the degree to which effort is self-improving.

Suppose D is a deliberation procedure in which some team member cannot unilaterally choose disclosure, and suppose d is an equilibrium team disclosure strategy given outcome distribution $\mu(\cdot : e_N)$, where d differs from full disclosure. Because $\hat{\mu}(\cdot;e_N) = \mu(\cdot : e_N)$, we know that d is an equilibrium disclosure strategy given deliberation procedure D and outcome distribution $\mu^\alpha(\cdot;e_N)$ for all $\alpha \in [0,1]$. Given d, $\mu^\alpha(\cdot;e_N)$, and $\mu^\alpha(\cdot;e_{N\setminus i})$, we can calculate

$$\begin{aligned}
\mathbb{E}^\alpha(\omega_i|ND;e_N) - \mathbb{E}^\alpha(\omega_i|ND;e_{N\setminus i}) = \mathbb{E}_\mu(\omega_i|ND;e_N) \\
- \frac{\alpha\hat{\mu}(ND;e_{N\setminus i})}{\alpha\hat{\mu}(ND;e_{N\setminus i}) + (1-\alpha)\alpha\mu(ND;e_{N\setminus i})}\mathbb{E}_{\hat{\mu}}(\omega_i|ND;e_{N\setminus i}) \\
- \frac{(1-\alpha)\mu(ND;e_{N\setminus i})}{\alpha\hat{\mu}(ND;e_{N\setminus i}) + (1-\alpha)\alpha\mu(ND;e_{N\setminus i})}\mathbb{E}_\mu(\omega_i|ND;e_{N\setminus i})
\end{aligned}$$

$$\begin{aligned}
&= \frac{\alpha\hat{\mu}(ND;e_{N\setminus i})}{\alpha\hat{\mu}(ND;e_{N\setminus i}) + (1-\alpha)\alpha\mu(ND;e_{N\setminus i})}\left[\mathbb{E}_{\hat{\mu}}(\omega_i|ND;e_N) - \mathbb{E}_{\hat{\mu}}(\omega_i|ND;e_{N\setminus i})\right] \\
&+ \frac{(1-\alpha)\mu(ND;e_{N\setminus i})}{\alpha\hat{\mu}(ND;e_{N\setminus i}) + (1-\alpha)\alpha\mu(ND;e_{N\setminus i})}\left[\mathbb{E}_\mu(\omega_i|ND;e_N) - \mathbb{E}_\mu(\omega_i|ND;e_{N\setminus i})\right],
\end{aligned} \tag{15}$$

where each step used the fact that $\mu^\alpha(\cdot;e_N) = \hat{\mu}(\cdot;e_N) = \mu(\cdot : e_N)$. Now consider two cases. First suppose $\mathbb{E}_\mu(\omega_i|ND;e_N) \leqslant \mathbb{E}_\mu(\omega_i|ND;e_{N\setminus i})$. We know from the proof of Proposition 2 that, because effort environment $\hat{\mu}$ is purely self-improving, $\mathbb{E}_{\hat{\mu}}(\omega_i|ND;e_N) < \mathbb{E}_{\hat{\mu}}(\omega_i|ND;e_{N\setminus i})$, and therefore (15) is weakly negative for all $\alpha \geqslant \alpha_i^d = 0$. Suppose instead that $\mathbb{E}_\mu(\omega_i|ND;e_N) \geqslant \mathbb{E}_\mu(\omega_i|ND;e_{N\setminus i})$. Then, because $\mathbb{E}_{\hat{\mu}}(\omega_i|ND;e_N) < \mathbb{E}_{\hat{\mu}}(\omega_i|ND;e_{N\setminus i})$, there exists some $\alpha_i^d \in [0,1)$ such that (15) is weakly negative for all $\alpha \geqslant \alpha_i^d$.

Now let $\bar{\alpha}(D) = \sup_{d\in\mathcal{D}} \max_{i\in N} \alpha_i$, where $\mathcal{D}$ is the set of equilibrium disclosure strategies for procedure D. Then, by Lemma 3, we know that the unilateral disclosure protocol dominates procedure D if and only if $\alpha \geqslant \bar{\alpha}(D)$. Further, let $\bar{\alpha}$ be the maximum $\bar{\alpha}(D)$ over all deliberation procedures in which some team member cannot unilaterally choose disclosure. Then the unilateral disclosure protocol dominates every other procedure if and only if $\alpha \geqslant \bar{\alpha}$.

A.9.2 Proof of Statement 2

The proof of statement 2 is analogous to that of statement 1. □

A.10 Proof of Proposition 4

Step 1. Fix $D' \neq D$, where D is the unilateral disclosure protocol. Remember that the outcome distribution when all team members exert effort is given by $\mu(\cdot; e_N) = (1-\epsilon)\mu + \epsilon\nu$. As our first step in the proof, we observe (in Lemma 8) that if ϵ is sufficiently large, an equilibrium of the team disclosure stage exists in which every team member favors disclosure if and only if they do not draw their worst outcome.

Lemma 8. *Suppose the outcome distribution is given by $\mu(\cdot; e_N) = (1-\epsilon)\mu + \epsilon\nu$, as stated in Proposition 4. There exists some $\epsilon' \in (0,1)$ such that if $\epsilon > \epsilon'$, there exists an equilibrium of the team disclosure stage — given full effort and deliberation procedure D' — where for every $i \in N$,*

$$x_i(\omega) = \begin{cases} 0, \text{ if } \omega_i = \min(\Omega_i) \\ 1, \text{ otherwise.} \end{cases} \tag{16}$$

Proof of Lemma. Conjecture an equilibrium of the team disclosure stage in which individual disclosure strategies are as given in (16); and suppose the implied equilibrium team disclosure strategy is $d(\omega) = D(x(\omega))$. Then we have for each $i \in N$, and each $\epsilon \in (0,1)$,

$$\omega_i^{ND,\epsilon} = \mathbb{E}^\epsilon\left(\omega_i | ND; e_N\right) = \mathbb{P}^\epsilon(\omega_i = \min(\Omega_i) | ND; e_N) \min(\Omega_i)$$

$$+ \mathbb{P}^\epsilon(\omega_i \neq \min(\Omega_i) | ND; e_N) \mathbb{E}^\epsilon\left(\omega_i | ND, \omega_i \neq \min(\Omega_i); e_N\right). \tag{17}$$

Note that, given the individual disclosure strategies in (16), no-disclosure happens only if at least one other team member $j \in N$ draws their worst possible outcome $\underline{\omega_j}$. But as $\epsilon \to 1$, it must be that for any $i, j \in N$, $\mathbb{P}(\omega_i = \min(\Omega_i) | \omega_j = \underline{\omega_j}) \to 1$. This, along with (17) and the fact that $\mathbb{E}^\epsilon\left(\omega_i | ND, \omega_i \neq \min(\Omega_i)\right)$ is bounded implies that for every $i \in N$,

$$\lim_{\epsilon \to 1} \omega_i^{ND,\epsilon} = \min(\Omega_i). \tag{18}$$

And consequently there is some ϵ' such that $\epsilon > \epsilon'$ implies that for every $i \in N$, $\omega_i^{ND,\epsilon} < \omega_i$ for all $\omega_i \in \Omega_i \setminus \{\min(\Omega_i)\}$. And therefore the individual disclosure strategy in (16) is individually rational and can be supported as an equilibrium of the team disclosure stage. □

Step 2. For $\epsilon > \epsilon'$ as given in Lemma 8, in the team disclosure equilibrium described in the lemma we have for some $i \in N$,

$$\mathbb{E}\left(\omega_i | ND; e_{N\setminus i}\right) > \min(\Omega_i).$$

And moreover, this value is independent of ϵ. These statements are true because (i) $\mu(\cdot; e_{N\setminus i})$ has full support over Ω and is independent of ϵ for every $i \in N$; and (ii) D' is not the unilateral disclosure deliberation procedure, and therefore given the individual disclosure strategies in (16) and D', for all $i \in N$ there exists some $\omega \in \Omega$ with $\omega_i \neq \min(\Omega_i)$ such that $d(\omega) < 1$.

Step 3. Fix $\epsilon > \epsilon'$ as given in Lemma 8 and consider the team disclosure equilibrium described in the lemma. By equation (18), and Step 2, we know that there is some $\bar{\epsilon} > \epsilon'$ such that, if $\epsilon > \bar{\epsilon}$,

$$\mathbb{E}\left[\omega_i | ND; e_{N\setminus i}\right] > \mathbb{E}\left[\omega_i | ND; e_N\right]$$

for all $i \in N$.

Step 4. As a consequence of Step 3, and using Lemma 3, we know that if $\epsilon > \bar{\epsilon}$, $fe(d') \subset fe(d)$ — where d' is the full disclosure rule and d is the equilibrium disclosure rule described in Lemma 8. Consequently, if $\epsilon > \bar{\epsilon}$,

$$FE(D) \subset FE(D'),$$

and so the unilateral disclosure protocol is strictly dominated by D'. □

A.11 Proof of Proposition 5

The first two statements in the proposition are implied directly by Theorem 1. For the third statement, suppose Assumption 4 holds and $D(\{i\}) \neq 1$. Conjecture an equilibrium in which, for each i,

$$x_i(\omega) = \begin{cases} 1, \text{ if } \omega_i = \omega^h, \\ 0, \text{ if } \omega_i = \omega^\ell. \end{cases} \tag{19}$$

Because D is symmetric, these conjectured strategies imply no-disclosure beliefs such that $\omega_i^{ND} = \omega_j^{ND}$ for all $i, j \in N$. Further, because $D(\{i\}) \neq 1$, we have that

$$\mathbb{P}(ND|\omega_i = \omega^h) \geqslant (1 - D(\{i\}))\mathbb{P}\left[\omega_j = \omega^\ell \ \forall j \in N \text{ with } j \neq i|\omega_i = \omega^h\right] > 0.$$

$$\text{and } \mathbb{P}(ND|\omega_i = \omega^\ell) \geqslant \mathbb{P}\left[\omega_j = \omega^\ell \ \forall j \in N \text{ with } j \neq i|\omega_i = \omega^\ell\right] > 0.$$

And so $\omega^h > \omega_i^{ND} > \omega^\ell$. Consequently, the conjectured recommendation strategy in (19) is as-if-pivotal optimal for each agent $i \in N$; thereby constituting an equilibrium.

Now suppose that a second equilibrium without full disclosure exists in which, for each $i \in N$, x_i depends only on ω_i. It must be that for some $i \in N$,

$$\mathbb{P}(ND|\omega_i = \omega^h) > 0 \text{ and } \mathbb{P}(ND|\omega_i = \omega^\ell) > 0,$$

and thus $\omega_i^{ND} \in (\omega^\ell, \omega^h)$. Therefore, in that equilibrium, i's recommendation strategy must satisfy (19). Because $\mathbb{P}(ND|\omega_i = \omega^h) > 0$, there must be a set $I \subset N$ of team members, with $i \notin I$ and $D(N \setminus I) < 1$, such that $\mathbb{P}(x_j(\omega) < 1|\omega_i = \omega^h) > 0$ for each $j \in I$.

Now fix some team member $k \in N$, with $k \neq i$; and let $K = (I \setminus \{k\}) \cup \{i\}$. By our initial assumption, we know that x_j depends only on ω_j for each $j \in K$. And therefore, for each $j \in K$ with $j \neq i$, $\mathbb{P}(x_j(\omega) < 1|\omega_k = \omega^h) = \mathbb{P}(x_j(\omega) < 1|\omega_i = \omega^h) > 0$. And, because x_i is as in (19), we know that $\mathbb{P}(x_i(\omega) < 1|\omega_k = \omega^h) > 0$. Moreover, because $|K| \geqslant |I|$, we have $D(N \setminus K) \leqslant D(N \setminus I) < 1$. And so we conclude that

$$\mathbb{P}(ND|\omega_k = \omega^h) > 0.$$

Similarly, we can conclude that $\mathbb{P}(ND|\omega_k = \omega^\ell) > 0$; and therefore $\omega_k^{ND} \in (\omega^\ell, \omega^h)$.

The same construction can be used to show that in the conjectured equilibrium $\omega_k^{ND} \in (\omega^\ell, \omega^h)$ for all $k \in N$. This implies that for every team member $i \in N$, recommendation strategies must be given by (19). Therefore the conjectured equilibrium must coincide with the equilibrium initially constructed in this proof, which is consequently the unique equilibrium without full disclosure in which, for each $i \in N$, x_i depends only on ω_i.

□

A.12 Proof of Lemma 4

Lemma 3 and Proposition 5 imply that the cost vector c belongs to the full effort set $iFE(D)$ if and only if $c_i \in (0, \bar{c}(D)]$ for each $i \in N$, where $\bar{c}(D)$ is defined as follows. If $D(\{i\}) = 1$ for every $i \in N$, then

$$\bar{c}(D) = \sum_{\Omega} \omega_i \mu(\omega; e_N) - \sum_{\Omega} \omega_i \mu(\omega; e_{N\setminus i}). \tag{20}$$

If instead $D(\{i\}) < 1$ for every $i \in N$, then

$$\bar{c}(D) = \sum_{\Omega} \omega_i \mu(\omega; e_N) - \sum_{\Omega} \omega_i \mu(\omega; e_{N\setminus i}) \tag{21}$$

$$+ \sum_{\Omega} (1 - d(\omega))\mu(\omega; e_{N\setminus i}) \left[\frac{\sum_{\Omega} \omega_i (1 - d(\omega))\mu(\omega; e_{N\setminus i})}{\sum_{\Omega} (1 - d(\omega))\mu(\omega; e_{N\setminus i})} - \frac{\sum_{\Omega} \omega_i (1 - d(\omega))\mu(\omega; e_N)}{\sum_{\Omega} (1 - d(\omega))\mu(\omega; e_N)} \right], \tag{22}$$

where d is the team disclosure strategy in the unique equilibrium without full disclosure described in Proposition 5. Note that, under the symmetry assumption, these expressions are independent of the particular choice of $i \in N$.

In the equilibrium without full disclosure described in Proposition 5, each individual's recommendation strategy is independent of the team's deliberation procedure D. Consequently, for each $\omega \in \Omega$, $d(\omega)$ is a continuous function of the the deliberation procedure D. Moreover, for any sequence $\{D^k\}$ of symmetric deliberation procedures with $D^k(\{i\}) \to 1$, it must be that $d^k(\omega) \to 1$ for every $\omega \in \Omega$. These two facts imply that $\bar{c}(D)$, as defined by (20) and (21), is a continuous function of D.

Additionally, in a symmetric deliberation procedure, $D(X)$ depends only on the cardinality of X, and by assumption $D(\varnothing)$ is fixed at 0 and $D(N) = 1$. Therefore, a symmetric deliberation procedure D is fully described by a vector in $[0, 1]^{n-1}$. Our assumption that deliberation procedures are monotone further requires that $D(X) \leqslant D(X')$ if $|X| \leqslant |X'|$. The space of deliberation procedures is thus a compact subset of $[0, 1]^{n-1}$.

Because $\bar{c}(D)$ is continuous, and the space of symmetric deliberation procedures is compact, there is a symmetric deliberation procedure D that maximizes $\bar{c}(D)$. As a consequence, a procedure D^* that maximizes $\bar{c}(D)$ is such that $iFE(D) \subseteq iFE(D^*)$ for every symmetric deliberation procedure D. And so D^* maximizes effort incentives among symmetric deliberation procedures.

□

A.13 Proof of Proposition 6

Because the team has two individuals, a symmetric deliberation procedure is fully described by the disclosure probability if exactly one team member recommends disclosure, $D(1)$.

From Lemmas 3 and 5, we know that full effort can be implemented in a symmetric equilibrium given a cost vector c and deliberation procedure D — $c \in SFE(D)$ — if and only if $c_i \in (0, \bar{c}(D)]$ for each $i \in N$, where $\bar{c}(D)$ is given by

$$\bar{c}(D) = \mathbb{E}\left(\omega_i | e_N\right) - \mathbb{E}\left(\omega_i | e_{N\setminus i}\right), \text{ if } D(1) = 1,$$

$$\text{and } \bar{c}(D) = \mathbb{E}\left(\omega_i | e_N\right) - \mathbb{E}\left(\omega_i | e_{N\setminus i}\right) + \mathbb{P}\left(ND | e_{N\setminus i}\right)\left[\mathbb{E}\left(\omega_i | ND; e_{N\setminus i}\right) - \mathbb{E}\left(\omega_i | ND; e_N\right)\right],$$

if $D(1) < 1$, where the disclosure/non-disclosure of each ω realization is given by the team disclosure strategy in the unique symmetric partial-disclosure equilibrium. The effort-maximizing procedure is therefore the one that maximizes the objective

$$\mathbb{P}\left(ND | e_{N\setminus i}\right)\left[\mathbb{E}\left(\omega_i | ND; e_{N\setminus i}\right) - \mathbb{E}\left(\omega_i | ND; e_N\right)\right]. \tag{23}$$

We can write the expression for each of the terms in this objective. We have:

$$\mathbb{P}\left(ND | e_{N\setminus i}\right) = \mu\left[(\omega_\ell, \omega_\ell); e_{N\setminus i}\right] + (1 - D(1))\left\{\mu\left[(\omega_h, \omega_\ell); e_{N\setminus i}\right] + \mu\left[(\omega_\ell, \omega_h); e_{N\setminus i}\right]\right\}.$$

$$\begin{aligned}
\mathbb{E}\left(\omega_1 | e_{N\setminus 1}\right) &= \omega_\ell + (\omega_h - \omega_\ell) \\
&\times \frac{(1 - D(1))\mu\left[(\omega_h, \omega_\ell); e_{N\setminus 1}\right]}{\mu\left[(\omega_\ell, \omega_\ell); e_{N\setminus 1}\right] + (1 - D(1))\left\{\mu\left[(\omega_h, \omega_\ell); e_{N\setminus 1}\right] + \mu\left[(\omega_\ell, \omega_h); e_{N\setminus 1}\right]\right\}} \\
&= \omega_\ell + (\omega_h - \omega_\ell)\frac{\mu\left[(\omega_h, \omega_\ell); e_{N\setminus 1}\right]}{\mu\left[(\omega_h, \omega_\ell); e_{N\setminus 1}\right] + \mu\left[(\omega_\ell, \omega_h); e_{N\setminus 1}\right]} \\
&\times \frac{(1 - D(1))\left\{\mu\left[(\omega_h, \omega_\ell); e_{N\setminus 1}\right] + \mu\left[(\omega_\ell, \omega_h); e_{N\setminus 1}\right]\right\}}{\mu\left[(\omega_\ell, \omega_\ell); e_{N\setminus 1}\right] + (1 - D(1))\left\{\mu\left[(\omega_h, \omega_\ell); e_{N\setminus 1}\right] + \mu\left[(\omega_\ell, \omega_h); e_{N\setminus 1}\right]\right\}} \\
&= \omega_\ell + (\omega_h - \omega_\ell)\frac{\mu\left[(\omega_h, \omega_\ell); e_{N\setminus 1}\right]}{\mu\left[(\omega_h, \omega_\ell); e_{N\setminus 1}\right] + \mu\left[(\omega_\ell, \omega_h); e_{N\setminus 1}\right]} \frac{1 - D(1)}{\frac{\mu\left[(\omega_\ell, \omega_\ell); e_{N\setminus 1}\right]}{\mu\left[(\omega_h, \omega_\ell); e_{N\setminus 1}\right] + \mu\left[(\omega_\ell, \omega_h); e_N\right]} + 1 - D(1)} \\
&= \omega_\ell + (\omega_h - \omega_\ell)\sigma\frac{1 - D(1)}{\rho + 1 - D(1)}.
\end{aligned}$$

Using analogous steps, we have

$$\mathbb{E}\left(\omega_1|e_N\right) = \omega_\ell + (\omega_h - \omega_\ell)\bar{\sigma}\frac{1-D(1)}{\bar{\rho}+1-D(1)}.$$

And therefore the objective in (23) can be rewritten as

$$\begin{aligned}&\left[\mu\left[(\omega_\ell,\omega_\ell);e_{N\setminus i}\right] + (1-D(1))\left\{\mu\left[(\omega_h,\omega_\ell);e_{N\setminus i}\right] + \mu\left[(\omega_\ell,\omega_h);e_{N\setminus i}\right]\right\}\right]\\ &\qquad\times\left[\omega_\ell + (\omega_h-\omega_\ell)\sigma\frac{1-D(1)}{\rho+1-D(1)} - \omega_\ell - (\omega_h-\omega_\ell)\bar{\sigma}\frac{1-D(1)}{\bar{\rho}+1-D(1)}\right] =\end{aligned}$$

$$\left[\mu\left[(\omega_\ell,\omega_\ell);e_{N\setminus i}\right] + (1-D(1))\left\{\mu\left[(\omega_h,\omega_\ell);e_{N\setminus i}\right] + \mu\left[(\omega_\ell,\omega_h);e_{N\setminus i}\right]\right\}\right]$$

$$(\omega_h-\omega_\ell)\left[\sigma\frac{1-D(1)}{\rho+1-D(1)} - \bar{\sigma}\frac{1-D(1)}{\bar{\rho}+1-D(1)}\right],$$

which is proportional to

$$\begin{aligned}&\left[\frac{\mu\left[(\omega_\ell,\omega_\ell);e_{N\setminus i}\right]}{\mu\left[(\omega_h,\omega_\ell);e_{N\setminus i}\right] + \mu\left[(\omega_\ell,\omega_h);e_{N\setminus i}\right]} + 1 - D(1)\right]\left[\sigma\frac{1-D(1)}{\rho+1-D(1)} - \bar{\sigma}\frac{1-D(1)}{\bar{\rho}+1-D(1)}\right],\\ &=(\rho+1-D(1))\left[\sigma\frac{1-D(1)}{\rho+1-D(1)} - \bar{\sigma}\frac{1-D(1)}{\bar{\rho}+1-D(1)}\right]\\ &=\left[\sigma - \frac{\rho+1-D(1)}{\bar{\rho}+1-D(1)}\bar{\sigma}\right](1-D(1)) \equiv \Psi(D(1)).\end{aligned} \tag{24}$$

We now want to maximize the objective in (24) with respect to $D(1)$. We consider two cases.

Case 1. $\rho > \bar{\rho}$. First, we verify that the objective is strictly convex for all $D(1) < 1$.

$$\Psi'(D(1)) = (1-D(1))\frac{\bar{\sigma}(\bar{\rho}-\rho)}{(\bar{\rho}+1-D(1))^2} - \left[\sigma - \bar{\sigma}\frac{\rho+1-D(1)}{\bar{\rho}+1-D(1)}\right] \tag{25}$$

$$\Psi''(D(1)) = -\frac{2\bar{\sigma}(\bar{\rho}-\rho)}{(\bar{\rho}+1-D(1))^2} + \frac{2(\bar{\rho}+1-D(1))(1-D(1))\bar{\sigma}(\bar{\rho}-\rho)}{(\bar{\rho}+1-D(1))^4} > 0$$

$$\Leftrightarrow \frac{1-D(1)}{\bar{\rho}+1-D(1)} < 1, \text{ which holds for all } D(1) < 1.$$

And so D^* that maximizes $\Psi(D(1))$ is either $D^* = 0$ or $D^* = 1$. And $D^* = 0$ if and only if $\Psi(0) \geqslant \Psi(1)$, or equivalently

$$\sigma \geqslant \bar{\sigma}\frac{\rho+1}{\bar{\rho}+1} \Leftrightarrow \frac{\sigma}{\bar{\sigma}} \geqslant \frac{\rho+1}{\bar{\rho}+1},$$

which yields statement 1 in the proposition.

Case 2. $\rho < \bar{\rho}$. By the same steps, we know that the objective is strictly concave for all $D(1) < 1$. And therefore

$$D^* = \begin{cases} 0, \text{ if } \Psi'(0) \leqslant 0, \\ D(1) \in (0,1), \text{ if } \Psi'(D(1)) = 0 \text{ for some } D(1) \in (0,1), \\ 1, \text{ if } \Psi'(1) \geqslant 0. \end{cases}$$

From equation (25), we have that

$$\Psi'(D(1)) = \left[\sigma - \bar{\sigma}\frac{\rho+1-D(1)}{\bar{\rho}+1-D(1)}\right] - (1-D(1))\frac{\bar{\sigma}(\bar{\rho}-\rho)}{(\bar{\rho}+1-D(1))^2},$$

which is increasing in σ, and therefore D^* is weakly decreasing in σ. We also have that

$$\frac{\partial \Psi'(D(1))}{\partial \rho} = -\frac{\bar{\sigma}}{\bar{\rho}+1-D(1)} + \frac{(1-D(1))\bar{\sigma}}{(\bar{\rho}+1-D(1))^2} \leqslant 0.$$

And so D^* is weakly increasing in ρ.

□

B Additional Results

B.1 Additional Results for Section 5.3

As in the proof of Proposition 6, we know that a deliberation protocol maximizes effort incentives if it maximizes the following objective:

$$\mathbb{P}(\omega_i|ND; e_{N\setminus i}) \left[\mathbb{E}\left(\omega_i|ND; e_{N\setminus i}\right) - \mathbb{E}\left(\omega_i|ND; e_N\right)\right],$$

which is proportional to

$$\mathbb{P}(\omega_i|ND;e_{N\setminus i})\left[\frac{Pr(\omega_i=1\cap ND;e_{N\setminus i})}{Pr(ND;e_{N\setminus i})}-\frac{Pr(\omega_i=1\cap ND;e_N)}{Pr(ND;e_N)}\right].$$

Using the binary structure and the given deterministic symmetric deliberation protocol, we can write expressions for each of these terms. If the protocol is such that disclosure occurs if at least K team members favor it, then no-disclosure occurs if and only if at least $N-K+1$ team members are against disclosure, i.e. obtain a bad outcome. This can occur if either all receive the same common bad outcome or if at least $N-K+1$ team members receive independently bad draws of their individual binary outcome. Using this additional structure, we write[32]

$$\mathbb{P}(\omega_i=1\cap ND;e_{N\setminus i})=(1-\rho)h_i\sum_{m=N-K+1}^{N-1}\binom{N-1}{m}(1-h_j)^m h_j^{N-1-m}.$$

$$\begin{aligned}\mathbb{P}(ND;e_{N\setminus i})=\rho(1-h_T)+(1-\rho)(1-h_i)\sum_{m=N-K}^{N-1}\binom{N-1}{m}(1-h_j)^m h_j^{N-1-m}\\+(1-\rho)h_i\sum_{m=N-K+1}^{N-1}\binom{N-1}{m}(1-h_j)^m h_j^{N-1-m}\end{aligned}$$

$$\begin{aligned}\Rightarrow\mathbb{P}(ND;e_{N\setminus i})=\rho(1-h_T)+(1-\rho)\sum_{m=N-K+1}^{N-1}\binom{N-1}{m}(1-h_j)^m h_j^{N-1-m}\\+(1-\rho)(1-h_i)\binom{N-1}{N-K}(1-h_j)^{N-K}h_j^{K-1}.\end{aligned}$$

And so

$$\begin{aligned}\mathbb{E}\left[\omega_i|ND;e_{N\setminus i}\right]=&\left[\frac{\rho(1-h_T)}{(1-\rho)h_i\sum_{m=N-K+1}^{N-1}\binom{N-1}{m}(1-h_j)^m h_j^{N-1-m}}+\frac{1}{h_i}\right.\\&\left.+\frac{(1-h_i)\binom{N-1}{N-K}}{h_i\sum_{m=N-K+1}^{N-1}\binom{N-1}{m}\left(\frac{1-h_j}{h_j}\right)^{m-(N-K)}}\right]^{-1}.\end{aligned}\tag{26}$$

[32]We adopt the convention that $\sum_{m=N}^{N-1}X(m)=0$ for any function X. This is relevant if $K=1$, in which case, following a good outcome for player i there is no possibility for no-disclosure.

And using the same steps, we have

$$\mathbb{E}\left[\omega_i|ND; e_N\right] = \Bigg[\frac{\bar{\rho}(1-h_T)}{(1-\bar{\rho})\bar{h}\sum_{m=N-K+1}^{N-1}\binom{N-1}{m}(1-\bar{h})^m\bar{h}^{N-1-m}} + \frac{1}{\bar{h}} + \frac{(1-\bar{h})\binom{N-1}{N-K}}{\bar{h}\sum_{m=N-K+1}^{N-1}\binom{N-1}{m}\left(\frac{1-\bar{h}}{\bar{h}}\right)^{m-(N-K)}}\Bigg]^{-1}. \tag{27}$$

By comparing the difference between (26) and (27), we can assess whether protocols with $K > 1$ provide more effort incentives than the unilateral protocol (with $K = 1$). Results are stated in Proposition 7 below.

Proposition 7. *The unilateral disclosure protocol ($K = 1$) is strictly dominated by all symmetric deterministic protocols with $K > 1$ if*

(i) Effort is purely team improving, that is, for each $i \in N$ and $j \neq i$,

$$\bar{h} > h_j, \bar{h} = h_i, \text{ and } \bar{\rho} = \rho.$$

(ii) Effort improves correlation between individual outcomes, that is, for every $i \in N$ and $j \neq i$,

$$\bar{\rho} > \rho, h_j = \bar{h}, \text{ and } h_i = \bar{h}.$$

The unilateral disclosure protocol ($K = 1$) dominates all K-majority protocols if

(iii) Effort is purely self-improving, that is, for each $i \in N$ and $j \neq i$,

$$\bar{h} > h_i, \bar{h} = h_j, \text{ and } \bar{\rho} = \rho.$$

The first two statements in the proposition are stronger versions of results in section 4. In this binary environment, if effort is team improving, then the unilateral disclosure protocol is dominated by *all* symmetric deliberation protocols such that disclosure requires more consensus. If effort improves the correlation between team members' outcomes — not necessarily to an extreme degree as in Proposition 4 — then all symmetric deliberation protocols dominate the unilateral disclosure protocol.

B.1.1 Proof of Proposition 7

In order to show the four statements, it suffices to sign the derivative of (26) with respect to the appropriate parameter. We begin with the first statement, so that we want to sign that derivative with respect to h_j. To do so, note that

$$\sum_{m=N-K+1}^{N-1} \binom{N-1}{m} (1-h_j)^m h_j^{N-1-m}$$

is decreasing in h_j, as it equals the probability of at least $N-K-1$ successes under a binomial with $N-1$ draws and success probability $1-h_j$. Moreover,

$$\sum_{m=N-K+1}^{N-1} \binom{N-1}{m} \left(\frac{1-h_j}{h_j}\right)^{m-(N-K)}$$

is also decreasing in h_j, as $m > N-K$ for the whole range of summation. Consequently, we have that $\mathbb{E}\left(\omega_i|ND\right)$ is decreasing in h_j, and therefore under the parametrization in statement (i), we have $\mathbb{E}\left(\omega_i|ND, e_{N\setminus i}\right) > \mathbb{E}\left(\omega_i|ND, e_N\right)$. This implies that all symmetric deliberation protocols with $K > 1$ strictly dominate the unilateral disclosure protocol (with $K = 1$).

Statements (ii)-(iv) follow from the same logic as statement (i), noting from equation (26) that $\mathbb{E}\left(\omega_i|ND\right)$ decreases in ρ, and increases in h_i.

□