Disclosure and Incentives in Teams*

Paula Onuchic   João Ramos
University of Oxford   USC Marshall

September, 2023

Abstract

We consider a team-production environment augmented with a stage in which the team decides how to communicate their 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 credit and blame for the team's successes and failures across team-members. Further, we show that, through this credit/blame-attribution channel, a team's deliberation procedure determines individuals' incentives to contribute to team production. We then characterize productive environments where effort-incentives are maximized by unilateral disclosure protocols or procedures such that disclosure require more consensus.

1 Introduction

Productive activities are increasingly conducted in teams. Startups are often founded by entrepreneurial partners, and in established firms new products are mostly developed and proposed by teams built and empowered within the company.[1] In policy-making or regulatory

*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, Michael Powell, Debraj Ray, Ludvig Sinander, Caroline Thomas, and Mark Whitmeyer. We also thank the audiences at seminars in several institutions, as well as the 2^{nd} Southeast Theory Festival at Nuffield College, 2023 ESSET - Gerzensee, 2023 SAET - Paris, and 2023 SITE - Dynamic Games, Contracts and Markets.

[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, Schwert 2021, and Jones (2021).

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 their 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; 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 disclosure is seen as an engineering success, but has negative reputational implications to the marketing team-members. If this team instead chooses to delay the product launch, then skeptical observers interpret this non-disclosure as a team failure, but blame for this negative collective outcome may be unevenly shared between the co-developers.

This paper proposes a model that incorporates such a disclosure stage into a team-production environment. In our model, the collective decision to disclose team-outcomes aggregates individual recommendations through some deliberation procedure. This procedure determines the allocation of *voice rights* across team-members — a voice right is defined by Zuckerman (2010) and Freeland and Zuckerman (2018) as "the right to speak on behalf of an organization." We show that the allocation of such rights impacts how team-outcomes are perceived by outside observers. Specifically, it determines how credit and blame for the team's successes and failures are shared by team-members. Through this channel, the allocation of voice rights within a team — at the disclosure stage — determines individuals' incentives to contribute to the team in the production stage. We leverage this insight to design team-deliberation protocols that incentivize individual effort provision.

Section 2 introduces the team-disclosure environment. A team is made up of two or more team-members who produce a team-outcome, drawn from a distribution known by all team-members and by an outside observer.[4] The team outcome is a piece of evidence which conveys

[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 (2022)).

[3]In the United States, at the start of congress, committees adopt and publish procedural rules which 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 determine 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."

[4]To study the team-disclosure problem in section 2, we take the team-outcome distribution as an exogenous

to the outside 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 member of the team, 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. Through a given *deliberation procedure* which aggregates individual disclosure recommendations, the team decides whether to reveal their outcome to the observer. If they do so, then each team member receives their respective individual value implied by the revealed evidence. If the outcome is not disclosed, then the observer forms a rational belief about each individual's value, accounting for the circumstances that may have led the team to choose not to reveal it. Individual payoffs are then equal to their respective *no-disclosure beliefs* held by the observer.

The deliberation procedure affects the team's disclosure decision directly by establishing whose individual recommendations are heeded by the team; but also indirectly, because it determines the equilibrium formation of the observer's no-disclosure beliefs. Take, for instance, a team with two team-members, and suppose each team-member can unilaterally decide that the team-outcome be disclosed. That is, the outcome is disclosed if at least one of the team-members recommends that decision. In any equilibrium, upon seeing that the team outcome is not disclosed, the observer must infer that the realized outcome was "bad news" to both team-members — for otherwise one of them would have chosen to disclose it — and must therefore form no-disclosure beliefs that are unfavorable to both team-members. This logic is typically referred to as *unravelling* in the single-agent evidence disclosure literature following Grossman (1981) and Milgrom (1981). It ensures that in equilibrium all outcomes are disclosed to the observer and that no-disclosure beliefs are "maximally skeptical" about every team-member.[5]

But suppose instead that the deliberation procedure is such that neither team-member can unilaterally disclose the team's outcome, that is, disclosure occurs only if neither team-member vetoes that decision. Then upon seeing no disclosure, the observer understands that the outcome must have been bad news about some team-member — for its disclosure was vetoed — but cannot definitively attribute the team's decision to an individual. A consequence is that the observer's equilibrium beliefs are not maximally skeptical about either team-member. Following this logic, Theorem 1 shows that in any equilibrium, the observer must be maximally skeptical about every team-member who has the ability to unilaterally disclose the team's out-

model primitive. At a later section, 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).

come (that is, who has complete voice rights). Conversely, there exists an equilibrium in which the observer is not maximally skeptical about any of the team-members who cannot unilaterally choose disclosure. The latter equilibrium must involve partial disclosure; specifically, the team does not disclose outcomes that are sufficiently bad news about sufficiently many team-members ("team failures").

The literature on single-agent evidence disclosure games has established a clear relation between model primitives, such as sender preferences and the evidence structure, and the equilibrium observer skepticism which triggers the unravelling of uninformative equilibria — see Hagenbagh, Koessler, and Perez-Richet (2014) and Rappoport (2023). In our model, individual preferences and the evidence structure are simple, so as to highlight the key new mechanism introduced by a disclosure model with a team, rather than a single individual, sender. In this context, equilibrium non-disclosures are met with observer-skepticism *about the team*. However, this aggregate skepticism is not sufficient to ensure unravelling; we highlight instead the importance of a new targeted notion of skepticism, which we call *blame*. We measure the blame attributed to an individual for a team-failure by the equilibrium no-disclosure belief that the observer assigns to that particular team-member. Proposition 1 and Corollaries 1 and 2 establish a relation between a team's deliberation procedure and the equilibrium assignment of blame across team-members. These results clarify that increasing an individual's voice rights (that is, their ability to speak on behalf of the team) also increases the observer's perception of their blame for team failures.

Theorem 2 completes our characterization of team-disclosure equilibria. Under any deliberation procedure, full-disclosure is one equilibrium in the equilibrium set; the theorem describes a refinement of the equilibrium set, which we use to evaluate the plausibility of the full-disclosure equilibrium. To that end, we introduce a refinement criterion bespoke to a team communication environment, which requires no-disclosure beliefs (even if off-path) to be *consistent with deliberation*. Intuitively, this criterion imposes the requirement that even off-path beliefs held by the outside observer should be justified by the aggregation of individual behavior of the team-members through the given deliberation protocol. The theorem shows that full-disclosure is consistent with deliberation if and only the procedure implies that disclosure requires less consensus than concealing evidence (this notion is formalized in the text).

Section 4 studies the full team-production and team-disclosure problem: e augment the environment in 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 positively affects not only the value of the team-outcome for themselves (what we call the individual's "own outcome"), 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 which now 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 that individual's expected own outcome with versus without their effort contribution, and a "blame misattribution" component. The blame misattribution component is the novel incentive mechanism introduced by strategic team-disclosure. In any equilibrium, the observer attributes blame for team failures (which are not disclosed) under the assumption that each individual exerted their respective equilibrium level of effort. If an individual deviates to another effort level, and the team draws a team failure, then blame for that failure will be misattributed across the team-members — each individual is either 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 extra production due to an individual's effort accrues mostly to that same individual — then the protocol that gives every individual the right to unilaterally disclose the team's outcome provides stronger effort incentives than any other deliberation protocol. On the other hand, if effort has high team-externalities — so 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.[6] Proposition 3 also shows that protocols that require more consensus for disclosure provide more incentives (relative to the unilateral disclosure protocol) for individuals to invest in a "common output component," which improves the correlation between all team-members' outcomes.

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 fully held 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," saying that "when used judiciously (...) blame

[6]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).

can prod people to put forth their best efforts." In contrast, the partial-disclosure equilibrium induced by consensus-disclosure protocols resembles a corporate culture where 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. Again 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 used 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.

In section 5, we consider binary-outcome environments, where every possible team-outcome can is seen by each team-member as either a "high outcome" or a "low outcome." In such an environment, we provide further characterization of team-disclosure and show stronger versions of the results in previous sections. In particular, we characterize "effort-maximizing deliberation" and show how the optimal degree of consensus required for disclosure varies with characteristics of the productive environment.

1.1 Related Literature

As already mentioned, our paper relates to the large literature on evidence disclosure, mainly stemming from Grossman (1981) and Milgrom (1981), and our paper is especially close to models with multidimensional evidence, such as Dziuda (2011) and Martini (2018).[7] We contribute to this literature by studying a problem of *team*-disclosure. Among other results, we show that team-disclosure equilibria often feature partial evidence revelation. The vast literature on single-agent disclosure games provides a variety of mechanisms that prevent "unravelling."[8] Our result is connected to Dye's (1985) observation that partial-disclosure equilibria exist in a single-agent problem when the observer is unsure whether the sender has access to

[7]Our equilibrium characterization is reminiscent of that in Martini's (2018) multi-dimensional disclosure model. Martini (2018) shows that if a single sender separably values the receiver's posterior about each dimensional 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 on which dimension 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 generally be mapped into instances of the latter through appropriately chosen sender preferences.

[8]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.

evidence. In the team context, despite senders always having access to evidence, they may be unable to disclose it because its disclosure is vetoed by other team-members. Our mechanism is also connected to that in Seidmann and Winter (1997) and Giovannoni and Seidmann (2006), which argue that partial-disclosure equilibria arise when, upon seeing no-disclosure, the is unsure whether the sender intended to bias their belief upwards or downwards. Similarly in the team context, the observer is unable to attribute the decision to not disclose to the interests of a particular individual in the team.

More broadly, our paper contributes to the literature on multi-sender communication. Using different communication protocols, Milgrom and Roberts (1986), Battaglini (2002), and Gentzkow and Kamenica (2016) study models where multiple senders communicate with a single receiver — Baumann and Dutta (2022), and Hu and Sobel (2019) are some more recent contributions that model multi-sender evidence disclosure. All those papers consider environments where senders "competitively" communicate with a receiver; that is, they unilaterally send 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 "cooperative" communication by a group using different deliberation procedures.

Our paper also connects to a small literature relating disclosure and incentives. Ben-Porath, Dekkel, 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 (2022), 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) — study 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 *evidence disclosure* by a team, rather than an environment where an action choice is delegated to a committee.[9] 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, follow-

[9] Bardhi and Bobkova (2023) also study an environment where the committee discloses evidence to a principal. However, in their model, all acquired evidence is necessarily disclosed to the outside observer.

ing Holmström (1999), and specifically to papers on career concerns in teams.[10] More generally, our results on incentive provision relate to the large literature on incentives in teams — 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 where blame for team-failures cannot be attributed across team-members.[11]

Finally, our study of blame for team failures connects to a small literature that considers credit attribution in teams. Onuchic ⓡ Ray (2023), Ray ⓡ Robson (2018), and Ozerturk and Yildirim (2021) study team production with unequal credit attribution to team-members. In the latter two papers, the attribution of credit is endogenously based on estimates of individual contributions. There are no reputational concerns, and credit attributed to each agent only determines their share in the physical outcome of the project.

2 Team Disclosure

2.1 Environment

A group $N = \{1, 2, ..., n\}$ of agents makes up a team, whose outcome $\omega = (\omega_1, ..., \omega_n)$ may be seen by an outside observer. The team's outcome is drawn from a distribution μ over a finite outcome space $\Omega \subset \mathbb{R}^n$. When the outcome ω realizes, the team has some piece of *hard evidence* that conveys to the observer the outcome that realized. For each $\omega \in \Omega$ and $i \in N$, ω_i should be interpreted as the value to team-member i of having the observer see outcome ω. This value may be derived from the observer making an assessment about i's ability, or by the observer learning about a state of the world which is relevant to the team-members; our general formulation allows for various interpretations of the individual implications of the team-product.

Each team member i, after seeing the team-product ω, makes an individual disclosure recommendation: $x_i(\omega) \in [0, 1]$ indicates the probability that agent i recommends the disclosure of outcome ω; with complementary probability $1 - x_i(\omega)$, agent i recommends that the outcome

[10] Jeon (1996) and Bar-Isaac (2007) show that pairing older and younger workers in teams may be beneficial for incentive provision. Ortega (2003) shows that the power allocation within a team — the distribution of how individual efforts affect team outcomes — affect effort incentives. Auriol, Friebel, and Pechlivanos (2002) and Chaliotti (2016) show that sabotage incentives arise when team-members are motivated by career concerns; and Arya and Mittendorf (2011) argue for the incentive benefits of aggregate performance measures in such environments.

[11] 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.

be concealed from the outside observer. Individual disclosure recommendation strategies define a distribution over the set of team-members who favor the disclosure of outcome ω. Formally, for every subset of team-members $X \subseteq N$

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

is the probability that the set of team-members who favor the disclosure of outcome ω is X. The teams' disclosure decision is then made according to some *deliberation process*. A deliberation process $D : \mathcal{P}(N) \to [0, 1]$ is procedure that aggregates individual disclosure decisions into a team disclosure decision. The team's disclosure decision $d(\omega) = \sum_{X \subseteq N} \Pi_X(\omega) D(X) \in [0, 1]$ represents the probability that the outcome ω is disclosed to the outside observer.

In a real-world scenario, deliberation is a perhaps lengthy process made up of formal rules and communication between team-members which somehow aggregates the interests of the group into a team decision. 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. In this model, we interpret our deliberation protocol D as a reduced form aggregation rule which 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 that decision is followed. And we require that the probability of disclosure be increasing in the set of people who favor the outcome's disclosure. Formally:

Assumption 1. *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 $D(X) \leqslant D(X')$.*

Given these assumptions, the deliberation procedure for a team with $N = \{1, 2\}$ 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 $D(\{2\}) \in [0, 1]$, the probability of team-disclosure when it is 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 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 where all team-members can unilaterally choose disclosure, so that $D(\{i\}) = 1$ for every $i \in N$. Conversely, we say disclosure decisions are made via *consensus* if the deliberation procedure is such that $D(\{i\}) = 0$ for every $i \in N$; this procedure mandates that the outcome be disclosed with any probability 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 highlights procedures where each of the team-members can unilaterally choose disclosure. In particular, we highlight the 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 team-outcome is 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 their mean-posterior about team-member i's outcome is equal to the realized ω_i, for each $i \in N$. If instead ω is not disclosed, then the observer's mean-posterior about i's outcome is 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 $\sum_\Omega (1 - d(\omega))\mu(\omega) > 0$. If no disclosure is an off-path (measure zero) event, then the observer's mean posterior is indeterminate. We refer to ω_i^{ND} as the observer's *no-disclosure belief* about team-member i. Agents' payoffs are equal to the observer's belief about their own outcomes, both when the team-outcome is disclosed and when it is not.[12]

2.2 Equilibrium

Definition 1 (Equilibrium)**.** *Given a deliberation procedure D, individual disclosure strategies x_i for $i \in N$, the team's disclosure decision d, and no-disclosure posteriors ω_i^{ND} for $i \in N$ constitute an equilibrium if*

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

$$\omega_i > \omega_i^{ND} \Rightarrow x_i(\omega) = 1 \textit{ and } \omega_i < \omega_i^{ND} \Rightarrow x_i(\omega) = 0.$$

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

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

[12] We assume that agents' payoffs depend on the posterior *mean* induced on the observer about their outcomes. However, we can also allow agents to value other moments of the outcome distribution, by renormalizing the outcomes. For example, if agents' payoffs are given by $\mathbb{E}(\omega_i^2)$, they can be equivalently expressed by $\mathbb{E}(\nu_i)$, where $\nu_i = \omega_i^2$. In that case, we would take μ to be the joint distribution of $(\nu_1, ..., \nu_N)$.

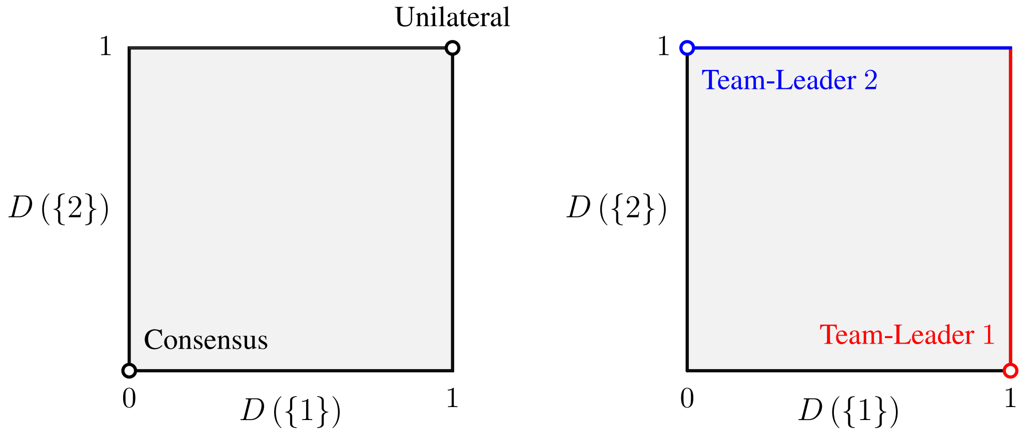


Figure 1: The set of deliberation procedures for two-person teams. The left-hand panel highlights the *unilateral* procedure, in which $D(\{1\}) = D(\{2\} = 1$ and the *consensus* procedure, in which $D(\{1\}) = D(\{2\} = 1$. The right-hand panel highlights in red the procedures such that team-member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

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

The equilibrium notion described above is Perfect Bayesian Equilibrium, in which all team members and the outside observer understand the deliberation process, and with the additional restriction that individuals make disclosure recommendations as if their recommendation is pivotal to the team's decision. With this requirement, we refine out equilibria where individuals position themselves for/against disclosure solely because they believe themselves not to bepivotal, and the equilibrium strategies indeed support that their recommendations are not pivotal. Condition 2 states that the teams' equilibrium disclosure strategy is reached by aggregating the individual equilibrium disclosure strategies, according to the given deliberation process. Finally, condition 3 imposes Bayes-consistency for beliefs reached on the equilibrium path.

We know from previous literature on disclosure with verifiable information — for a survey, see Milgrom (2008) — that when disclosure decisions are made by a single individual, the unique equilibrium involves full disclosure of all outcomes. The key insight supporting that result is that if the observer knows that an individual holds some hard evidence of their outcome, then the non-disclosure of that evidence makes the observer skeptical about the outcome realization. The observer's skepticism then generates an *unraveling* of any equilibrium with (partial) non-disclosure. We first remark in passing 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 μ is such that outcomes are perfectly correlated across team members. Then for any deliberation protocol D, the unique equilibrium outcome is full disclosure.*

To highlight the differences between individual- and team-disclosure problems, for the rest of the paper we assume that μ is such that team members outcomes are "not too correlated."

Assumption 2. *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.*

Given this assumption, 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 often fails because the outside observer cannot fully attribute a non-disclosure decision to a specific team member.

3 Equilibrium Team Disclosure

A deliberation procedure D determines which team-members have complete "voice rights," and therefore can choose unilaterally to speak on behalf of the team. Specifically, remember that we say a team-member i can unilaterally choose disclosure if $D(\{i\}) = 1$. Theorem 1 shows that equilibrium team-disclosure distinguishes between these team-members, to whom the observer can fully attribute the team's decision to not disclose an outcome, and those who cannot unilaterally choose disclosure.

We say 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 $d(\omega) < 1$. An equilibrium has partial disclosure if it does not have full disclosure.

Theorem 1. *The following statements are 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*

$$\omega_i^{ND} = \min(\Omega_i) \textit{ in every partial-disclosure equilibrium.}$$

3. *Conversely, if $I \subseteq N$ is the set of team-members who cannot unilaterally choose disclosure, then there exists a* partial-disclosure equilibrium *where*

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

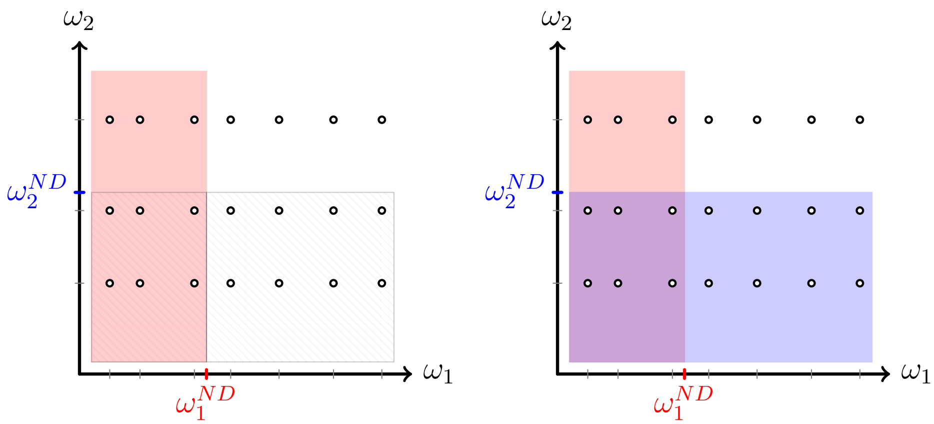


Figure 2: Both panels depict *candidate* team-disclosure equilibria in a team with $n = 2$. The left-hand panel supposes a deliberation protocol where team-member 1 is the team-leader. The right-hand panel supposes a protocol whereby disclosure decisions are made via consensus.

A full proof of Theorem 1 is available in the Appendix. 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, suppose 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 surely the realized outcome corresponded 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 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, let's refer to Figure 2. In each panel, the figure pictures the space of possible team outcomes in a two-people team, which correspond to outcome-values for person 1 (pictured in the x-axis) and outcome-values for person 2 (pictured on the y-axis). Conjecture an equilibrium where the observer is not maximally skeptical about either team-member, so that $\omega_1^{ND} > \min(\Omega_1)$ and $\omega_2^{ND} > \min(\Omega_2)$. In such an equilibrium, each team-member must recommend the disclosure of an outcome if and only if the realized value of their respective outcome component is larger than the conjectured no-disclosure beliefs about that component. In both panels, the red-shaded area depicts team-outcomes that team-member 1 recommends to conceal; and the blue-shaded area represents those that team-member 2 recommends to conceal. These recommendations are then aggregated according to the team's deliberation procedure. The left-hand side figure supposes that team-member 1 is the team-leader, and the right-hand side panel supposes that the deliberation procedure is such that disclosure is chosen via consensus.

In the former case, 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. But note that all outcomes in that area are such that ω_1 is smaller than the originally conjectured no disclosure belief ω_1^{ND}. And therefore it cannot be that the conjectured belief is Bayes-consistent; which thus implies that there is no equilibrium in which $\omega_1^{ND} > \min(\Omega_1)$, as described in statement 2 of the theorem.

Instead if disclosure is chosen via consensus, 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 those outcomes be concealed. But note that, 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 some "good news" about individual 2 are also not shown for their disclosure 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 a initially conjectured pair of no-disclosure beliefs, which is not maximally skeptical about either team-member, that satisfies Bayes-consistency.

Note that, when team-members cannot unilaterally choose to disclose the team's outcome, the existing partial-disclosure equilibrium is supported by the fact that the observer is not able to attribute "blame" for an outcome's non-disclosure to one particular team-member (or to all team-members). In this equilibrium, the team chooses not to disclose outcomes when the result of the team project is 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 Shin (1994). Accordingly, upon seeing that an outcome is not disclosed, the observer becomes skeptical *about the team*, and thus interprets non-disclosure as a *team failure*. Importantly, the team takes the negative repercussions of this failure as a group, instead of precisely revealing each team-member's contribution to the team's failure. In contrast, if a team-member can unilaterally disclose, then they are seen as "fully to blame" for a team's decision to not disclose, and the observer must thus be maximally skeptical about them in any equilibrium.

The results in Theorem 1 highlight that in a team context, a notion of aggregate skepticism (about the team's outcome) is not sufficient to determine the team's communication behavior in equilibrium. In fact, we see that the team's disclosure behavior is intimately connected to a targeted notion of skepticism, which we refer to as *blame* and its determination via the team's deliberation procedure. In section 3.1, we further explore the relation between the deliberation procedure, blame, and aggregate skepticism. 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 *de*creases in ω_i^{ND} as *in*creases in team-member i's blame; or equivalently as increases in the observer's skepticism that is targeted at team-member i. Note that our chosen measure of blame, ω_i^{ND}, is not normalized by the distribution of i's outcomes, but our results remain unchanged if we use the measure $\omega_i^{ND}/\left[\max(\Omega_i) - \min(\Omega_i)\right]$ instead. Our measure is also not normalized by the total aggregate blame, as for example $\omega_i^{ND}/\sum_{j\in N}\omega_j^{ND}$. Our choice reflects that in a Bayesian learning model as ours, blame is not a "zero-sum game." Because agents are not just "sharing a pie," but rather each signaling their outcome realizations to the observer, it is possible for all team-members to be blamed for a group failure, or for none of them to be blamed.

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

Proposition 1 considers the effect of marginal changes in the deliberation procedure D on the equilibrium blame vector ω^{ND} around a particular starting strict equilibrium. Lemmas 1 and 2 ensure that this is a well defined exercise. Also remember that a deliberation procedure can be thought of as a vector with 2^n entries, specifying for each subset $I \subseteq N$ a probability of disclosure $D(I) \in [0,1]$. Our assumptions require that $D(\varnothing) = 0$, $D(N) = 1$, and $I \subseteq I' \Rightarrow D(I) \leqslant D(I')$; and 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 $\mathcal{E}_\mu^D \neq \varnothing$. 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 procedure D and a distribution μ such that $\mathcal{E}_\mu^D \neq \varnothing$, and a strict equilibrium $\varepsilon \in \mathcal{E}_\mu^D$. In a neighborhood of D, there is a unique continuous selection E of the strict-equilibrium correspondence $\mathcal{E}_\mu^D$ such that $E(D) = \varepsilon$.*

Specifically, 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 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 Appendix contains proofs of both lemmas in this section and of Proposition 1. 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-diclosure belief about team-member i's outcome must decrease. An interpretation is that condition (2) ensures that team-member i's *voice rights* — their right to speak on behalf of the team — increase; and as a consequence i's blame for team failures also increase. 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

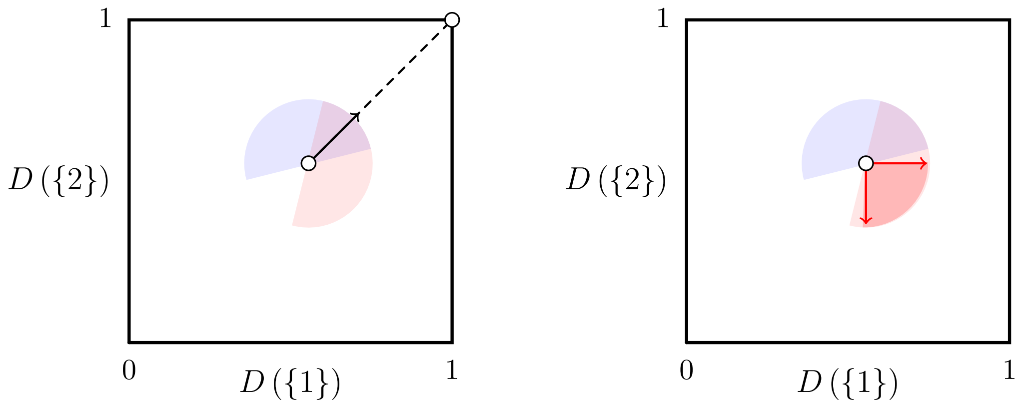


Figure 3: Both panels show the effect of changes in the deliberation procedure in a two-person team on the observer's no-disclosure belief about each team-member. Red-shaded areas indicate directions of changes in D that decrease ω_1^{ND} (increase team-member 1's blame) and blue-shaded areas indicate directions that increase team-member 2's blame. The left panel shows 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).

in a proportional way. Because the probability of disclosure after all 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 right 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 increased opportunity to enforce the disclosure of a given outcome, and therefore interprets the team's choice to not disclose an outcome as worse news about every individual team-member. In other words, a "aggregate increase" in all team-members' voice rights leads to a corresponding "aggregate increase" in their blame for team failures. Corollary 2 instead shows that a *relative* increase in a particular team-member's voice rights — which makes 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 $d\omega_i^{ND}/dD \leqslant 0$.

The right panel in Figure 3 shows the directions that increase a team-member's pivotality in a two-person team. Note that in a two-person team, 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 increases 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 partial-disclosure equilibria, when present, coexist with a full disclosure equilibrium. Theorem 2 below is a result refining the equilibrium set: we characterize deliberation procedure environments under which a "plausible" full disclosure equilibrium exists. By definition, full-disclosure equilibria are such that no-disclosure does not happen on the path of play — or perhaps only when all team members draw their worst-possible outcome — and therefore no-disclosure posteriors are vacuously Bayes-consistent. As usual with "forward induction" refinements, we wish to evaluate whether these off-path posteriors which support the full-disclosure equilibrium can be justified by some reasonable off-path behavior. In the context of team disclosure, we posit that even off-path beliefs should be justified by some behavior that is consistent with the team's deliberation protocol.

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

1. *For each $i, j \in N$ with $j \neq i$, $x_i(\omega)$ is constant with respect to ω_j.*

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).*

The definition states that a vector of no-disclosure beliefs is consistent with deliberation if there is some set of individual disclosure strategies (that ensure no-disclosure happens with positive probability), and such that the no-disclosure beliefs are Bayes-plausible given these strategies. Condition 1 also demands that the disclosure strategy each team-member uses to justify the vector ω^{ND} be independent of other team-members' outcomes — so that the overall team decision can be interpreted as purely aggregating stated individual preferences and nothing further. Note that, unlike in the equilibrium notion in Definition 1, we do not require the individual strategies to be optimal for any of the team-members. Our refinement therefore places a very weak requirement: that there be *any* set of such individual strategies that, once aggregated through the team's deliberation procedure, ensure the consistency of the conjectured no-disclosure beliefs.

Theorem 2 shows a condition on the deliberation protocol that is necessary and sufficient for full-disclosure equilibria to be consistent with deliberation. Informally, this condition requires the deliberation protocol to be such that a decision to disclose is easier to reach than a decision to conceal an outcome with some 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 $D(N \setminus J) < 1$ but $D(J) \neq 1$. For example, if we consider a team of only two individuals, then disclosure requires more consensus than concealing if and only if disclosure only happens with certainty if decided via perfect consensus between the two team members — or equivalently, iff neither team-member can unilaterally choose disclosure — so that $D(\{1\}), D(\{2\}) < 1$.

Theorem 2. *A full-disclosure equilibrium that is consistent with the deliberation procedure D exists if and only if D is such that disclosure requires more consensus than concealing.*

To understand the result, suppose there are only two team members, and suppose disclosure requires more consensus than concealing — for example, $D(\{1\}) = D(\{2\}) = 0$ — so that each team member can unilaterally choose to conceal a realization. Now suppose there is a full-disclosure equilibrium supported by a pair 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, and they would be able to unilaterally impose such non-disclosure. This would contradict 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 strategies consistent with the deliberation process; and therefore full-disclosure is not consistent with deliberation.

With some work shown in the Appendix, this argument generalizes to teams with more than two members, so long as the deliberation process is such that disclosing requires more consensus than concealing.[13] To see the other direction of Theorem 2, let's again consider that there are only two team members, and now suppose that disclosure does not require more consensus than concealing. That is, either $D(\{1\}) = 1$ or $D(\{2\}) = 1$ — suppose the former is true for the sake of this argument. Then there exists a full disclosure equilibrium where $\omega_1^{ND} = \min(\Omega_1)$ and $\omega_2^{ND} > \min(\Omega_2)$. Moreover, these off-path 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 $\omega \in \Omega$.

Together, Theorems 1 and 2 characterize how equilibrium disclosure "decreases" after an increase in the degree of consensus required for the team to disclose. In the first result, we see that unless disclosure is very easy — in the sense that it can be chosen unilaterally by all team-members — then full-disclosure is not the unique equilibrium outcome. Theorem 2 strengthens the observation by delineating a necessary and sufficient condition under which not only a partial-disclosure equilibrium exists, but also it is more plausible than full-disclosure. The predictions made by Theorems 1 and 2 are particularly clear when we consider anonymous deliberation (so that the team's disclosure decision depends only on the *number* of team-members who recommend disclosure). Within the class of anonymous deliberation procedures, disclosure can be chosen unilaterally if $D(I) = 1$ whenever $|I| \geqslant 1$; and disclosure requires more consensus than concealing if $D(I) = 1$ whenever $|I| \geqslant n/2$.

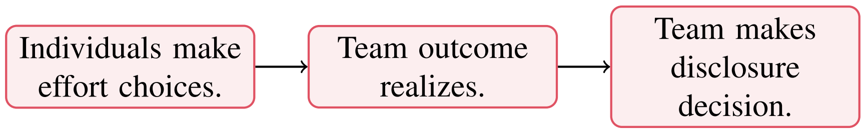
Corollary 3 (to Theorems 1 and 2)**.** *Suppose D is an anonymous deliberation procedure, with $D(I) = 1$ if and only if $|I| \geqslant k$. Full-disclosure is the unique equilibrium outcome if and only if $k = 1$; and full-disclosure is consistent with deliberation if and only if $k \leqslant n/2$.*

4 Deliberation and Incentives

So far in the paper, we studied a problem where a team chooses how to communicate about their group outcomes with outside observers, taking team production as given by the outcome distribution μ. In this section, we explore the effects of features of equilibrium team disclosure

[13] 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 $i \in I$. But 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, $D(J) < 1$.

on individual incentives to contribute to team production in the first place. To do so, we add a pre-disclosure stage to the team's problem.



Formally, at the initial stage, each agent unilaterally and covertly makes an effort decision: team-member $i \in N$ chooses $e_i \in \{0, 1\}$, incurring in 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 distribution $\mu(\cdot; e)$, which now depends on the effort vector e chosen by the team. After the team outcome realizes, the communication 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 team's deliberation procedure D. The observer sees the disclosed/not-disclosed outcome, according tot he 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 $\Omega = \Omega_1 \times ... \times \Omega_n$, where $\Omega_i \subset \mathbb{R}$ has at least 2 elements for all $i \in N$. Moreover, effort is productive, so that*[14]

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

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, we understand that the deliberation procedure is fixed at some exogenously given D, and commonly understood by all team-members and by the observer. One interpretation is that, even prior to the effort-stage, team-members collectively pick a deliberation procedure — that is, they agree on a governance structure to guide their future communication decisions. From that lens, our exercise in this section describes what types of such governance structures a team should implement if they wish to incentivize individuals to contribute effort to

[14]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).

team production. Such a "deliberation design" problem is relevant in applied scenarios where team-members cannot contract on outcomes of team-production — which may be especially hard in career concerns environments, where each individual's outcome is given by the observer's perception of their type, which is not an easily measurable variable — but can contract on how to map observable individual recommended decisions into overall team decisions.

A stage in which "voice rights" are decided indeed occur in the applied scenarios we consider. For example, as mentioned in the introduction, at the beginning of congress, congressional committees explicitly write down and pledge to abide by a set of formal rules guiding how the committee will later make decisions regarding the publication of reports and gathered evidence. In entrepreneurial and other productive teams, much importance is placed in establishing a "corporate culture" and governance to guide future team communications. Indeed, corporate teams also strongly value communicating the team's culture and hierarchical structures to third parties — in our model, ensuring that the observer understands the team's deliberation procedure. We comment further on the interpretation of deliberation as corporate culture later in this section.

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 do so in terms of whether — and for what cost vectors — each deliberation procedure can implement an equilibrium in which all team-members exert costly effort.[15] 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 below establishes the basis for this analysis, clarifying the relation between disclosure strategies implemented in 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 any team-disclosure strategy $d : \Omega \to [0,1]$, let $fe(d) \subset \mathbb{R}^n_{++}$ be its corresponding *full-effort set*. That is, $c \in fe(d)$ if, given the team-disclosure strategy d, there is an equilibrium of the effort-choice stage in which $e_i = 1$ for all $i \in N$. 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.

[15] Our results in section 2 show that there are often multiple equilibria in the team-disclosure game; and 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.

Lemma 3. *A team-disclosure strategy $d : \Omega \to [0,1]$ implements full effort for a given cost vector $c \in \mathbb{R}^N$ — that is, $c \in fe(d)$ — if and only if, for every $i \in N$,*[16]

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

The expression in (4) clarifies how the selective disclosure of the teams 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: it 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.

The second term corresponds to extra incentives provided by selective non-disclosure. Specifically, it measures how the observer misattributes blame to team-member i when they do not exert effort and a team-failure ensues. More 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 chooses instead to deviate to no effort, so the true effort vector is $e = e_{N\setminus i}$; and suppose that the drawn outcome after that deviation is such that the team chooses not to disclose it. Because i's effort is covert, the observer still calculates i's "blame" in that case under the presumption that all team-members exerted effort — and therefore i's value is $\mathbb{E}\left(\omega_i|ND; e_N\right)$, as opposed to the "correct" blame assessment $\mathbb{E}\left(\omega_i|ND; e_{N\setminus i}\right)$.[17]

The misattributed blame term is positive — and therefore selective non-disclosure through 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 no effort. Intuitively, this is the case when the team's disclosure decisions are more correlated with i's outcome under the full-effort vector e_N than under effort $e_{N\setminus i}$.[18]

The decomposition of individual effort incentives in Lemma 3 parallels incentive decom-

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

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

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

[17] 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 effort is chosen covertly by each agent.

[18] Indeed, the following rewriting of the left-hand side of (4) expresses i's effort gains directly in terms of the

positions 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. And moreover, the efficient allocation of control over the relationship asset is the one that best aligns this second term with efficient effort incentives.[19] 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 Effort Environments and Effective Deliberation

We now use the characterization in Lemma 3 to study which types of deliberation procedures effectively incentivize effort in a variety of effort environments — where an effort environment describes the impact of each individual's effort on the distribution of team outcomes. We say effort is *purely self-improving* if, for every $i \in N$ and every $I \subset N$,[20]

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

The notation $\mu_{N\setminus i}(\cdot; e_I)$ indicates the joint distribution of outcomes of all team members except 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 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 thus 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 team-mates.

improvement of the covariance between i's outcome and disclosure:

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

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

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

[20] 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 $\mathbb{P}(X \in U) > \mathbb{P}(Y \in U)$ for every upper set $U \in \Omega$.

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

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

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 where 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 then 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-externalities environments parallels the distinction between selfish and cooperative investments in a hold-up context proposed by Che and Hausch (1999).

Proposition 2 ranks deliberation procedures in both of these effort environments. To that end, we formally introduce an order over deliberation procedures. For a given procedure D, we let $FE(D) \subset \mathbb{R}^n_{++}$ be its corresponding full-effort set: $c \in FE(D)$ if, given D, there is a team-disclosure strategy d with $c \in fe(d)$ that can be sustained in a team-disclosure equilibrium under the full-effort outcome distribution $\mu(\cdot; e_N)$. And we say that a deliberation procedure D *dominates* a deliberation procedure D' if $FE(D') \subseteq FE(D)$. Also remember that D is the *unilateral disclosure* protocol if every team-member can unilaterally choose disclosure; and the *consensus disclosure* protocol is such that every team-member can unilaterally veto disclosure.

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.*

The proposition as stated describes purely self-improving and purely team-improving environments, but the respective statements also hold for environments close enough to either of these extreme cases. Likewise, Proposition 2 holds for comparisons between almost-unilateral and almost-consensus deliberation procedures.

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. Indeed, we show that 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; and 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 draw an outcome for which they favor disclosure; therefore improving the odds that i's disclosure recommendation is pivotal to the 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 where the disclosure of their own good outcome realizations is vetoed by others.

Proposition 3 below considers effort environments such that effort increases the correlation between team-members' outcomes. We can interpret these as investments in some common component of team-production. The proposition states that if each team-member's effort sufficiently improves the correlation between all team-member's outcomes, then the unilateral disclosure procedure is dominated by all other procedures. To that end, 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 $\omega_1 = ... = \omega_n$.[21]

[21] These assumptions are made for notational convenience. Proposition 3 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 $i, j \in N$.

Proposition 3. *Suppose 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 $\nu \succsim_{FOS} \mu \succsim \mu(\cdot; e_{N\setminus i})$ for every $i \in N$. Consider varying the correlation in $\mu(\cdot; e_N)$ by letting, for $\epsilon \in (0, 1)$,*

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

Let D be a 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 $\epsilon > \bar{\epsilon}$ and $\mu_\epsilon(\cdot; e_N)$ is the full-effort distribution, then D' strictly dominates D.

One way to interpret deliberation procedures in real-world team production environments is as the "corporate culture" in 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 fully held 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 not to "play the blame game." The article "When Transparency Backfires, and How to Prevent It," Harvard Business Review, acknowledges the benefits of such cultures, 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 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.[22]

5 Disclosure and Incentives in Binary Environments

This section further characterizes team disclosure, and its relation to effort incentives, in an environment where each individual's outcome is binary. We say the distribution μ has *binary outcomes* if $\Omega_i = \{\omega_{\ell,i}, \omega_{h,i}\}$, with $\omega_{\ell,i} < \omega_{h,i}$ for each $i \in N$.

5.1 Equilibrium Team Disclosure with Binary Outcomes

Proposition 4 characterizes, in the binary outcome setting, the equilibrium with the least disclosure for a given deliberation procedure D. We say an equilibrium is the least-informative equilibrium (or has the least disclosure) for a given deliberation procedure D if its team-disclosure strategy d is such that $d(\omega) \leqslant d'(\omega)$ for every other equilibrium team-disclosure strategy d' for the same procedure D.

Proposition 4. *Suppose μ has binary outcomes. Then, for a given deliberation procedure D, the least informative equilibrium is such that*

(i) Each team-member recommends disclosure if and only if they draw a high outcome:

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

(ii) The observer's no-disclosure belief about team-member i is maximally skeptical if and only if i can unilaterally choose disclosure:

$$\omega_i^{ND} = \omega_{\ell,i} \Leftrightarrow D(\{i\}) = 1.$$

[22] 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 documentations, 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.

When outcomes are binary it is a weakly dominant strategy for each agent i to recommend disclosure if and only if $\omega_i = \omega_{h,i}$. Moreover, in any equilibrium, every agent must recommend disclosure weakly more than under that strategy — for it is always individually optimal to recommend the disclosure of an agent's own high outcomes with certainty. Therefore, the equilibrium implied by all agents using this weakly dominant strategy must be the equilibrium with least disclosure. This is stated in part (i) of the proposition. Part (ii) states the no-disclosure beliefs implied by the team-disclosure strategy given by the aggregation of these individual disclosure recommendation strategies via protocol D. See the Appendix for a more detailed proof of the proposition.

We know from Theorem 1 that for any deliberation procedure D, the most informative equilibrium (the equilibrium with most disclosure) invoves full-disclosure of every outcome. The equilibrium set includes other equilibria beyond the most informative and the least informative, because each team-member about whom the observer is maximally skeptical is indifferent between disclosing and not disclosing their low outcome realizations.

In this binary environment, we can evaluate the effect of increasing the correlation between team-members' outcomes on the probability that outcomes are disclosed and on the equilibrium blame vector. This complements the analysis in section 3.1, which evaluates the effect of changes in the deliberation procedure on equilibrium blame. Corollary 4 follows from Proposition 4 and shows that when the correlation between team-members' outcomes increases, equilibrium disclosure becomes more correlated with each agent's outcome — in the sense that the disclosure of an agent's high outcome becomes more likely and the disclosure of that same agent's low outcome becomes less likely. A consequence is that every agent becomes "more to blame" for team failures. To state this result, we propose an ordering on the correlation of team-members' outcomes. Suppose μ and μ' are two outcome distributions with the same marginal distributions for every $i \in N$. We say μ' features *more outcome correlation* than μ if for every $i \in N$, $\mu_{N\setminus i}(\cdot|\omega_{\ell,i}) \succsim_{FOS} \mu'_{N\setminus i}(\cdot|\omega_{\ell,i})$ and $\mu'_{N\setminus i}(\cdot|\omega_{h,i}) \succsim_{FOS} \mu_{N\setminus i}(\cdot|\omega_{h,i})$. (Where $\mu_{N\setminus i}$ and $\mu'_{N\setminus i}$ indicate the distribution of outcomes for team-members other than i.)

Corollary 4. *Let μ and μ' have the same marginal distributions over $\Omega_i = \{\omega_{\ell,i}, \omega_{h,i}\}$ for every $i \in N$, and suppose μ' has more outcome correlation than μ. The least-informative equilibrium team-disclosure strategy is the same under μ and μ' — denote it d. The following hold:*

(i) Each i's high (low) outcomes are disclosed with higher (lower) probability under μ':

$$d_{\mu'}(\omega_{h,i}) = \sum_{\{\omega:\omega_i=\omega_{h,i}\}} \mu'(\omega)d(\omega) \geqslant \sum_{\{\omega:\omega_i=\omega_{h,i}\}} \mu(\omega)d(\omega) = d_{\mu}(\omega_{h,i}).$$

$$d_{\mu'}(\omega_{\ell,i}) = \sum_{\{\omega:\omega_i=\omega_{\ell,i}\}} \mu'(\omega)d(\omega) \geqslant \sum_{\{\omega:\omega_i=\omega_{\ell,i}\}} \mu(\omega)d(\omega) = d_\mu(\omega_{\ell,i}).$$

(ii) All team-members are more to blame for team failures under μ':

$$\omega_i^{ND}(\mu') \leqslant \omega_i^{ND}(\mu) \text{ for all } i \in N.$$

For this result, we consider increasing the correlation between agents' outcomes without ever making them perfectly correlated — remember that we maintain the assumption of Ω being a product space throughout. However, suppose we consider a sequence of "correlation increases" that converges to fully correlated outcomes; that is, for all $i, j \in N$, $\mu^k(\omega_{h,i}|\omega_{h,j}) \to 1$ and $\mu^k(\omega_{\ell,i}|\omega_{\ell,j}) \to 1$. In that limit, we have that in the least-informative equilibrium $d_k(\omega_{h,i}) \to 1$ and $d_k(\omega_{\ell,i}) \to 0$ for any $i \in N$. In words, all agents' high outcomes are with certainty and all agents' low outcomes are not disclosed, also with certainty, in the limit as correlation becomes perfect. And therefore in the least informative equilibrium — and consequently in all equilibria — $\omega_i^{ND}(\mu^k) \to \omega_{\ell,i}$.

5.2 Effort-Maximizing Deliberation in Binary Environments

The results in section 4 rank different deliberation procedures in terms of their effort-incentives provision. Using the further tractability implied by binary environments, we can strengthen those results and provide a characterization of deliberation procedures that *maximize* effort incentives. In a fully general binary environment, such an effort-maximizing deliberation procedure may not always exist, since a deliberation procedure may improve effort incentives for some team-members, but decrease them for other individuals. To ensure that effort-maximizing procedures exist, we consider only symmetric environments:

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

(i) The deliberation procedure is symmetric: for all $X \subseteq N$, $D(X)$ depends only on the cardinality of X, $|X|$.

(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, then the outcome distribution is symmetric:

$$|\{i : \omega_i = \omega_h\}| = |\{i : \omega'_i = \omega_h\}| \Rightarrow \mu(\omega; e_N) = \mu(\omega'; e_N).$$

(iv) Agents' efforts affect the outcome distribution symmetrically:

$$\textit{For every } i, j \in N, \mu(\omega_i = w, \omega_j = w', \omega_{N\setminus i,j}; e_{N\setminus i}) = \mu(\omega_i = w', \omega_j = w, \omega_{N\setminus i,j}; e_{N\setminus j}).$$

The symmetry assumption requires the deliberation procedure to be symmetric, as well as the outcome distribution and 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.

Lemma 4. *Under Assumption 4, there are at most two symmetric team-disclosure equilibria in the subgame where all team-members exert effort. First, a full-disclosure equilibrium with $\omega_i^{ND} = \omega_\ell$ for all $i \in N$. Second, if $D(\{i\}) \neq 1$, there exists a unique symmetric partial-disclosure equilibrium.*

Let $SFE(D)$ be the symmetric full-effort set of the symmetric deliberation procedure D: a cost vector $c \in SFE(D)$ if, given D, there is a team-disclosure strategy d, with $c \in fe(d)$, that can be sustained in a symmetric team-disclosure equilibrium under the symmetric full-effort distribution $\mu(\cdot; e_N)$. We say the symmetric deliberation procedure D maximizes effort incentives if for any symmetric procedure D', $SFE(D') \subseteq SFE(D)$. By symmetry, we also know that for every D, the set $SFE(D)$ 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'.

Lemma 4 implies that, to evaluate the effort incentives provided by a symmetric procedure D, it suffices to consider the incentives provided by its unique partial-disclosure equilibrium of the team-disclosure subgame. 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 unique symmetric partial-disclosure equilibrium under deliberation procedure D; and $\bar{c}(D) = 0$ if D is such that all team-members can unilaterally disclose. The maximum cost $\bar{c}(D)$ is thus the unique value of c that satisfies equation (4) with equality. Lemma 5 follows from the fact that $\bar{c}(D)$ is a continuous function of D, and that the space of symmetric deliberation procedures is compact.

Lemma 5. *In a binary symmetric environment, there exists a symmetric deliberation procedure D that maximizes effort incentives. That is, for every symmetric deliberation procedure D', $SFE(D') \subseteq SFE(D)$.*

5.2.1 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 these measure the ratio between the probability that both agents had a low outcome relative to the probability that exactly one of them did. (Note that ρ does not depend on the selected agent i because of symmetry.) 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 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' outcomes.

Proposition 5 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 5 shows that the effort-maximizing value of $D(1)$ is increasing in Δ_σ and decreasing in Δ_ρ.

Proposition 5. *Fix $\bar{\rho}$ and $\bar{\sigma}$, and let ρ and σ vary. Denote by D^* the level of $D(1)$ in the effort-maximizing deliberation procedure.*

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-increasing in σ and non-decreasing in ρ; or equivalently.

2. *If $\Delta_\rho > 0$, then D^* is a continuous non-increasing (non-decreasing) function of σ (of ρ).*

The proof of Proposition 5 is in the Appendix. This proposition complements the results in section 4 by showing that, in a binary-outcome environment, effort-maximizing deliberation procedures are "more unilateral" when effort is more self-improving or when effort correlates individuals' outcomes to a lesser extent.

5.2.2 Effort-Maximizing Deliberation in Larger Teams

We now consider, in a numerical exercise, effort-maximizing deliberation in teams with more than two team-members. To that end, we specify outcome distributions as follows. First suppose 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 what is 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 incen-

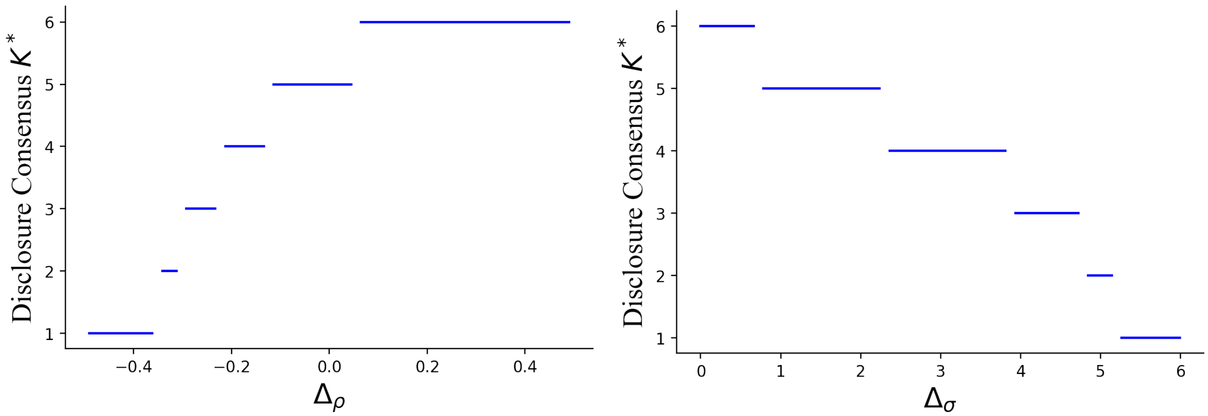


Figure 4: Effort-maximizing degree of consensus required for disclosure as a function of Δ_ρ and Δ_σ. In the left panel, $\bar{\rho} = .5$ and changes in ρ create the variation in Δ_ρ. Other parameters are fixed at $h_T = h_i = h_j = .5$ and $\bar{h} = .6$. In the right panel, $\bar{h} = .6$ and $h_j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

tives provided by different consensus levels K. We also 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 simmulation are specified in the figure, but the results 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**: 777-795.

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

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

Austen-Smith, David, and Timothy Feddersen. (2005) "Deliberation and Voting Rules." in: David Austen-Smith, and John Duggan (eds.), *Social Choice and Strategic Decisions*, Springer:

269-316.

Baldwin, David G. (2001) “How to Win the Blame Game.” *Harvard Business Review 79*: 144-144.

Bardhi, Arjada, and Nina Bobkova. (2023) “Local Evidence and Diversity in Minipublics.” *Journal of Political Economy*, forthcoming.

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

Battaglini, Marco. (2022) “Multiple Referrals and Multidimensional Cheap Talk,” *Econometrica* **70**: 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**: 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**: 125-147.

de Cremer, David. (2016) “When Transparency Backfires, and How to Prevent It.” *Harvard Business Review*.

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

Dranove, David, and Ginger Zhe Jin. (2010) “Quality Disclosure and Certification: Theory and Practice.” *Journal of Economic Literature* **48**: 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**: 1362-1397.

Fortunato, S., Bergstrom, C.T., Börner, K., Evans, J.A., Helbing, D., Milojević, S., Petersen, A.M., Radicchi, F., Sinatra, R., Uzzi, B. and Vespignani, A. (2018). “Science of Science,” *Science*, **359**.

Freeland, Robert F., and Ezra W. Zuckerman. (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**: 300-322.

Gerardi, Dino, and Leeat Yariv. (2007) "Deliberative Voting." *Journal of Economic Theory* **134**: 317-338.

Gerardi, Dino, and Leeat Yariv. "Information Acquisition in Committees." *Games and Economic Behavior* **62**: 436-459.

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

Giovannoni, Francesco, and Daniel J. Seidmann. (2007) "Secrecy, Two-Sided Bias and the Value of Evidence." *Games and Economic Behavior* **59**: 296-315.

Grossman, Sanford J. (1981) "The Informational Role of Warranties and Private Disclosure about Product Quality." *Journal of Law and Economics* **24**: 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**: 691-719.

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

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

Haghpanah, Nima, Aditya Kuvalekar, and Elliot Lipnowski. (2022) "Selling to a Group." *working paper*.

Holmström, Bengt. "Managerial Incentive Problems: A Dynamic Perspective." *Review of Economic Studies* **66**: 169-182.

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

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

Itoh, Hideshi. (1991) "Incentives to Help in Multi-Agent Situations." *Econometrica*: 611-636.

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

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

Lazear, Edward P., and Kathryn L. Shaw. (2007) “Personnel Economics: the Economist’s View of Human Resources.” *Journal of Economic Perspectives* **21**: 91-114.

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

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

Martini, Giorgio. (2018) ”Multidimensional Disclosure,” *working paper*.

Matthews, Steven, and Andrew Postlewaite. (1985) “Quality Testing and Disclosure,” *RAND Journal of Economics*: 328-340.

Milgrom, Paul (1981) “Good News and Bad News: Representation Theorems and Applications.” *Bell Journal of Economics*: 380-391.

Milgrom, Paul. (2008) ”What the Seller Won’t Tell You: Persuasion and Disclosure in Markets.” *Journal of Economic Perspectives* **22**: 115-131.

Milgrom, Paul, and John Roberts. (1986) “Relying on the Information of Interested Parties.” *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**.

Onuchic, Paula. (2022) “Advisors with Hidden Motives,” *working paper*.

Onuchic, Paula, and Debraj Ray. (2023) “Signaling and Discrimination in Collaborative Projects.” *American Economic Review* **113**: 210-252.

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

Rappoport, Daniel. (2023) “Evidence and Skepticism in Verifiable Disclosure Games,”*working paper*.

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

Schwert, G. William. (2021) “The Remarkable Growth in Financial Economics, 1974–2020,” *Journal of Financial Economics* **140**: 1008-1046.

Seidmann, Daniel J., and Eyal Winter. (1997) “Strategic Information Transmission with Verifiable Messages.” *Econometrica*: 163-169.

Shaked, Moshe, and J. George Shanthikumar, eds. *Stochastic Orders*. New York, NY: Springer

New York.

Shin, Hyun Song. "The Burden of Proof in a Game of Persuasion." *Journal of Economic Theory* **64**: 253-264.

Shishkin, Denis. (2021) "Evidence Acquisition and Voluntary Disclosure," *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**: 337-372.

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

Winter, Eyal. (2004) "Incentives and Discrimination." *American Economic Review* **94**: 764-773.

Zuckerman, Ezra W. (2010) "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 $\hat{\omega}_j = \omega$ 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 $\mathbb{E}[\omega_i|$ 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.

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

First fix a full-support outcome distribution μ. 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.

Claim 1. *Consider the following class of deliberation procedures: $D(\varnothing) = 0$, $D(N) = 1$, and, for some $\epsilon \in (0, 1)$, $\epsilon \leqslant D(I) < 1$ for every $I \notin \{\varnothing, N\}$. For every $i \in N$ and every $k \in \{2, ..., |\Omega_i|\}$, there exists an $\epsilon_{ik} \in (0, 1)$ such that if $\epsilon > \epsilon_{ik}$, then no partial-disclosure equilibrium exists in which*

$$\omega_i > \omega_i^k \Rightarrow x_i(\omega) = 1 \text{ and } \omega_i \leqslant \omega_i^k \Rightarrow x_i(\omega) < 1. \tag{6}$$

Proof of Claim. Fix a team-member i and some $k \in \{2, ..., |\Omega_i|\}$. Consider a candidate partial-disclosure equilibrium, under a deliberation protocol in the class defined in the statement of the claim, in which

$$\omega_i > \omega_i^k \Rightarrow x_i(\omega) = 1 \text{ and } \omega_i \leqslant \omega_i^k \Rightarrow x_i(\omega) < 1.$$

In such an equilibrium, we must 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 remember $\Pi_X(\omega)$ is the probability that exactly the subset X of team-members recommend the disclosure of outcome ω. Because i recommends disclosure of all outcomes with $\omega_i > \omega_i^k$, it must be that $d(\omega) \geqslant \epsilon$ for all such ω. In contrast, there exist outcomes with $\omega_i < \omega_i^k$ — for example, $\omega = (\min(\Omega_1), ..., \min(\Omega_n))$ — for which $d(\omega) = 0$ in any such equilibrium. Consequently, if $\epsilon \in (0, 1)$ is sufficiently large — above some $\epsilon_{ik} \in (0, 1)$ — then we must have $\omega_i^{ND} < \omega_i^k$, which contradicts the initially assumed equilibrium behavior of individual i. □

Let $\bar{\epsilon} = \max_{i,k} \epsilon_{ik}$ and consider deliberation procedures satisfying $\bar{\epsilon} \leqslant D(I) < 1$. By Claim 1, we know that there is no partial disclosure equilibrium in which equilibrium disclosure recommendations satisfy (6) for any $i \in N$ and any $k \in \{2, ..., |\Omega_i\}$. But e know from Theorem 1 that a partial-disclosure equilibrium exists, in which $\omega_i^{ND} > \min(\Omega_i)$ for all $i \in N$. So it must be that in such an equilibrium, for each i,

$$\omega_i > \min(\Omega_i) \Rightarrow x_i(\omega) = 1 \text{ and } \omega_i = \min(\Omega_i) \Rightarrow x_i(\omega) = 0.$$

And consequently, in such partial-disclosure equilibrium, we must have $\min(\Omega_i) < \omega_i^{ND} < \omega_i^2$ for every team-member $i \in N$. The equilibrium is therefore strict.

Now for the second statement in the lemma, fix a deliberation procedure D with $D(\{i\}) < 1$ for every $i \in N$. The proof uses the following claim, analogous to Claim 1. The proof of the claim, and that of the statement follow analogously to the proof of the first statement.

Claim 2. *Consider the following class of outcome distributions: $\mu(\omega) > 0$ for all $\omega \in \Omega$ and $\mu(\omega) < \epsilon$ for all $\omega \neq (\min(\Omega_1), ..., \min(\Omega_n))$. For every $i \in N$ and every $k \in \{2, ..., |\Omega_i|\}$, there exists an $\epsilon_{ik} \in (0, 1)$ such that if $\epsilon < \epsilon_{ik}$, no partial-disclosure equilibrium exists in which*

$$\omega_i > \omega_i^k \Rightarrow x_i(\omega) = 1 \textit{ and } \omega_i \leqslant \omega_i^k \Rightarrow x_i(\omega) < 1.$$

□

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 $i \in N$,

$$\hat{\omega}_i^{ND} = \frac{\sum_\Omega \omega_i(1-\hat{d}(\omega))\mu(\omega)}{\sum_\Omega(1-\hat{d}(\omega))\mu(\omega)}, \text{ with } \hat{d}(\omega) = \sum_{X \subseteq N} \Pi_X(\omega)\hat{D}(X) \text{ for every } \omega \in \Omega, \quad (7)$$

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 there exists a continuous selection E of the strict-equilibrium correspondence such that $E(D) = \varepsilon$.

Further note that for any continuous selection E' of the strict-equilibrium correspondence, for any $\hat{D}$ close enough to D, $\hat{\omega}^{ND} = E(\hat{D})$ must satisfy $\omega_i^k < \hat{\omega}_i^{ND} < \omega_i^{k+1}$ for each i — where ω_i^k and ω_i^{k+1} are such that $\omega_i^k < \omega_i^{ND} < \omega_i^{k+1}$. So 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 $\omega^{\hat{N}D}$ must satisfy (7), and so $E(\hat{D}) = E'(\hat{D})$. Therefore, in a neighborhood of D, there exists a *unique* continuous selection of the strict-equilibrium correspondence. □

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 $\mu(H(\omega) = I)$ be the probability that an outcome ω realizes which is a high realization for exactly team-members I.

Lemma 6. *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 $i \in N$,*

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

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.

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

□

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

$$\begin{aligned} d\omega_i^{ND} = &\sum_{I \subseteq N} \frac{\mu(H(\omega) = I)(1 - D(I))}{\sum_{I' \subseteq N} \mu(H(\omega) = I')(1 - D(I'))} \left[\omega_i^{ND} - \mathbb{E}\left(\omega_i | H(\omega) = I\right)\right] \frac{dD(I)}{(1 - D(I))} \\ \leqslant &\, m \left[\sum_{i \in I \subseteq N} \frac{\mu(H(\omega) = I)(1 - D(I))}{\sum_{I' \subseteq N} \mu(H(\omega) = I')(1 - D(I'))} \left(\omega_i^{ND} - \mathbb{E}\left(\omega_i | H(\omega) = I\right)\right) \right] \\ &+ M \left[\sum_{i \notin I \subseteq N} \frac{\mu(H(\omega) = I)(1 - D(I))}{\sum_{I' \subseteq N} \mu(H(\omega) = I')(1 - D(I'))} \left(\omega_i^{ND} - \mathbb{E}\left(\omega_i | H(\omega) = I\right)\right) \right] \\ \leqslant &\, m \left[\sum_{I \subseteq N} \frac{\mu(H(\omega) = I)(1 - D(I))}{\sum_{I' \subseteq N} \mu(H(\omega) = I')(1 - D(I'))} \left(\omega_i^{ND} - \mathbb{E}\left(\omega_i | H(\omega) = I\right)\right) \right] \\ = &\, m \left[\omega_i^{ND} - \sum_{I \subseteq N} \frac{\mu(H(\omega) = I)(1 - D(I))\mathbb{E}\left(\omega_i | H(\omega) = I\right)}{\sum_{I' \subseteq N} \mu(H(\omega) = I')(1 - D(I'))} \right] = 0. \end{aligned}$$

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 7. *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 7. One direction ($\Rightarrow$) is trivial given the definition of "disclosing requires more consensus than concealing." Consider the other direction ($\Leftarrow$). 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 some subset $I \subseteq N$ such that $D(I) = 1$ and $D(N \setminus I) < 1$, and take some $J \subset I$ such that $D(N \setminus J) < 1$. If $D(J) < 1$, then we know that there is a set $J \subset I$ such that $D(N \setminus J) < 1$ and $D(J) < 1$. 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 we know that there is a set $K \subset I$ such that $D(N \setminus K) < 1$ and $D(K) < 1$. If not, we can repeat this procedure until we find such a subset of I — the repetition of the procedure must return a subset of I that satisfies the conditions, because we know that $D(N) = 1$. □

We can now prove the theorem in two parts:

Part 1 ($\Rightarrow$) . If disclosing does not require more consensus than concelaing, 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 7, 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

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

Claim 3. *The set I is non-empty, and $D(I) = 1$.*

Proof of Claim. Suppose towards a contradiction that $D(I) < 1$ (which would vacuously hold if I were empty). Then every member of subgroup $N \setminus I$ 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. □

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

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

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

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

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

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

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

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

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

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

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 8. *If effort is purely self-improving, then for any full-effort equilibrium disclosure rule d with $d(\omega) < 1$ for some $\omega \in \Omega$, we have that for all $i \in N$*

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

Proof of Lemma. For notational convenience, we fix some team-member $i \in N$, and let $\nu_w = \mu_i(\cdot|w; e_N)$, $\hat{\nu}_w = \mu_i(\cdot|w; e_{N\setminus i})$ for all $w \in \Omega_{N\setminus i}$ and $\eta = \mu_{N\setminus i}(\cdot; e_N) = \mu_{N\setminus i}(\cdot; e_{N\setminus i})$. Therefore, for any equilibrium disclosure rule d with d with $d(\omega) < 1$ for some $\omega \in \Omega$, we have

$$\mathbb{E}\left(\omega_i|ND; e_{N\setminus i}\right) = \frac{\sum_{\Omega_{N\setminus i}} \sum_{\Omega_i} \omega_i (1 - d(\omega_i, w)) \hat{\nu}_w(\omega_i) \eta(w)}{\sum_{\Omega_{N\setminus i}} \sum_{\Omega_i} (1 - d(\omega_i, w)) \hat{\nu}_w(\omega_i) \eta(w)}. \tag{9}$$

$$\mathbb{E}\left(\omega_i|ND; e_N\right) = \frac{\sum_{\Omega_{N\setminus i}} \sum_{\Omega_i} \omega_i (1 - d(\omega_i, w)) \nu_w(\omega_i) \eta(w)}{\sum_{\Omega_{N\setminus i}} \sum_{\Omega_i} (1 - d(\omega_i, w)) \nu_w(\omega_i) \eta(w)}. \tag{10}$$

Now observe that for every such disclosure rule d, and every realization $w \in \Omega_{N\setminus i}$ of the outcomes of team-members $N \setminus i$, there is some probability $\alpha(w)$ that the outcome is *not disclosed* regardless of the realization ω_i, some probability $\beta(w)$ that the realization ω_i (and therefore i's individual disclosure strategy) is pivotal for the team-disclosure decision, and some probability $1 - \alpha(w) - \beta(w)$ that the outcome is disclosed, regardless of the realization ω_i. Using this, we can rewrite (9) as

$$\mathbb{E}\left(\omega_i|ND;e_{N\setminus i}\right)=\frac{\sum_{\Omega_{N\setminus i}}\left[\alpha(w)\sum_{\Omega_i}\omega_i\hat{\nu}_w(\omega_i)+\beta(w)\sum_{\omega_i\leqslant\omega_i^{ND}}\omega_i\hat{\nu}_w(\omega_i)\right]\eta(w)}{\sum_{\Omega_{N\setminus i}}\left[\alpha(w)\sum_{\Omega_i}\hat{\nu}_w(\omega_i)+\beta(w)\sum_{\omega_i\leqslant\omega_i^{ND}}\hat{\nu}_w(\omega_i)\right]\eta(w)},\qquad(11)$$

Of course, we can rewrite (10) analogously. Let $\hat{\mathcal{V}}_w$ be the cdf implied by $\hat{\nu}_w$, so that for each $\omega_i\in\Omega_i$, $\hat{\mathcal{V}}_w(\omega_i)=\sum_{v_i\leqslant\omega_i}\hat{\nu}_w(v_i)$. And let $\hat{\mathcal{V}}_w^{-1}$ be the quantile function implied by $\hat{\mathcal{V}}_w$: for each $q\in[0,1]$, $\hat{\mathcal{V}}_w(q)=\inf\{\omega_i:\hat{\mathcal{V}}_i(\omega_i)\geqslant q\}$. We can rewrite (11) as

$$\mathbb{E}\left(\omega_i|ND;e_{N\setminus i}\right)=\frac{\sum_{\Omega_{N\setminus i}}\left[\alpha(w)\int_0^1\hat{\nu}_w^{-1}(q)dq+\beta(w)\int_0^{\hat{\mathcal{V}}_w(\omega_i^{ND})}\hat{\nu}_w^{-1}(q)dq\right]\eta(w)}{\sum_{\Omega_{N\setminus i}}\left[\alpha(w)+\beta(w)\hat{\mathcal{V}}_w(\omega_i^{ND})\right]\eta(w)},\qquad(12)$$

where $\omega_i^{ND}=\mathbb{E}\left(\omega_i|ND;e_N\right)$ is the conjectured equilibrium observer's no-disclosure belief. Suppose by contradiction that $\mathbb{E}\left(\omega_i|ND;e_{N\setminus i}\right)\geqslant\mathbb{E}\left(\omega_i|ND;e_N\right)=\omega_i^{ND}$. Because effort is purely self-improving, we have $\hat{\mathcal{V}}_w(\omega_i^{ND})>\mathcal{V}_w(\omega_i^{ND})$ and therefore

$$\mathbb{E}\left(\omega_i|ND;e_{N\setminus i}\right)<\frac{\sum_{\Omega_{N\setminus i}}\left[\alpha(w)\int_0^1\hat{\mathcal{V}}_w^{-1}(q)dq+\beta(w)\int_0^{\mathcal{V}_w(\omega_i^{ND})}\hat{\mathcal{V}}_w^{-1}(q)dq\right]\eta(w)}{\sum_{\Omega_{N\setminus i}}\left[\alpha(w)+\beta(w)\mathcal{V}_w(\omega_i^{ND})\right]\eta(w)},\qquad(13)$$

where we used the fact that $\mathbb{E}\left(\omega_i|ND;e_{N\setminus i}\right)\geqslant\omega_i^{ND}$; which implies that the right-hand side of (13) can be reached by removing "worse than average" realizations of ω_i from the average computed in (12). But also note that, because $\hat{\mathcal{V}}_w\succ_{FOS}\mathcal{V}_w$ for every $w\in\Omega_{N\setminus i}$, we have $\mathcal{V}_w^{-1}(q)\geqslant\hat{\mathcal{V}}_w^{-1}(q)$ for every $q\in[0,1]$. And consequently

$$\frac{\sum_{\Omega_{N\setminus i}}\left[\alpha(w)\int_0^1\hat{\mathcal{V}}_w^{-1}(q)dq+\beta(w)\int_0^{\mathcal{V}_w(\omega_i^{ND})}\hat{\mathcal{V}}_w^{-1}(q)dq\right]\eta(w)}{\sum_{\Omega_{N\setminus i}}\left[\alpha(w)+\beta(w)\mathcal{V}_w(\omega_i^{ND})\right]\eta(w)}\leqslant$$

$$\frac{\sum_{\Omega_{N\setminus i}}\left[\alpha(w)\int_0^1\mathcal{V}_w^{-1}(q)dq+\beta(w)\int_0^{\mathcal{V}_w(\omega_i^{ND})}\mathcal{V}_w^{-1}(q)dq\right]\eta(w)}{\sum_{\Omega_{N\setminus i}}\left[\alpha(w)+\beta(w)\mathcal{V}_w(\omega_i^{ND})\right]\eta(w)}=\mathbb{E}\left(\omega_i|ND;e_N\right).$$

Now combining this with (13), we have $\mathbb{E}\left(\omega_i|ND;e_{N\setminus i}\right)<\mathbb{E}\left(\omega_i|ND;e_N\right)$, which 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 lemma. □

Using Lemma 8 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, $fe(d)\subset fe(d')$, where d' is the full-disclosure rule. And consequently $FE(D) \subset FE(D')$, where D' is the unilateral disclosure protocol and D is any other deliberation procedure. $\square$

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

Step 1. Fix $D' \neq D$, where D is the unilateral disclosure protocol. As our first step in the proof, we observe (in Lemma 9) 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 9. *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 $i \in N$,*

$$x_i(\omega) = \begin{cases} 0, \text{ if } \omega_i = \underline{\omega}_i \equiv \min(\Omega_i) \\ 1, \text{ otherwise.} \end{cases} \tag{15}$$

Proof of Lemma. Conjecture an equilibrium of the team-disclosure stage in which individual disclosure strategies are as given in (15); and suppose the implied equilibrium team-disclosure strategy is $d(\omega) = D(x(\omega))$. Then we have for each $i \in N$, and each $\epsilon \in (0, 1)$,

$$\omega_i^{ND,\epsilon} = \mathbb{E}^\epsilon\left(\omega_i | ND; e_N\right)$$
$$= \mathbb{P}^\epsilon(\omega_i = \underline{\omega}_i | ND; e_N)\underline{\omega}_i + \mathbb{P}^\epsilon(\omega_i \neq \underline{\omega}_i | ND; e_N)\mathbb{E}^\epsilon\left(\omega_i | ND, \omega_i \neq \underline{\omega}_i; e_N\right). \tag{16}$$

Note that, given the individual disclosure strategies in (15), no-disclosure happens only if at least one team-member $j \in N$ draws their worst possible outcome $\underline{\omega}_j$. But as $\epsilon \to 1$, it must be that for any $i, j \in N$, $\mathbb{P}(\omega_i = \underline{\omega}_i | \omega_j = \underline{\omega}_j) \to 1$. This, along with (16) and the fact that $\mathbb{E}^\epsilon\left(\omega_i | ND, \omega_i \neq \underline{\omega}_i\right)$ is bounded implies that for every $i \in N$,

$$\lim_{\epsilon \to 1} \omega_i^{ND,\epsilon} = \underline{\omega}_i. \tag{17}$$

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 $\omega_i \in \Omega_i \setminus \{\underline{\omega}_i\}$. And therefore the individual disclosure strategy in (15) is individually rational and can be supported as an equilibrium of the team-disclosure stage. □

Step 2. For $\epsilon > \epsilon'$ as given in Lemma 9, in the team-disclosure equilibrium described in the lemma we have for some $i \in N$,

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

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 (15) and D', for all $i \in N$ there exists some $\omega \in \Omega$ with $\omega_i \neq \underline{\omega}_i$ such that $d(\omega) < 1$.

Step 3. Fix $\epsilon > \epsilon'$ as given in Lemma 9 and consider the team-disclosure equilibrium described in the lemma. By equation (17), 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 9. Consequently, if $\epsilon > \bar{\epsilon}$,

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

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

A.10 Proof of Proposition 4

An equilibrium exists in which each team-member i recommends disclosure with certainty if $\omega_i = \omega_{h,i}$ and recommends concealing the outcome with certainty if $\omega_i = \omega_{\ell,i}$. To see, note that in any equilibrium, it must be that $\omega_i^{ND} \in [\omega_{\ell,i}, \omega_{h,i})$, and therefore the proposed strategy is consistent with each agent's "as if pivotal" rationality; and so the proposed strategies constitute an equilibrium. Moreover, as in statement (i), they must constitute the least-informative equilibrium. This is true because there is no equilibrium in which any team-member recommends the concealment of their own high outcomes — because $\omega_i^{ND} < \omega_{\ell,i}$ and therefore concealing individual outcomes would not satisfy individual "as if pivotal" rationality — and the deliberation procedure is monotonic.

The team's disclosure strategy in the least-informative equilibrium is thus given by $d(\omega) = D(\{i : \omega_i = \omega_{h,i}\})$. It is easy to see that, if i is a team-member who can unilaterally choose disclosure, then $D(\{i\}) = 1$, and therefore by monotonicity it must be that $d(\omega) = 1$ for all ω with $\omega_i = \omega_{h,i}$. Therefore all concealed realizations are such that $\omega_i = \omega_{\ell,i}$, and so $\omega_i^{ND} = \omega_{\ell,i}$. If instead i is a team-member who cannot unilaterally choose disclosure, then for $\hat{\omega}$ with $\hat{\omega}_i = \omega h, i$ and $\hat{\omega}_j = \omega_{\ell,j}$, it must be that $d(\hat{\omega}) = D(\{i\}) < 1$; and therefore a "high" realization for team-member i is concealed with positive probability, yielding thus $\omega_i^{ND} > \omega_{\ell,i}$. □

A.11 Proof of Corollary 4

If μ' has more outcome correlation than μ, then

$$d_{\mu'}(\omega_{h,i}) = \sum_{\{\omega:\omega_i=\omega_{h,i}\}} \mu'(\omega)d(\omega) = \mu'_i(\omega_{h,i}) \sum_{\{\omega:\omega_i=\omega_{h,i}\}} \mu'(\omega_{N\setminus i}|\omega_{h,i})d(\omega)$$

$$\leqslant \mu_i(\omega_{h,i}) \sum_{\{\omega:\omega_i=\omega_{h,i}\}} \mu(\omega_{N\setminus i}|\omega_{h,i})d(\omega) = d_\mu(\omega_{h,i}),$$

where the inequality is due to the fact that $\mu'_i(\omega_{h,i}) = \mu_i(\omega_{h,i})$, $\mu'(\omega_{N\setminus i}|\omega_{h,i})$ first-order stochastically dominates $\mu(\omega_{N\setminus i}|\omega_{h,i})$ and $d(\omega)$ is an increasing function of ω. The second inequality in the first statement of the corollary follows analogously, using the fact that $\mu'(\omega_{N\setminus i}|\omega_{\ell,i})$ is first-order stochastically dominated by $\mu(\omega_{N\setminus i}|\omega_{\ell,i})$.

The second statement in the corollary follows directly from the first statement and the fact that μ and μ' have the same marginal distributions. □

A.12 Proof of Lemma 4

Theorem 1 implies that a full-disclosure equilibrium exists, which is also symmetric with $\omega_i^{ND} = \omega_\ell$ for all $i \in N$. If $D(\{i\}) = 1$ for every i, we know that this is the unique symmetric equilibrium. If instead $D(\{i\}) \neq 1$, we want to argue that there exists a unique partial-disclosure symmetric equilibrium. First note that there is a symmetric partial-disclosure equilibrium in which $\omega_i^{ND} > \omega_\ell$ is equal across all team members — this is implied by the least-informative equilibrium described in Proposition 4. Now observe that there is no other symmetric partial-disclosure equilibrium in which $\omega_i^{ND} > \omega_\ell$ for all $i \in N$, for any such no-disclosure beliefs imply that the unique individually rational "as if pivotal" recommendation strategy for each $i \in N$ is to recommend the disclosure of their "high" outcomes and the non-disclosure of their "low" outcomes, both with certainty.

Finally note that in any partial disclosure equilibrium, it must be that $\omega_i^{ND} > \omega_\ell$ for some team-member i, for otherwise the unique non-disclosed outcome must be $\omega = (\omega_\ell, ..., \omega_\ell)$, which then implies that the equilibrium has full-disclosure. But by symmetry, it must then be that $\omega_i^{ND} > \omega_\ell$ for all team-members; and so there exists no symmetric partial-disclosure equilibrium in which $\omega_i^{ND} = \omega_\ell$ for some $i \in N$. □

A.13 Proof of Lemma 5

From Lemmas 3 and 4, 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

$$\bar{c}(D) = \sum_{\Omega} \omega_i \mu(\omega; e_N) - \sum_{\Omega} \omega_i \mu(\omega; e_{N\setminus i}), \text{ if } D(\{i\}) = 1,$$

$$\text{and if } D(\{i\}) < 1, \bar{c}(D) = \sum_{\Omega} \omega_i \mu(\omega; e_N) - \sum_{\Omega} \omega_i \mu(\omega; e_{N\setminus i})$$

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

where d is the team-disclosure strategy in the unique symmetric partial-disclosure equilibrium. Note that, by symmetry, these expressions are independent of the particular choice of $i \in N$.

Because the individual disclosure recommendation strategies are independent of D in the unique symmetric partial-disclosure equilibrium, we know that for each $\omega \in \Omega$, $d(\omega)$ changes continuously with changes in the deliberation procedure D. And 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 $\bar{c}(D)$ is a continuous function of D.

Moreover, a symmetric deliberation procedure D is fully determined by a $n-1$-dimensional vector with each entry between 0 and 1 — remember that $D(X)$ depends only on the cardinality of X, and we fixed $D(\varnothing) = 0$, $D(N) = 1$. Monotonicity further requires that $D(X) \leqslant D(X')$ if $|X| \leqslant |X'|$. The space of deliberation procedures is thus a compact subset of $[0,1]^{n-1}$.

Consequently, there is a deliberation procedure D that maximizes $\bar{c}(D)$, and therefore this procedure is such that $SFE(D') \subseteq SFE(D)$ for every symmetric deliberation procedure D.

□

A.14 Proof of Proposition 5

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 4, 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

$$\mathbb{P}\left(ND | e_{N\setminus i}\right)\left[\mathbb{E}\left(\omega_i | ND; e_{N\setminus i}\right) - \mathbb{E}\left(\omega_i | ND; e_N\right)\right]. \tag{18}$$

We can write the expression for each of the terms in this objective. To that end, we will use the notation $C = 1 - D(1)$. We have:

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

$$\mathbb{E}\left(\omega_1|e_{N\setminus 1}\right) = \omega_\ell + (\omega_h - \omega_\ell)\frac{C\mu\left[(\omega_h,\omega_\ell);e_{N\setminus 1}\right]}{\mu\left[(\omega_\ell,\omega_\ell);e_{N\setminus 1}\right] + C\left\{\mu\left[(\omega_h,\omega_\ell);e_{N\setminus 1}\right] + \mu\left[(\omega_\ell,\omega_h);e_{N\setminus 1}\right]\right\}}$$

$$= \omega_\ell + (\omega_h - \omega_\ell)$$

$$\times\frac{\mu\left[(\omega_h,\omega_\ell);e_{N\setminus 1}\right]}{\mu\left[(\omega_h,\omega_\ell);e_{N\setminus 1}\right] + \mu\left[(\omega_\ell,\omega_h);e_{N\setminus 1}\right]}\frac{C\left\{\mu\left[(\omega_h,\omega_\ell);e_{N\setminus 1}\right] + \mu\left[(\omega_\ell,\omega_h);e_{N\setminus 1}\right]\right\}}{\mu\left[(\omega_\ell,\omega_\ell);e_{N\setminus 1}\right] + C\left\{\mu\left[(\omega_h,\omega_\ell);e_{N\setminus 1}\right] + \mu\left[(\omega_\ell,\omega_h);e_{N\setminus 1}\right]\right\}}$$

$$= \omega_\ell + (\omega_h - \omega_\ell)\frac{\mu\left[(\omega_h,\omega_\ell);e_{N\setminus 1}\right]}{\mu\left[(\omega_h,\omega_\ell);e_{N\setminus 1}\right] + \mu\left[(\omega_\ell,\omega_h);e_{N\setminus 1}\right]}\frac{C}{\frac{\mu\left[(\omega_\ell,\omega_\ell);e_{N\setminus 1}\right]}{\mu\left[(\omega_h,\omega_\ell);e_{N\setminus 1}\right]+\mu\left[(\omega_\ell,\omega_h);e_N\right]} + C}$$

$$= \omega_\ell + (\omega_h - \omega_\ell)\sigma\frac{C}{\rho + C}.$$

Using analogous steps, we have

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

And therefore the objective in (18) can be rewritten as

$$\left[\mu\left[(\omega_\ell,\omega_\ell);e_{N\setminus i}\right] + C\left\{\mu\left[(\omega_h,\omega_\ell);e_{N\setminus i}\right] + \mu\left[(\omega_\ell,\omega_h);e_{N\setminus i}\right]\right\}\right]$$
$$\times\left[\omega_\ell + (\omega_h - \omega_\ell)\sigma\frac{C}{\rho + C} - \omega_\ell - (\omega_h - \omega_\ell)\bar{\sigma}\frac{C}{\bar{\rho} + C}\right] =$$

$$\left[\mu\left[(\omega_\ell,\omega_\ell);e_{N\setminus i}\right] + C\left\{\mu\left[(\omega_h,\omega_\ell);e_{N\setminus i}\right] + \mu\left[(\omega_\ell,\omega_h);e_{N\setminus i}\right]\right\}\right](\omega_h - \omega_\ell)\left[\sigma\frac{C}{\rho + C} - \bar{\sigma}\frac{C}{\bar{\rho} + C}\right],$$

which is proportional to

$$\left[\frac{\mu\left[(\omega_\ell,\omega_\ell);e_{N\setminus i}\right]}{\mu\left[(\omega_h,\omega_\ell);e_{N\setminus i}\right] + \mu\left[(\omega_\ell,\omega_h);e_{N\setminus i}\right]} + C\right]\left[\sigma\frac{C}{\rho + C} - \bar{\sigma}\frac{C}{\bar{\rho} + C}\right],$$

$$=(\rho + C)\left[\sigma\frac{C}{\rho + C} - \bar{\sigma}\frac{C}{\bar{\rho} + C}\right] = \left[\sigma - \frac{\rho + C}{\bar{\rho} + C}\bar{\sigma}\right]C \equiv \Psi(C). \tag{19}$$

We now want to maximize the objective in (19) with respect to C. We consider two cases.

Case 1. $\rho > \bar{\rho}$. First, we verify that the objective is strictly convex for all $C > 0$.

$$\Psi'(C) = \left[\sigma - \bar{\sigma}\frac{\rho + C}{\bar{\rho} + C}\right] - C\frac{\bar{\sigma}(\bar{\rho} - \rho)}{(\bar{\rho} + C)^2} \tag{20}$$

$$\Psi''(C) = -\frac{2\bar{\sigma}(\bar{\rho}-\rho)}{(\bar{\rho}+C)^2} + \frac{2(\bar{\rho}+C)C\bar{\sigma}(\bar{\rho}-\rho)}{(\bar{\rho}+C)^4} > 0 \Leftrightarrow \frac{C}{\bar{\rho}+C} < 1, \text{ which holds for all } C > 0.$$

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

Case 2. $\rho < \bar{\rho}$. By the same steps, we know that the objective is strictly concave for all $C > 0$. And therefore

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

From equation (20), we have that

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

which is increasing in σ, and therefore C^* is weakly increasing in σ (or equivalently, D^* is weakly decreasing in σ). We also have that

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

And so C^* is weakly decreasing in ρ.

□

B Additional Results

B.1 Additional Results for Section 5.2.2

As in the proof of Proposition 5, 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[23]

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

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

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

And so

$$\mathbb{E}\left[\omega_i|ND;e_{N\setminus i}\right]=\left[\frac{\rho(1-h_T)}{(1-\rho)h_i\sum_{m=N-K+1}^{N-1}\binom{N-1}{m}(1-h_j)^m h_j^{N-1-m}}+\frac{1}{h_i} + \frac{(1-h_i)\binom{N-1}{N-K}}{h_i\sum_{m=N-K+1}^{N-1}\binom{N-1}{m}\left(\frac{1-h_j}{h_j}\right)^{m-(N-K)}}\right]^{-1}. \quad (21)$$

[23] 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

$$\mathbb{E}\left[\omega_i|ND;e_N\right]=\Bigg[\frac{\bar{\rho}(1-h_T)}{(1-\bar{\rho})\bar{h}\sum_{m=N-K+1}^{N-1}\binom{N-1}{m}(1-\bar{h})^m\bar{h}^{N-1-m}}+\frac{1}{\bar{h}} \\ +\frac{(1-\bar{h})\binom{N-1}{N-K}}{\bar{h}\sum_{m=N-K+1}^{N-1}\binom{N-1}{m}\left(\frac{1-\bar{h}}{\bar{h}}\right)^{m-(N-K)}}\Bigg]^{-1}. \tag{22}$$

By comparing the difference between (21) and (22), we can assess whether protocols with $K>1$ provide more effort incentives than the unilateral protocol (with $K=1$). Results are stated in Proposition 6 below.

Proposition 6. *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 3 — then all symmetric deliberation protocols dominate the unilateral disclosure protocol.

B.1.1 Proof of Proposition 6

In order to show the four statements, it suffices to sign the derivative of (21) 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 (21) that $\mathbb{E}\left(\omega_i|ND\right)$ decreases in ρ, and increases in h_i.

□