Modes of persuasion toward unanimous consent*

Arjada Bardhi[†] Yingni Guo[‡]

October 23, 2017

Abstract

A fully committed sender seeks to sway a collective adoption decision through designing experiments. Voters have correlated payoff states and heterogeneous thresholds of doubt. We characterize the sender-optimal policy under unanimity rule for two persuasion modes. Under general persuasion, evidence presented to each voter depends on all voters' states. The sender makes the most demanding voters indifferent between decisions, while the more lenient voters strictly benefit from persuasion. Under individual persuasion, evidence presented to each voter depends only on her state. The sender designates a subgroup of rubber-stampers, another of fully informed voters, and a third of partially informed voters. The most demanding voters are strategically accorded high-quality information.

JEL: D71, D82, D83, G28, K20, O32

Keywords: information design, collective decision-making, unanimity rule, information guard

A tremendous share of decision-making in economic and political realms is made within collective schemes. We explore a setting in which a sender seeks to get the unanimous approval of a group for a project he promotes. Group members care about different aspects of the project and might disagree on whether the project should be implemented. They might also vary in the loss they incur if the project is of low quality in their respective aspects. The sender designs experiments to

*For helpful comments and suggestions, we thank Thomas Mariotti, two anonymous referees, Eddie Dekel, Harry Di Pei, Georgy Egorov, Jeffrey Ely, Emir Kamenica, Navin Kartik, Alessandro Pavan, Nicola Persico, Bruno Strulovici, Jean Tirole, Asher Wolinsky, and seminar participants and conference audiences at University of Chicago, Columbia University, University of Miami, 2017 UK PolEcon Conference (LSE), 2017 NASM (St Louis), 2017 SAET (Faro), 2017 Social Networks and Information Conference (Johns Hopkins University), Summer School of Econometric Society 2017 (Seoul), UBC Summer Theory Conference, Transatlantic Theory Workshop 2016 (Oxford), 27th International Conference in Game Theory (Stony Brook), and EconCon 2016 (Princeton).

[†]Department of Economics, Northwestern University, abardhi@u.northwestern.edu.

[‡]Department of Economics, Northwestern University, yingni.guo@northwestern.edu.

persuade the members to approve. When deciding as part of a group, individuals understand the informational and payoff interdependencies among their decisions.

Previous literature has focused mostly on the aggregation and acquisition of (costly) information from exogenous sources in collective decision-making. In contrast, our focus is on optimal persuasion of a heterogeneous group by a biased sender who is able to design the information presented to each group member. We aim to understand the optimal information design, the extent to which group decision-making might be susceptible to information manipulation, and the welfare implications of persuasion for each voter under unanimity rule. Moreover, we contrast optimal persuasion under unanimity with that under nonunanimous rules.

Let us briefly discuss two examples captured by the model. Consider first an industry representative that aims to persuade multiple regulators to approve a project. This representative could be a trade association or an industry-wide self-regulatory authority that interacts directly with regulators.[1] Each regulator is concerned about different but correlated aspects of the project. Typically a successful approval entails the endorsement of all regulators. The representative provides evidence to each regulator by designing informative experiments about the project.

A second example concerns the flow of innovative ideas within organizations. Such ideas are typically born in the R&D department, but they are required to find broad support from other departments, with potentially varied interests, before implementation. The R&D department provides tests to persuade them, and may vary these tests to fit the particular concerns of the department being addressed.

The sender seeks to maximize the ex-ante probability that the project is approved by the entire group.[2] He establishes and commits to an institutionalized standard for the amount of information to be provided to each voter. Modifying institutional standards on an ad hoc basis is costly and difficult due to legal constraints. At the time of the design the sender is uncertain about the quality of projects to be evaluated using this standard.[3]

[1]In the context of the U.S. financial industry, such an industry representative is the Financial Industry Regulatory Authority, the private self-regulatory authority created by the industry and serving its internal needs. FINRA provides information to different regulatory agencies such as the Securities and Exchange Commission (SEC), the Consumer Financial Protection Bureau (CFPB), the Federal Deposit Insurance Corporation (FDIC) etc. It arguably has wide authority in determining the precision of the evidence presented to the regulators (McCarty (2015)).

[2]An example in which majority (rather than unanimous) approval suffices is that of a lobbyist who persuades multiple legislators but only needs a share of them to support the cause.

[3]There are two sources of commitment in the examples we present. R&D units are naturally uninformed about the quality of their innovation at the time of test design. In contrast, in regulation, the industry representative commits on behalf of the entire industry to certain guidelines of information disclosure for any projects to be presented to regulators in the future. In particular, FINRA has a consolidated rulebook on disclosure rules and standards; for an example, see

In both examples, the sender is a key source of information for the receivers. In complex policy environments, the regulators are highly dependent on the industry for expertise and knowledge on how to evaluate the project under investigation.[4] Within innovative organizations, often R&D units exclusively have the required expertise and background information to test the quality of their innovations. They occupy the superior ground of designing tests to persuade other organizational units, which lack the expertise to find independent informational sources.[5]

We take an information design approach in analyzing how the sender optimally persuades a unanimity-seeking group. We consider two scenarios: general persuasion, in which evidence presented to each voter depends on the aspects of all voters; and individual persuasion, in which evidence presented to each voter depends only on her own aspect. We show that the form of the optimal policy as well as the welfare distribution among voters differ drastically under these two persuasion modes.

The two modes describe natural forms of evidence presentation. Within the context of regulation, there is an ongoing debate about the relative effectiveness of comprehensive and targeted evidence to different regulators (Harris and Firestone 2014). General (individual) persuasion is akin to comprehensive (targeted) evidence. Individual persuasion is more natural in regulatory contexts with greater independence and more clearly defined areas of authority across regulators, while general persuasion naturally describes contexts in which each regulator obtains varied evidence on many aspects of a project. Our analysis of the two modes sheds light on the implications of different forms of evidence presentation for optimal information design by the sender and the welfare of different regulators.

Our model features a sender and n voters. A voter's preference is characterized by her binary payoff state. Her payoff from the project is positive if her state is high and negative if the state is low. We assume that the distribution of the voters' payoff states is affiliated and symmetric. With perfectly correlated states, voters agree about the right decision if their states are commonly known. Away from this extreme case, they might disagree about the right decision even if the realized states are commonly known. The magnitude of the loss suffered from approval by a low-state voter differs across voters. We interpret these varying magnitudes as

FINRA's Regulatory Notice 10-41 (http://www.finra.org/industry/notices/10-41).

[4]See McCarty (2015), Omarova (2011), Woodage (2012). Also, McCarty (2015) and McCarty, Poole, and Rosenthal (2013) have argued that within the financial sector in the United States, federal agencies such as SEC and FDIC are demonstrably reliant on the expertise and information provided by FINRA.

[5]Evaluative authorities might attempt to enforce some minimal standards of disclosure on the sender (e.g. Sarbanes-Oxley Act in the finance sector). Our analysis can be interpreted as the design problem after the sender fulfills the minimal requirement imposed by the authorities.

heterogeneous thresholds of doubt; the higher the threshold, the more demanding the voter is. The thresholds and the distribution over state profiles are commonly known by all. Ex ante, neither the voters nor the sender know the realized state profile.

Under general persuasion, the sender designs a mapping from the set of state profiles to the set of distributions over signal profiles. Two features are worth emphasizing: the distribution over the signal profiles is conditioned on the states of all voters; and the signals across voters can be correlated. Under individual persuasion, each voter's signal distribution depends only on her own state. For each voter the sender designs a mapping from the two possible payoff states to the set of distributions over this voter's signal space. We interpret general persuasion as one grand experiment and individual persuasion as voter-specific targeted experiments. Under each mode, we assume that the designed policies are public information, based on the observation that in regulatory settings standards of information are public knowledge either by legal or other institutional requirements. Furthermore, we assume that each voter observes her signal privately. This assumption is based on the observation that confidential information communicated to a specific regulator and pertaining to a specific project is not observable by all parties.[6] Yet, the results remain valid even if all signals were publicly observed.

Under general persuasion, all players are concerned only with the set of state profiles in which all voters receive a recommendation to approve. In the optimal policy the sender chooses a group of the most demanding voters such that they are made indifferent between approval and rejection whenever the sender recommends approval. The more lenient voters obtain a positive payoff merely due to the presence of the more demanding voters. In the extreme case of perfectly correlated states, this group consists of only the most demanding voter. The sender achieves the same payoff as if he faced this voter alone.

A short detour considers the case in which the sender is required to draw the signals independently across voters, conditional on the entire state profile, which we refer to as *independent general* policies. We show that any general policy can be replicated with an independent general policy for unanimity rule.

Under individual persuasion, each voter learns about her state directly from her policy and indirectly from the conjectured decisions of others. With independent states, no voter learns payoff-relevant information from the approval of others.

[6]The assumption of private observability of signals is motivated by the fact that regulatory agencies often face legal and bureaucratic obstacles in sharing information smoothly with each other. Moreover, the principle of strict independence of different regulators is often used as justification for the lack of information sharing.

Hence the optimal policy consists of the single-voter policy for each voter. In contrast, for perfectly correlated states, the sender persuades the most demanding voter as if he needed only her approval, while all other voters always approve.

When the states are imperfectly correlated, each voter receives the sender's recommendation to approve with certainty if her payoff state is high. Intuitively, a higher probability of approval by a high-state voter benefits the sender while also boosting the beliefs of all other voters about their own states. We show that, when a voter's state is low, the probability that she is asked to approve decreases in her threshold of doubt. Thus, the optimal individual policy provides more precise information to more demanding voters.

The optimal individual policy divides the voters into at most three subgroups: the most lenient voters who rubber-stamp the project, the most demanding voters who learn their states fully, and an intermediate subgroup who are partially-informed. Interestingly, the sender does not persuade all voters to approve as frequently as possible, contrary to the case with only one voter. For moderate correlation of states, the most demanding voter(s) learn their states fully and reject for sure when their state is low. Full revelation of the individual state is more informative than what is necessary to persuade the strictest voter(s), but it allows the sender to persuade other voters more effectively. The most demanding and the least demanding voters might obtain a strictly positive expected payoff. For the former, the payoff is due to the information externality they generate for others by acting as *information guards* for the collective decision. For the latter, the positive expected payoff is due to their willingness to rubber-stamp other voters' informed decisions. The intermediate voters obtain a zero expected payoff.

Under either persuasion mode, the sender prefers smaller to larger groups. Also, he weakly prefers general to individual persuasion: any approval probability attained by an individual policy is also achieved through a general policy. The sender's payoffs across the two modes coincide when the states are sufficiently correlated. The most demanding voter weakly prefers individual persuasion, while the rest of the voters might disagree on the preferred persuasion mode.

When moving away from unanimity rule, the results change drastically. For nonunanimous voting rules, the sender achieves a payoff of one under general and independent general persuasion. The project is approved with certainty, so there is no meaningful check on the adoption decision by the voters.[7] In contrast, individual persuasion cannot achieve a certain approval. The voters unambiguously

[7] For any certain-approval policy, there is a nearby policy such that each voter is pivotal with positive probability and strictly prefers to follow an approval recommendation.

prefer individual persuasion to general persuasion because the former allows them to partially discriminate between favorable and unfavorable projects.

The rest of this section discusses the related literature. Section 1 presents the formal model. Sections 2 through 4 analyze the two persuasion modes. Section 5 compares and contrasts these modes. Section 6 briefly discusses two extensions. In particular, subsection 6.2 characterizes sender-optimal persuasion for nonunanimous rules. We conclude and discuss directions for future research in section 7.

Related literature. This paper is immediately related to the literature on persuasion. Rayo and Segal (2010) and Kamenica and Gentzkow (2011) study optimal persuasion between a sender and a single receiver.[8] We study the information design problem of a sender who persuades a group of receivers. Bergemann and Morris (2016a, 2016b), Taneva (2014), Mathevet, Perego, and Taneva (2016), and Bergemann, Heumann, and Morris (2015) also focus on information design with multiple receivers in non-voting contexts. In our setting, voters interact with the sender without any prior private information. The incentive-compatibility constraints for general persuasion characterize the entire set of Bayes correlated equilibria (BCE), while those for individual persuasion characterize a subset of the set of BCE. We identify the sender-optimal BCE within these two sets. In contrast to general persuasion, once attention is restricted to individual persuasion the sender's problem is no longer a linear program.

More specifically, our paper is closely related to the recent literature on persuading voters. Alonso and Câmara (2016b) explore general persuasion when the sender is restricted to public persuasion. We focus on private persuasion. The differences are threefold. Firstly, we show in subsection 6.1 that whether persuasion is private or public is inconsequential under unanimity rule. Hence, under unanimity our general persuasion setting is a special case of Alonso and Câmara. We strengthen their implications for unanimity by characterizing the general persuasion solution in more detail for a broad class of state distributions. Our main result for general persuasion and unanimity is not implied by their analysis.[9] Secondly, we also study individual persuasion under unanimity to examine how the optimal policy changes when the evidence presented to each voter depends only on this voter's state. Lastly, when the voting rule is nonunanimous, we characterize the optimal policy under private persuasion which is drastically different from the optimal policy that Alonso and

[8] More broadly, the paper is related to the literature on communicating information through cheap talk (Crawford and Sobel (1982)), disclosure of hard evidence (see Milgrom (2008) for a survey), and verification (Glazer and Rubinstein (2004)).

[9] In particular, it is not a special case of their proposition 3 which assumes that all voters rank states in the same order. In our setting, voters might disagree even if the state profile were commonly known.

Câmara identify for public persuasion.

Schnakenberg (2015) shows that the sender can achieve certain approval through public persuasion under certain prior distributions if and only if the voting rule is noncollegial. The unanimity rule, which we focus on, is collegial since the approval of all voters is required. Wang (2015) and Chan, Gupta, Li, and Wang (2016) focus on persuading voters who agree under complete information but have different thresholds. In contrast, we examine an environment in which voters might have heterogeneous preferences even under complete information. Moreover, Chan, Gupta, Li, and Wang (2016) characterize the optimal design among information structures that rely on minimal winning coalitions. Arieli and Babichenko (2016) study the optimal group persuasion by a sender who promotes a product. Each receiver makes her own adoption decision so, unlike our setting, there is no payoff externality among receivers.

Another closely related paper is Caillaud and Tirole (2007). In the language of our paper, they also consider individual persuasion. Their setting differs from ours in two aspects: (1) the sender can either reveal or hide a voter's state perfectly; (2) a voter pays a cost to investigate the evidence if provided by the sender. Due to the cost, only voters with moderate beliefs find investigation worthwhile. They show that the sender optimally provides information to a moderate voter so that a more pessimistic voter, who is not willing to investigate if alone, agrees to rubber-stamp the other's approval. Both their analysis and ours rely on the observation that each voter learns about her state from the decisions of others. Yet, we are interested in how the sender adjusts the information precision for each individual voter. We show that the sender provides better information to the most demanding voters in order to convince the more lenient ones to follow suit. More surprisingly, he may find it optimal to fully reveal their states to the most demanding voters.[10]

Our paper also relates to a large literature on information aggregation in collective decisions with exogenous private information, following Austen-Smith and Banks (1996) and Feddersen and Pesendorfer (1996, 1997), as well as on information acquisition in voting games. Li (2001) allows the voters to choose the precision of their signals through costly effort. He argues that groups might choose to commit to more stringent decision-making standards than what is ex-post optimal so as to avoid free riding at the information-acquisition stage. Persico (2004) considers the

[10]Proposition 6 of Caillaud and Tirole shows that the sender might do better facing two randomly drawn voters than one. The cost of investigation is the key. A single pessimistic voter never investigates or approves. Adding a moderate voter who is willing to investigate might induce the pessimistic voter to rubber-stamp. In our environment, there is no cost and the sender can choose the information precision, so more voters always hurt him.

optimal design of a committee, both in terms of its size and its threshold voting rule, so as to incentivize private acquisition of information. Gerardi and Yariv (2008) and Gershkov and Szentes (2009) look at a broader class of mechanisms that incentivize costly information acquisition within a committee. Our focus, in contrast, is on the sender's design of the information structure so as to influence the group decision.

1 Model

Players and payoff states. We consider a communication game between a sender (he) and n voters $\{R_i\}_{i=1}^n$ (she). The voters collectively decide whether to adopt a project promoted by the sender. The payoff of each voter from adopting the project depends on her individual payoff state. In particular, R_i's payoff state, denoted by $\theta_i \in \{H, L\}$, can be either high or low. Let $\theta = (\theta_1, ..., \theta_n)$ denote the state profile of the group, and let $\Theta := \{H, L\}^n$ denote the set of all such state profiles.

Before the game begins, nature randomly draws the state profile θ according to a distribution f. The realized state profile is initially unobservable to all players and f is common knowledge. Throughout our analysis, we assume that the random variables $(\theta_1, ..., \theta_n)$ are exchangeable, in the sense that for every θ and for every permutation ρ of the set $\{1, 2, ..., n\}$, the following holds:

$$f(\theta_1, ..., \theta_n) = f\left(\theta_{\rho(1)}, ..., \theta_{\rho(n)}\right).$$

The analysis of individual persuasion in section 4 and 6.2 also assumes that the voters' states are affiliated. For any two state profiles $\theta, \theta' \in \Theta$, let $\theta \vee \theta'$ and $\theta \wedge \theta'$ denote the componentwise maximum and minimum state profiles respectively.[11] Affiliation of states requires that, for any $\theta, \theta' \in \Theta$,

$$f(\theta \vee \theta')f(\theta \wedge \theta') \geq f(\theta)f(\theta').$$

Let θ^H and θ^L be the state profiles such that $\theta_i^H = H$ for all i and $\theta_i^L = L$ for all i, respectively. The voters' states are perfectly correlated if and only if $f(\theta^H)+f(\theta^L) = 1$. More generally, for $f(\theta^H) + f(\theta^L) < 1$, the states are imperfectly correlated.

Decisions and payoffs. The sender designs an information policy which generates individual signals about the realized state profile. At the time of the design, the sender is uninformed of the realized state profile and fully commits to the chosen policy. The signal intended for R_i is observed only by her.[12]

[11] Formally, $\theta \vee \theta' := (\max\{\theta_1, \theta_1'\}, ..., \max\{\theta_n, \theta_n'\})$ and $\theta \wedge \theta' := (\min\{\theta_1, \theta_1'\}, ..., \min\{\theta_n, \theta_n'\})$.

[12] Under unanimity, it makes no difference whether the signal profile is public or private, as we show in subsection 6.1. Each R_i behaves as if all other voters had received signals that induce them to approve. The private-observability assumption is crucial when we discuss nonunanimous rules in subsection 6.2.

After observing their signals, voters simultaneously decide whether to approve the project. We let $d_i \in \{0, 1\}$ represent R_i's approval decision, where $d_i = 1$ denotes approval. Analogously, $d \in \{0, 1\}$ denotes the collective adoption decision. Under unanimous consent, the project is adopted ($d = 1$) if and only if $d_i = 1$ for every i.

The sender prefers approval of the project regardless of the realized θ: his payoff is normalized to one if the project is approved and zero otherwise. The payoff of R_i depends only on her own state θ_i. The project yields a payoff of one to R_i if $\theta_i = H$ and $-\ell_i < 0$ if $\theta_i = L$. Here, ℓ_i captures the magnitude of the loss incurred by R_i when the project is adopted and R_i's state is L. If the project is rejected, the payoffs of all voters are normalized to zero. Without loss, we assume that no two voters are identical in their thresholds: $\ell_i \neq \ell_j$ for $i \neq j$. For notational convenience, the voters are indexed in increasing order of leniency:

$$\ell_i > \ell_{i+1} \text{ for all } i \in \{1, ..., n-1\}.$$

Each R_i prefers adoption if and only if her state is H. Let $\Theta_i^H := \{\theta \in \Theta : \theta_i = H\}$ and $\Theta_i^L := \{\theta \in \Theta : \theta_i = L\}$ be the set of state profiles with R_i's state being H and L respectively. Therefore, Θ_i^H and Θ_i^L contain R_i's favorable and unfavorable state profiles, respectively. For any i, $\Theta_i^H \cup \Theta_i^L = \Theta$, and $\Theta_i^H \cap \Theta_i^L = \emptyset$.

Voter R_i's prior belief of her state being H is $\sum_{\theta \in \Theta_i^H} f(\theta)$. Due to exchangeability of f, all voters share the same prior belief of their state being H.[13] We focus on parameter values for which none of the voters prefers approval under the prior belief:

Assumption 1. *For any $i \in \{1, ..., n\}$, R_i strictly prefers to reject the project under the prior belief, i.e.:*

$$\ell_i > \frac{\sum_{\theta \in \Theta_i^H} f(\theta)}{\sum_{\theta \in \Theta_i^L} f(\theta)}.$$

We make assumption 1 to simplify exposition. It is without loss of generality. We show in online appendix B.2 that, if some voters prefer to approve ex ante, then the sender designs the optimal policy for those who are reluctant to approve. Those who prefer to approve ex ante are always willing to rubber-stamp.

Modes of persuasion. Let S_i denote the signal space for R_i and, correspondingly, $\prod_i S_i$ the space of signal profiles for the entire group. The game has three stages. The sender first commits to an information policy. Subsequently, nature draws θ according to f and then draws the signals $(s_i)_{i=1}^n$ according to the chosen information policy. Finally, each voter receives her signal and chooses d_i.

[13] By assuming exchangeability, we can focus on the impact of different thresholds on the optimal information policy.

We consider three major classes of information policies: 1) general policies, 2) independent general policies, and 3) individual policies. Formally, a *general* policy is a mapping from the set of state profiles to the set of probability distributions over the signal space:

$$\pi : \Theta \to \Delta\left(\prod_i S_i\right).$$

For each state profile θ, it specifies a distribution over all possible signal profiles, so the signals sent to the voters could be correlated conditional on the state profile. We let Π^G denote the set of all such policies.

An *independent general* policy specifies, for each voter R_i, a mapping from the set of state profiles to the set of probability distributions over R_i's signal space:

$$\pi_i : \Theta \to \Delta\left(S_i\right).$$

As in a general policy, each voter's signal distribution depends on the entire state profile. But unlike in a general policy, conditional on θ the signals across voters are independently drawn. Let Π^{IG} denote the set of all independent general policies, with $(\pi_i)_{i=1}^n$ being a typical element of it.

An *individual* policy specifies, for each voter R_i, a mapping from R_i's state space to the set of probability distributions over S_i:

$$\pi_i : \{H, L\} \to \Delta(S_i).$$

Unlike in a general or independent general policy, each voter's signal distribution depends only on her own state. Let Π^I denote the set of all individual policies, with $(\pi_i)_{i=1}^n$ being a typical element.

The information policy adopted by the sender determines the information structure of the voting game played among voters. Let Π denote the set of policies available to the sender. If the sender is allowed to use general policies, then $\Pi = \Pi^G$. If the sender is constrained to independent general policies or individual policies, then $\Pi = \Pi^{IG}$ or $\Pi = \Pi^I$ respectively.[14] The strategy σ_i determines the probability R_i approves, given the chosen policy and the realized signal:

$$\sigma_i : \Pi \times S_i \to [0, 1].$$

Without loss, we focus on *direct obedient policies*, for which (i) S_i coincides with the action space $\{0, 1\}$, and (ii) each voter receives action recommendations, with which she complies.[15] We use $\hat{d}_i \in \{0, 1\}$ to represent the sender's recommendation

[14] By definition, any individual policy can be replicated by an independent general policy. Any independent general policy can be replicated by a general policy.

[15] The first part of online appendix B.3 provides a proof that the restriction to direct obedient policies is without loss of generality.

Modes	Definition	Notation
General	$\pi : \Theta \to \Delta(\{0,1\}^n)$	$(\pi(\cdot\|\theta))_{\theta\in\Theta}$
Independent general	$\pi_i : \Theta \to \Delta(\{0,1\})\,,\ \forall i$	$(\pi_i(\theta))_{\theta\in\Theta}\,,\ \forall i$
Individual	$\pi_i : \{H,L\} \to \Delta(\{0,1\})\,,\ \forall i$	$(\pi_i(H),\pi_i(L))\,,\ \forall i$

Table 1: Modes of persuasion

to R_i, and $\hat{d} = (\hat{d}_i)_i \in \{0,1\}^n$ to represent the profile of action recommendations.

A direct general policy specifies a distribution over action-recommendation profiles as a function of the realized state profile, i.e., it specifies $\pi(\cdot|\theta) \in \Delta(\{0,1\}^n)$ for each θ. A direct independent general policy specifies for each R_i and for each θ the probability $\pi_i(\theta)$ with which R_i is recommended to approve, since the action space is binary. For individual persuasion, any direct policy specifies $(\pi_i(H), \pi_i(L)) \in [0,1]^2$ for each R_i, where $\pi_i(\theta_i)$ is the probability that the sender recommends approval to R_i when her state is θ_i. Table 1 summarizes the definition and notation for these three different modes of persuasion.

Refinement. We allow for any policy that is the limit of a sequence of direct obedient policies with full support. The full-support requirement demands that for any state profile, all possible recommendations are sent with positive probabilities.[16] We impose this refinement so that along the sequence (i) no recommendation is off the equilibrium path; and (ii) a voter is always pivotal with positive probability.[17] For each persuasion mode, we solve for the sender-optimal policy.

2 General persuasion

2.1 General formulation

In the regulatory process for complex industries, the most general form that the evidence presented to the regulators can take is as an experiment that generates correlated action recommendations conditional on the realized states of all regulators. The structure of such an experiment is formally captured by a *general policy*.

Recall that $\pi(\hat{d}|\theta)$ is the probability that the recommendation profile $\hat{d}$ is sent, given the state profile θ. If R_i rejects, the project is definitively rejected. If R_i approves, the project is collectively approved if and only if all other voters approve as well; this is the only event in which R_i's decision matters. Let $\hat{d}^a$ be the recommen-

[16]The second part of online appendix B.3 provides a formal definition of full-support policies.

[17]For unanimity rule, we can easily construct a sequence of full-support direct obedient policies that approaches the optimal policy. Therefore, we do not invoke this refinement explicitly in the main discussion. However, when we discuss nonunanimous rules in subsection 6.2, we construct explicitly the sequence of full-support policies that approaches the optimal policy.

dation under which all voters receive a recommendation to approve. We refer to $\hat{d}^a$ as the unanimous (approval) recommendation. Voter R_i obeys a recommendation to approve if

$$\sum_{\theta \in \Theta_i^H} f(\theta)\pi(\hat{d}^a|\theta) - \ell_i \sum_{\theta \in \Theta_i^L} f(\theta)\pi(\hat{d}^a|\theta) \geq 0. \tag{IC^a-i}$$

Let $\hat{d}^{r,i}$ be the recommendation under which all voters except R_i receive a recommendation to approve, i.e., $\hat{d}_i^{r,i} = 0$ and $\hat{d}_j^{r,i} = 1$ for all $j \neq i$. Voter R_i obeys a recommendation to reject if

$$\sum_{\theta \in \Theta_i^H} f(\theta)\pi(\hat{d}^{r,i}|\theta) - \ell_i \sum_{\theta \in \Theta_i^L} f(\theta)\pi(\hat{d}^{r,i}|\theta) \leq 0. \tag{IC^r-i}$$

The sender chooses $(\pi(\cdot|\theta))_{\theta \in \Theta}$ so as to maximize the probability of the project being approved, $\sum_{\theta \in \Theta} f(\theta)\pi(\hat{d}^a|\theta)$, subject to (i) approval and rejection IC constraints, and (ii) the following feasibility constraints: $\pi(\hat{d}|\theta) \geq 0$, for all $\hat{d}, \theta$, and $\sum_{\hat{d} \in \{0,1\}^n} \pi(\hat{d}|\theta) = 1$ for all θ.

We first analyze the relaxed problem in which IC^r constraints are ignored:

$$\max_{(\pi(\hat{d}^a|\theta))_{\theta \in \Theta}} \sum_{\theta \in \Theta} f(\theta)\pi(\hat{d}^a|\theta), \tag{1}$$
$$\text{subject to} \quad (\text{IC}^a\text{-}i)_i \quad \text{and} \quad \pi(\hat{d}^a|\theta) \in [0,1],\ \forall \theta \in \Theta.$$

A solution to this relaxed problem specifies, for each θ, only the probability that all voters are recommended to approve, i.e., $(\pi(\hat{d}^a|\theta))_{\theta \in \Theta}$. For any such solution, we can easily construct the probabilities $(\pi(\hat{d}^{r,i}|\theta))_{i,\theta}$ so that the IC^r constraints are satisfied as well.[18] Therefore, focusing on the relaxed problem is without loss.

2.2 Characterization of the optimal policy

Each R_i learns about her state from the relative frequency with which the unanimous recommendation $\hat{d}^a$ is generated in Θ_i^H rather than in Θ_i^L. The posterior belief that R_i holds about her state being H conditional on $\hat{d}^a$ having been drawn is:

$$\Pr(\theta_i = H|\hat{d}^a) = \frac{\sum_{\theta \in \Theta_i^H} f(\theta)\pi(\hat{d}^a|\theta)}{\sum_{\theta \in \Theta} f(\theta)\pi(\hat{d}^a|\theta)}.$$

Each IC^a-i can be rewritten in terms of this posterior belief as:

$$\Pr(\theta_i = H|\hat{d}^a) \geq \frac{\ell_i}{1+\ell_i}.$$

[18] Lemma 2.2 establishes that for any policy that satisfies the approval IC constraints and for any R_i there exists $\theta' \in \Theta_i^L$ such that $\pi(\hat{d}^a|\theta') < 1$. The sender can specify $\pi(\hat{d}^{r,i}|\theta') = \varepsilon$ and $\pi(\hat{d}^{r,i}|\theta) = 0$ for all $\theta \neq \theta'$. Such a specification guarantees that IC^r-i holds for each R_i.

This posterior belief has to be sufficiently high for R_i to obey an approval recommendation. The cutoff value for this posterior belief increases in the threshold of doubt: naturally, the larger the loss a voter experiences if her state is L, the higher the posterior belief about $\theta_i = H$ needed in order for this voter to prefer approval.

We first examine the optimal policy for perfectly correlated states. Only two state profiles, θ^H and θ^L, are possible to realize.

Proposition 2.1 (Perfect correlation)**.** *Suppose the voters' states are perfectly correlated. The unique optimal policy, for which only IC^a-1 binds, is given by:*

$$\pi(\hat{d}^a|\theta^H) = 1, \quad \pi(\hat{d}^a|\theta^L) = \frac{f(\theta^H)}{f(\theta^L)}\frac{1}{\ell_1}.$$

The unanimous recommendation is sent with certainty given θ^H. The probability of unanimous approval in θ^L is determined only by the threshold of the most demanding voter R_1. Due to perfect correlation, all voters share the same posterior belief about their respective states being H. The highest cutoff on this posterior belief is imposed by R_1. Thus, the sender provides sufficiently accurate recommendations so as to leave R_1 indifferent between approval and rejection. Being more lenient than R_1, all other voters receive a strictly positive expected payoff.

We now generalize our discussion to imperfectly correlated states. Our first observation is that the sender recommends approval with certainty to all voters when all their respective states are high. The intuition is straightforward: increasing $\pi(\hat{d}^a|\theta^H)$ strengthens the posterior belief $\Pr(\theta_i = H|\hat{d}^a)$ of each voter R_i, while also strictly improving the probability of a collective approval.

Lemma 2.1 (Certain approval for θ^H)**.** *The sender recommends with certainty that all voters approve when every voter's state is high, i.e., $\pi(\hat{d}^a|\theta^H) = 1$.*

We next show that, for any R_i, an optimal policy does not set $\pi(\hat{d}^a|\theta) = 1$ for all $\theta \in \Theta_i^L$. If there existed such a voter who knows that the group receives $\hat{d}^a$ for sure whenever her state is L, her posterior belief conditional on $\hat{d}^a$ would be lower than the prior belief. She would not be willing to approve given assumption 1.

On the other hand, there does not exist a voter who, given that $\hat{d}^a$ is sent, is fully confident that her state is H. Put differently, every voter mistakenly obeys a unanimous recommendation for a project for which her state is low with positive probability. If indeed some voter R_i learned her state fully, her IC^a-i would be slack. Moreover, full revelation would require that $\pi(\hat{d}^a|\theta) = 0$ for $\theta \in \Theta_i^L$ such that $\theta_j = H$ for all $j \neq i$. By increasing the frequency with which $\hat{d}^a$ is generated in this state profile, the sender improves his payoff while still satisfying the previously slack IC^a-i and strictly relaxing all other IC^a constraints.

Lemma 2.2 (No certain approval or rejection for Θ_i^L)**.** *For each i, there exist $\theta, \theta' \in \Theta_i^L$ such that $\pi(\hat{d}^a|\theta) < 1$ and $\pi(\hat{d}^a|\theta') > 0$.*

A natural next question concerns the pattern of binding and slack IC^a constraints across voters. Under the unanimous rule with an outside option normalized at zero, a binding IC^a constraint implies that the corresponding voter's expected payoff is exactly zero. Hence, an analysis of the subset of binding IC^a constraints has immediate implications for the welfare of the voters. The following proposition establishes that there exists an index $i' \geq 1$ such that all voters who are more demanding than $R_{i'}$ have binding IC^a constraints and a zero expected payoff in any optimal policy. The only voters who might obtain a positive expected payoff are the most lenient voters in the group.[19]

Proposition 2.2 (The strictest voters' IC^a constraints bind)**.** *Suppose f is exchangeable. In any optimal policy, a subgroup of the strictest voters' IC^a constraints bind, i.e., IC^a-i binds iff $i \in \{1, ..., i'\}$ for some $i' \geq 1$.*

This proposition holds as long as f is exchangeable. Ex ante, the voters differ only in their thresholds. The proof makes use of the dual problem corresponding to (1). Thinking of each voter's IC^a as a resource constraint, we show that granting positive surplus to a tough voter is more expensive than to a lenient one. Intuitively, the voters with the highest thresholds are the hardest to persuade; hence, the sender provides sufficiently precise information about the strictest voters' states in order to leave them indifferent between approval and rejection.

Let us now briefly touch upon the multiplicity of optimal policies that arises under general persuasion. The dual problem corresponding to (1) identifies the set of binding IC^a constraints. It also identifies the state profiles for which the project is approved or rejected for sure. These conditions pin down the sender's payoff. As a result, he has flexibility in designing how frequently approval is recommended to all voters in other state profiles. For voters with slack IC^a constraints, their expected payoffs vary across different optimal policies. The following example with $n = 3$ and independent payoff states illustrates this multiplicity.

Example 1 (Multiplicity of optimal policies)**.** *We suppose that (ℓ_1, ℓ_2, ℓ_3) equals $(20, 15, \frac{2291}{229})$. Voters' states are independent. The state of each voter is H with probability $9/10$. By examining the dual problem (which is a linear program), we determine that in any optimal policy:*

$$\pi(\hat{d}^a|HHH) = \pi(\hat{d}^a|HHL) = 1, \quad \pi(\hat{d}^a|LLH) = \pi(\hat{d}^a|LLL) = 0.$$

[19] We say that R_i's IC^a constraint binds if the dual variable associated with this constraint is strictly positive.

Moreover, both IC^a-1 and IC^a-2 bind. These pin down the sender's payoff. Subject to the binding IC^a for R_1 and R_2, the feasibility constraints, and IC^a-3, the sender has flexibility in specifying the rest of the unanimous recommendation probabilities. Among the optimal policies, R_3's payoff is the highest if

$$\pi(\hat{d}^a|HLH) = \frac{210}{299},\ \pi(\hat{d}^a|LHH) = \frac{160}{299},\ \pi(\hat{d}^a|HLL) = \pi(\hat{d}^a|LHL) = 0.$$

R_3 receives a strictly positive payoff from this policy. R_3's payoff is the lowest if

$$\pi(\hat{d}^a|HLH) = \frac{47}{69},\ \pi(\hat{d}^a|LHH) = \frac{160}{299},\ \pi(\hat{d}^a|HLL) = \frac{57}{299},\ \pi(\hat{d}^a|LHL) = 0,$$

which grants R_3 a zero expected payoff. □

In this example, the optimal policy in which R_3 receives a zero payoff is Pareto dominated by the optimal policy in which R_3 receives a positive payoff. The sender may very well choose an optimal policy that is also Pareto efficient. The set of optimal and Pareto efficient policies is given by maximizing the weighted sum of voters' payoffs subject to their IC^a constraints and the constraint that the sender obtains her optimal payoff.

Example 1 also illustrates that even when the voters' payoff states are entirely independent, the general persuasion problem faced by the sender is not separable across voters. For example, the probability with which R_3 is recommended to approve the project when her state is L depends on the realized states of the other voters: in any optimal policy in example 1, $\pi(HHL) = 1$ but $\pi(LLL) = 0$.[20]

3 Equivalence of general and independent general policies

In the context of our motivating example, the regulatory process might require the sender to conduct an independent experiment for each regulator. If the sender is allowed to condition the recommendations to each regulator on the entire state profile, the sender designs for each regulator a mapping from the set of state profiles to the set of distributions over this regulator's signal space. Formally, such a profile of independent experiments is an *independent general policy.* We show in this section that under unanimity, for any general policy, there is an independent general policy that achieves the same payoffs for all players.

Recall that $\pi_i(\theta)$ denotes the probability of an approval recommendation made to R_i when the realized state profile is θ. The approval and rejection incentive-compatibility constraints have to be slightly modified for an independent general

[20]This observation stands in sharp contrast to the case of individual persuasion, under which no voter can possibly receive a strictly positive payoff when the states are independent. In section 4, we show that the optimal individual policy exploits the lack of information externalities between voters to set all IC^a-i constraints binding.

policy $(\pi_i)_{i=1}^n$:

$$\sum_{\theta\in\Theta_i^H} f(\theta)\prod_{j=1}^n \pi_j(\theta) - \ell_i \sum_{\theta\in\Theta_i^L} f(\theta)\prod_{j=1}^n \pi_j(\theta) \geq 0, \tag{2}$$

$$\sum_{\theta\in\Theta_i^H} f(\theta)\,(1-\pi_i(\theta))\prod_{j\neq i} \pi_j(\theta) - \ell_i \sum_{\theta\in\Theta_i^L} f(\theta)\,(1-\pi_i(\theta))\prod_{j\neq i} \pi_j(\theta) \leq 0. \tag{3}$$

Similarly, the objective of the sender changes to

$$\sum_{\theta\in\Theta} f(\theta)\prod_{i=1}^n \pi_i(\theta).$$

We argue that a payoff is attainable under independent general persuasion if and only if it is attainable under general persuasion. One direction of this statement is trivial: any independent general policy can be formulated as a general policy. The next propositions shows that the other direction holds as well.

Proposition 3.1 (Equivalence of general/independent general policies)**.** *Under unanimity, the set of attainable payoffs for the sender and the voters is the same under general policies and independent general policies.*

The key assumption for this result is the unanimity rule. The distribution f need not be exchangeable or affiliated. Moreover, the result holds even if each voter's state space is not binary. Under unanimity, the sender cares only about the event in which all voters receive an approval recommendation; so does each voter when she contemplates obeying an approval recommendation. Hence, for each θ, the sender can choose the approval probability for each voter so that the product of all these probabilities equals the unanimous approval probability in a general policy. When we consider nonunanimous rules, the equivalence does not hold any more: In subsection 6.2, we show that the sender is strictly worse off under independent general persuasion when voters' states are perfectly correlated. Moreover, this equivalence result is not expected to hold in other group persuasion games beyond voting games.[21]

4 Individual persuasion

4.1 General formulation

Restrictive regulatory processes might require that a regulator be provided only evidence directly pertaining to her area of interest. Such a requirement is often justified on grounds of protecting the independence of different regulatory agencies

[21]See Arieli and Babichenko (2016).

in their evaluations. Other times law assigns separate and disjoint areas of authority to different regulators: they decide based on evidence pertaining to their area of authority. Targeted experiments allow each regulator to focus her limited resources and have full authority over the evaluation of one aspect in parallel regulatory processes (i.e. processes that involve more than one regulator). Such experiments are formally captured by an *individual policy*.

This section characterizes optimal individual persuasion. We show that a more demanding voter enjoys a more informative policy. The sender essentially divides the group into (at most) three subgroups: (i) the most demanding voters fully learn their states; (ii) the intermediate voters are partially informed; and (iii) the most lenient voters rubber-stamp. We further show that only the extreme voters might obtain a positive payoff: the most demanding voters due to their role as informational guards, and the least demanding voters due to their willingness to rubber-stamp.

Recall that $\pi_i(\theta_i)$ denotes the probability that R_i receives an approval recommendation when her state is θ_i. Let $\Pr(\theta_i = H|R_{-i} \text{ approve})$ denote the probability that $\theta_i = H$, conditional on all voters other than R_i approving:

$$\Pr(\theta_i = H|R_{-i} \text{ approve}) = \frac{\sum_{\theta \in \Theta_i^H} f(\theta) \prod_{j \neq i} \pi_j(\theta_j)}{\sum_{\theta \in \Theta} f(\theta) \prod_{j \neq i} \pi_j(\theta_j)}.$$

R_i's incentive-compatibility constraint when she receives an approval recommendation is:

$$\Pr(\theta_i = H|R_{-i} \text{ approve})\pi_i(H) - \ell_i(1 - \Pr(\theta_i = H|R_{-i} \text{ approve}))\pi_i(L) \geq 0. \quad (\text{IC}^a\text{-}i)$$

The sender maximizes the probability of a collective approval:

$$\sum_{\theta \in \Theta} f(\theta) \prod_{i=1}^{n} \pi_i(\theta_i). \quad (\text{OBJ})$$

We focus on the relaxed problem of maximizing (OBJ), subject to the set of IC^a-i constraints.[22] Focusing on such a relaxed problem is without loss: we show later in lemma 4.1 that recommendation to reject in any optimal policy for the relaxed problem is conclusive news that the voter's own state is low; hence she always obeys a rejection recommendation.

[22]Under optimal individual persuasion, every voter approves with positive probability, i.e., $\pi_i(H) + \pi_i(L) > 0$, $\forall i$. Otherwise, the sender's payoff is zero. The sender could do better by fully revealing her payoff state to each voter. Hence, $\Pr(\theta_i = H|R_{-i} \text{ approve})$ is well defined.

4.2 Characterization of the optimal policy

The IC^a-i constraint, rewritten in the following form, emphasizes the informational externalities among the voters' decisions:

$$\Pr(\theta_i = H | R_{-i} \text{ approve}) \geq \frac{\ell_i \pi_i(L)}{\ell_i \pi_i(L) + \pi_i(H)}.$$

The left-hand side is R_i's belief that her state is high when she conditions on the others' approvals. This belief depends on the policies of all voters other than R_i.[23] Due to affiliation, an increase in $\pi_j(H)$ of another voter R_j boosts the posterior belief of R_i, while an increase in $\pi_j(L)$ makes R_i more pessimistic about her state. The right-hand side depends only on R_i's own policy and her threshold of doubt. It decreases in $\pi_i(H)$ and increases in $\pi_i(L)$ and ℓ_i. The more likely that R_i receives an approval recommendation when her state is high, the easier it is to induce compliance. The more frequently R_i receives an approval recommendation when her state is low or the more demanding R_i is, the more difficult it is to induce compliance. So, an increase in $\pi_i(H)$ relaxes not only IC^a-i but also IC^a-j for all other R_j, while an increase in $\pi_i(L)$ increases the cutoff for R_i and lowers the posterior belief of all other voters, thus tightening all IC^a constraints.

Before approaching the more general problem of imperfectly correlated states, we first solve the polar cases of perfectly correlated and independent states.

Proposition 4.1 (Perfectly correlated or independent states)**.** *Suppose the voters' states are perfectly correlated. Any optimal policy is of the form:*

$$\pi_i(H) = 1, \ \forall i, \quad \textit{and} \quad \prod_{i=1}^{n} \pi_i(L) = \frac{f(\theta^H)}{f(\theta^L)} \frac{1}{\ell_1}.$$

Suppose the states are independent. The policy for each voter is the same as if the sender were facing only this voter:

$$(\pi_i(H), \pi_i(L)) = \left(1, \frac{\sum_{\theta \in \Theta_i^H} f(\theta)}{\sum_{\theta \in \Theta_i^L} f(\theta)} \frac{1}{\ell_i}\right) \ \textit{for all } i.$$

For perfectly correlated states, the project is approved for sure when every voter's state is H, while the probability of approval in state L is chosen so that only IC^a-1 binds. One such optimal policy is the one in which the sender persuades the most demanding R_1 and recommends that all other voters rubber-stamp the decision of R_1.[24] For this policy, $\pi_1(L) = f(\theta^H)/(f(\theta^L)\ell_1)$ and $\pi_i(L) = 1$ for all $i \geq 2$.

For independent states, each voter R_i's posterior belief $\Pr(\theta_i = H | R_{-i} \text{ approve})$ equals her prior belief of her state being H regardless of the other voters' policies.

[23] With slight abuse of the term "policy," we call $(\pi_i(H), \pi_i(L))$ the policy of R_i under individual persuasion.

[24] A voter rubber-stamps if she approves with probability one in both states.

In the absence of information externalities across voters, the sender sets $\pi_i(L)$ as high as IC^a-i allows for each voter R_i.

We next show that for imperfectly correlated states it continues to be the case that a high-state voter obtains an approval recommendation with probability one.

Lemma 4.1 (High states approve for sure)**.** *Suppose that f is exchangeable and affiliated. In any optimal policy, the sender recommends approval to each R_i with probability one when $\theta_i = H$, i.e., $\pi_i(H) = 1$ for every i.*

The main step of the proof relies on the Ahlswede-Daykin inequality, through which we show that increasing $\pi_i(H)$ for R_i relaxes the IC^a constraints for voters other than R_i as well. The interests of the sender and all voters are aligned when it comes to an increase of $\pi_i(H)$ for any R_i: such an increase improves the sender's payoff, makes R_i more compliant with an approval recommendation, and makes all other voters more optimistic given that R_i approves. Thus, lemma 4.1 reduces the problem to simply choosing $(\pi_i(L))_{i=1}^n$.

Lemma 4.2 (At least one IC^a binds)**.** *Suppose that f is exchangeable and affiliated. In any optimal policy, $\pi_i(L) < 1$ for some i. Moreover, at least one voter's IC^a constraint binds, so $\pi_j(L) > 0$ for some j.*

Lemma 4.2 rules out the possibility that all voters approve without any additional information from the policy, as a direct consequence of assumption 1. Moreover, any optimal policy accords a zero expected payoff to at least one voter. If that were not the case, the sender could strictly improve the approval probability by slightly increasing $\pi_i(L)$ for some R_i. Hence, lemma 4.2 rules out the possibility of the optimal policy being fully revealing about all states. Yet, it may well be the case that the optimal policy is fully revealing about the states of some voters. Indeed, we present in example 2 an optimal policy for which $\pi_i(L) = 0$ for some R_i.

Example 2 (Full revelation to some voter)**.** *Let f be: $f(HHH) = 6/25$, $f(HHL) = 1/250$, $f(HLL) = 7/750$, and $f(LLL) = 18/25$. The thresholds are $(\ell_1, \ell_2, \ell_3) = (41, 40, 39)$. Based on lemma 4.1, the sender chooses $(\pi_i(L))_{i=1}^3$ to maximize the probability of unanimous approval subject to IC^a constraints. The solution is:*

$$(\pi_1(L), \pi_2(L), \pi_3(L)) = (0, 0.606, 0.644).$$

The most demanding voter R_1 learns her state fully. The project is never approved when $\theta_1 = L$, so R_1's IC^a is slack. R_1 takes the role of a very accurate veto player: the more lenient voters depend on R_1 to veto a bad project. The sender could recommend the low-state R_1 to approve more frequently, but he optimally chooses to fully reveal θ_1 so that he can persuade R_2 and R_3 more effectively.

Example 2 highlights the crucial role of the most demanding voter(s): they serve as information guards vis-a-vis the more lenient voters who either partially or fully rubber-stamp the project. This example also suggests a monotonicity feature of the optimal policy $(\pi_i(L))_{i=1}^n$ with respect to the threshold: for $\ell_i > \ell_j$, $\pi_i(L) \leq \pi_j(L)$. More demanding voters are less likely to receive an approval recommendation when their states are low. Hence, the approval recommendations made to more demanding voters are more informative of their respective states being H. Any voter is more optimistic about her state being high when taking into account the approval of another more demanding voter than the approval of a more lenient voter. The following proposition establishes this monotonicity property.

Proposition 4.2 (Monotonicity of persuasion)**.** *Suppose that f is exchangeable and affiliated. There exists an optimal policy in which more demanding voters' policies are more informative:*

$$\pi_i(L) \leq \pi_{i+1}(L) \text{ for all } i \in \{1, ..., n-1\}.$$

Moreover, in any optimal policy in which R_i's IC^a constraint binds, those who are more demanding than R_i must have strictly more informative policies, i.e., $\pi_j(L) < \pi_i(L)$ for all $j < i$.

Proposition 4.2 states that the sender essentially divides the group into (at most) three subgroups. The most demanding voters learn their states fully. The intermediate voters are partially manipulated. The most lenient voters rubber-stamp the collective decision.

To prove the first part of proposition 4.2, we show that if there is a pair of voters for which the more lenient voter enjoys a more informative policy, we can swap the individual policies between the two. The new policy remains incentive-compatible. Intuitively, when the stricter voter is compliant with a less informative policy, she continues to comply when assigned the more informative policy. After the swap, the stricter voter's belief that her state is H is weakened when conditioning only on the approval of all other voters, but the more accurate information acquired through her own policy offsets this increased pessimism.[25] After the swap, the more lenient voter is assigned the less informative policy accorded previously to the stricter voter. Since the stricter voter was willing to comply with an approval recommendation from this policy, the more lenient voter is willing to do so as well. Therefore, there always exists an optimal policy such that stricter voters have more informative policies. The second part of proposition 4.2 follows naturally: if IC^a-i

[25] After all, each voter learns about her own state indirectly from the others' approval decisions but she learns about her own state directly from her own policy.

binds for some R_i, any voter who is more demanding than R_i must have a strictly more informative policy. Otherwise, this stricter voter's IC^a is not satisfied.

Based on the monotonicity property and the previous examples, a natural conjecture is that any voter with a slack IC^a constraint either fully learns her state or rubber-stamps. This conjecture is not true, as the following example 3 demonstrates. Nonetheless, among those voters who are partially informed (i.e., $\pi_i(L) \in (0,1)$, at most one of them has a slack IC^a. If such a voter exists, she must be the strictest among those voters with partially informative policies. We state this result formally in proposition 4.3 below.

Example 3 (Slack IC^a with interior $\pi_i(L)$). *Let f be the same as in example 2 and the threshold profile be $(31, 30, 24)$. The optimal policy is given by:*

$$(\pi_1(L), \pi_2(L), \pi_3(L)) = (0.003, 0.516, 1) .$$

Only IC^a-1 is slack. The sender provides very precise information to R_1 so as to be able to recommend approval more frequently to R_2 and R_3. Decreasing $\pi_1(L)$ further does not benefit the sender once the recommendation to R_3 becomes fully uninformative, i.e., $\pi_3(L) = 1$. □

Proposition 4.3. *Suppose distribution f is exchangeable and satisfies strict affiliation for any 3-voter subgroup. Among those voters who are partially informed, at most one has a slack IC^a constraint. If such a voter exists, she is the strictest voter among those who are partially informed.*

Taken together, propositions 4.2 and 4.3 establish that only the extreme voters might obtain a positive payoff from persuasion: the most demanding voters due to their role as informational guards, and the least demanding voters due to their willingness to rubber-stamp.

We conclude this section by discussing the proof of proposition 4.3. Suppose, to the contrary, that we can find two partially informed voters with thresholds $\ell_i < \ell_j$ whose IC^a constraints are slack. From proposition 4.2, it is without loss that $\pi_i(L) \geq \pi_j(L)$. Due to the slack IC^a-i and IC^a-j, the sender can slightly increase $\pi_i(L)$ and decrease $\pi_j(L)$ so that, given R_i and R_j's approvals, every other voter's posterior belief of her respective state being H is at least as high as prior to the change. Because $\pi_i(L)$ is greater than $\pi_j(L)$ to start with, the boost in another voter's belief from a lower $\pi_j(L)$ has a stronger effect than the drop in the belief from a higher $\pi_i(L)$. Therefore, only a small decrease in $\pi_j(L)$ is required to offset the change in $\pi_i(L)$. Importantly, this necessary decrease is sufficiently small so that the sender's payoff strictly improves from this policy perturbation.

Hence, at most one partially informed voter has a slack IC^a. The second part of proposition 4.3 strengthens this observation: if some partially informed voter receives a positive payoff, she must be the strictest voter among all those who are partially informed. If the sender ever provides to some voter more precise information than what is required by her IC^a constraint, he prefers to do so with a voter who will necessarily be assigned a low $\pi_i(L)$. Reducing an already low $\pi_i(L)$ generates a stronger optimism boost among other voters.

4.3 When do some voters fully learn their states?

Individual persuasion introduces the possibility that the strictest voter(s) learn their own states fully. This stands in contrast to general persuasion. Lemma 2.2 established that no voter fully learns her state under general persuasion. The following discussion identifies necessary conditions for full revelation to the strictest voters to be optimal under individual persuasion. More importantly, we characterize the parameter region under which full revelation to some voters arise for a subclass of state distributions.

First, the group size must be at least three. We show in online appendix B.4 that when facing two voters, the sender sets $\pi_1(L)$ and $\pi_2(L)$ as high as their IC^a constraints permit; he never fully reveals the state to any voter. When the sender provides more precise information to some voter than what this voter's IC^a constraint requires, the probability that this voter approves the project is reduced. On the other hand, the sender can persuade the other voters to approve in their low state more frequently. This information externality is a public good so the benefit is larger for a larger group. The group size has to be at least three for the sender to find it worthwhile to provide more precise information than necessary.

Secondly, when the states are sufficiently independent or correlated, the sender does not fully reveal to any voter her state. Propositions 4.4 and 4.5 establish this fact. If the states are sufficiently independent, the information externality is not sufficiently strong for the benefits of full revelation to offset its cost in terms of forgone approval probability. On the other hand, if the states are sufficiently correlated, full revelation is not necessary: due to the strong correlation, the sender does not need to reduce by much the probability with which the strictest voter approves in her low state before the other voters are willing to rubber-stamp her decision.[26]

[26] In a sense, proposition 4.4 establishes a conceptual continuity between (i) full revelation to some voter for imperfectly correlated states, and (ii) lack of full revelation to any voter for perfectly correlated states.

Proposition 4.4 (No full revelation if states are sufficiently correlated). *For a fixed threshold profile* $(\ell_i)_{i=1}^n$*, there exists a critical degree of correlation above which full revelation to any voter is not optimal:*

$$(\pi_j(H), \pi_j(L)) \neq (1, 0) \text{ for any } j.$$

Proposition 4.5 (No full revelation if states are sufficiently independent). *For sufficiently independent states, all* IC^a *constraints bind in the optimal individual policy.*

Lastly, we fully characterize the parameter region under which some voters learn their own states fully for a relatively broad class of state distributions and a group in which all voters have the same threshold $\ell > 1$.[27]

Each distribution in this class is parametrized by $\lambda_1 \in [1/2, 1]$, which measures the degree of correlation among the voters' states. More specifically, let there be a grand state $\omega \in \{G, B\}$, for which $\Pr(\omega = G) = p_0$. The state of each voter is drawn conditionally independently according to the following probabilities: $\Pr(H|G) = \Pr(L|B) = \lambda_1$. In particular, $\lambda_1 = 1/2$ corresponds to independent payoff states, and $\lambda_1 = 1$ to perfectly correlated payoff states. Without loss, we assume $\pi_i(L) \leq \pi_j(L), \forall i < j$.

Figure 1 summarizes the optimal policy when the group size n is sufficiently large. When the states are sufficiently independent (i.e. $\lambda_1 \in (1/2, \ell/\ell + 1]$), the optimal policy is symmetric across all voters (Proposition B.4). No voter learns her state fully for such low λ_1. Moreover, every voter's IC^a constraint binds. If the states are sufficiently correlated (i.e., $\lambda_1 \in (\lambda_1^*, 1)$), the sender provides one voter with more precise information than her IC^a constraint requires so that every other voter is willing to rubber-stamp (Proposition B.2). Due to sufficiently high correlation, the better informed voter does not have to learn her state fully before the others are willing to rubber-stamp. When the state correlation is intermediate (i.e. $\lambda_1 \in (\ell/\ell + 1, \lambda_1^*]$), the optimal policy is to have some voters who are fully revealed their own states, one partially informed voter, and all other voters as rubber-stampers (Proposition B.3). This analysis shows that full revelation to some voters is optimal for a broad range of parameter values.

5 Comparison of persuasion modes

In this subsection we compare the different persuasion modes analyzed in sections 2 through 4. Proposition 3.1 has shown that independent general persuasion and

[27] The formal analysis and results for this case can be found in the online appendix B.5.

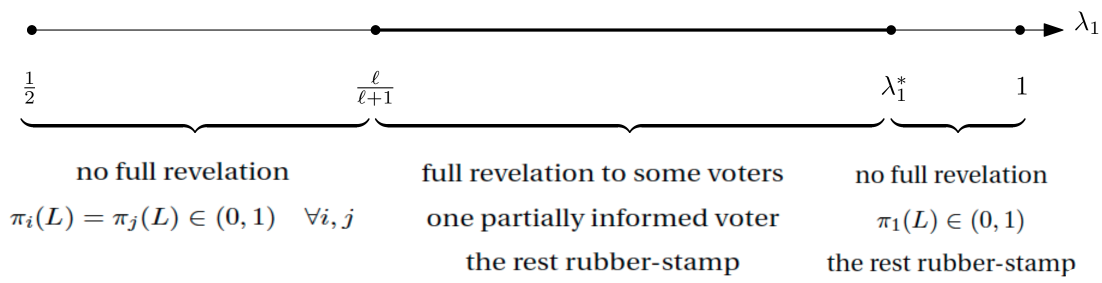

Figure 1: Optimal individual persuasion with homogeneous thresholds

general persuasion are equivalent under unanimity. Therefore, we focus on the comparison of general and individual persuasion.

Under both general and individual persuasion, more voters hurt the sender. We show that, under both modes, any approval probability that can be achieved by n voters can also be achieved after a voter is removed from the group. Hence, the sender is weakly better off after the removal of a voter. Intuitively, due to the required unanimity for a collective approval, a greater number of voters can only hurt the probability of a unanimous approval.

Lemma 5.1 (More voters hurt the sender). *For any fixed threshold profile, the sender attains a weakly higher payoff with $(n-1)$ voters than with n voters under both general and individual persuasion.*

When the states are sufficiently correlated, the sender attains the same payoff under general and individual persuasion. Under the latter, the sender persuades only the most demanding voter R_1, whereas all other more lenient voters are willing to rubber-stamp R_1's decision. This obviously induces the highest attainable payoff for the sender since persuading more voters than just R_1 can only leave him worse off. Therefore, general persuasion cannot improve upon the payoff from individual persuasion. There exists an optimal policy under general persuasion such that the approval probability in each state profile is the same across the two modes.

Lemma 5.2 (Equivalence with sufficiently high correlation). *Given a perfect correlation distribution f', there exists $\varepsilon > 0$ such that for any f with $||f - f'|| < \varepsilon$, the approval probability for any $\theta \in \Theta$ is the same across the two modes, i.e.,*

$$\pi(\theta) = \prod_{i=1}^{n} \pi_i(\theta_i),$$

where π and $(\pi_i)_{i=1}^n$ denote, respectively, an optimal general policy and an optimal individual policy associated with f.

As the correlation among the states weakens, the sender does strictly better under general persuasion than under individual persuasion. The sender is able to pool different state profiles more freely under general persuasion in order to obtain a higher payoff. In contrast, the strictest voter always weakly prefers individual persuasion. This is because any optimal general policy accords the strictest voter a zero payoff, while the optimal individual policy might accord her a strictly positive payoff. Moreover, if IC^a-i binds for $i \in \{1, ..., i'\}$ in any optimal general policy, the strictest voters $\{R_1, ..., R_{i'}\}$ are better off with individual persuasion.

While the strictest voter(s) unequivocally prefer individual persuasion, the ranking of the two persuasion modes by more lenient voters can go either way. For instance, in example 1, R_3 obtains a positive payoff under general persuasion when the sender chooses an optimal and Pareto efficient policy. Due to independent states, R_3's payoff must be zero under individual persuasion. In this example, the more lenient voter prefers general persuasion. We can find another example in which the most lenient voter obtains a stricter higher payoff under individual persuasion than under general persuasion, even if we restrict attention to optimal general policies that are Pareto efficient. As a result, the more lenient voters might disagree on the preferred persuasion mode.

6 Extensions

6.1 Public/sequential persuasion and observability of policies

This subsection discusses the robustness of the results as we relax two assumptions of the benchmark model: the private observability of recommendations, and the simultaneous structure of voting. We argue that the relaxation of either assumption does not affect the results from sections 2 through 5. Moreover, we briefly discuss the implications of privately observed policies under individual persuasion.

Let us first assume that the recommendations are announced publicly. Given any optimal policy under the previous assumption of private communication, the voting game that follows after voters publicly observe the same recommendations admits an obedient equilibrium in which all voters follow their respective recommendations. This is true for any of the three persuasion modes. If the public recommendation profile is the unanimous recommendation $\hat{d}^a$, the previous IC^a constraints ensure that all voters comply with the approval recommendation. If $\hat{d}^{r,i}$ is realized instead, the previous IC^r-i ensures that R_i prefers to reject. The other voters comply as well since they are no longer pivotal. If two or more voters receive a recommendation to reject, no voter can overturn the collective rejection given that all other voters follow their individual recommendations.[28] The following proposition summarizes this reasoning.

Proposition 6.1 (Public persuasion)**.** *Fix an optimal general or individual policy. Suppose that the recommendation profile $\hat{d}$ is observable by all voters. There exists an equilibrium in which each voter complies with the recommendation.*

Suppose now that the signals are privately observed but the sender encounters the voters sequentially in a particular order, one at a time. Each voter perfectly observes the decisions of preceding voters. We analyze how well the sender performs, compared to simultaneous voting, and whether the payoff achieved by the sender depends on the particular order in which the voters are approached.

Under unanimity rule, the decision of R_i is significant for the collective decision only if all preceding voters have already approved and all succeeding voters will approve as well. This observation—that in sequential voting too each voter decides as if all other voters have already approved the project—suggests that sequential voting is equivalent to simultaneous voting.[29] Indeed, the set of incentive-compatible policies under sequential voting is the same as that under simultaneous voting. Moreover, the order in which the voters are encountered is immaterial for the sender's payoff.

Proposition 6.2 (Sequential persuasion)**.** *Fix an optimal general or individual policy with simultaneous voting. This policy is also optimal under sequential voting for any order of voters.*

Furthermore, it follows from propositions 6.1 and 6.2 combined that the set of incentive-compatible policies remains the same if we relax the assumptions of private communication and simultaneous voting concurrently.

Finally, we discuss how results change if under individual persuasion, each voter privately observes her own policy. First of all, if we use Nash equilibrium as our solution concept, the optimal policy stays the same. This is because if the sender deviates and alters the equilibrium policy for R_i, then R_i will reject for sure. However, rejecting for sure on R_i's side regardless of the deviation policy and the realized signal is not sequentially rational. After the deviation, R_i needs to form a belief

[28] Moreover, if two or more voters receive a recommendation to reject, it cannot be the case that all voters prefer to approve. If that were so, the sender could recommend that all voters approve under private communication and increase his payoff. This contradicts the presumption that the policy is optimal under private communication.

[29] The reasoning clearly does not extend to nonunanimous voting rules.

about the sender's policies toward other voters and their behavior, in order to form a belief about their respective states. The most pessimistic belief R_i can hold is that all other voters' states are L upon approval. If the deviation policy toward R_i is incentive compatible under this most pessimistic belief, then R_i is willing to follow the recommendation. This gives an upper bound on how informative R_i's equilibrium policy can be. If $\pi_i(H) = 1$, this gives a positive lower bound on how low $\pi_i(L)$ can be in order for the policy to be sustained in equilibrium. The sender's problem is essentially the same as before except that each $\pi_i(L)$ now must stay above its lower bound.

By this reasoning, the optimal policy identified in example 2 is not an equilibrium. The sender prefers to deviate from the fully revealing policy π_1 and increase $\pi_1(L)$ by a very small amount: for any belief R_1 might hold, there exists a sufficiently small $\pi_1(L)$ for which R_1 continues to obey the recommendation. Full revelation does not arise with privately observed policies.

However, the optimal policy identified in example 3 is credible under the most pessimistic off-path beliefs described above. R_1 obeys a recommendation of an off-path policy $\tilde{\pi}_1$ with $\tilde{\pi}_1(H) = 1$ if and only if $\tilde{\pi}_1(L) < f(HLL)/(\ell_1 f(LLL)) = 0.00042$. This imposes a lower bound on any credible $\pi_1(L)$. Hence, the sender does not deviate from the optimal policy $\pi_1^*(L) = 0.003$. Similarly, the sender prefers not to deviate from π_2^* and π_3^* as well. This example shows that even with privately observed individual policies, there are environments $(f, (\ell_i)_{i=1}^n)$ for which the sender provides to the strictest voters more precise information than what is needed to persuade them in order to persuade others more often.

6.2 Nonunanimous decision rules

This subsection considers nonunanimous voting rules, i.e., when $k < n$ votes suffice for project adoption. We show that the project is approved with certainty under general persuasion, and with probability strictly less than one under individual persuasion. Voters, on the other hand, strictly prefer individual to general persuasion.

A general policy that trivially achieves a certain approval is one that always recommends all voters to approve for all state profiles. No voter is ever pivotal, so each voter trivially follows the approval recommendation. This construction, however, relies on the failure to be pivotal, so each voter is indifferent when asked to approve. We take a different route.

As noted in section 1, we allow only for any policy that is the limit of a sequence of full-support incentive-compatible policies. For each limiting policy, there exists a nearby full-support policy for which each voter is pivotal with positive probability for both approval and rejection recommendations, and each voter's incentive-compatibility constraints are strictly satisfied.[30] We can construct a sequence of such full-support policies that achieves a payoff arbitrarily close to one for the sender. This shows that the certain-approval result does not rely on the voters' failure to be pivotal.

This certain-approval result stands in contrast to the optimal policy under unanimity where the sender's payoff is strictly below one. The key observation is that the sender benefits from the event that *at least* k voters approve, whereas each voter cares about the event that *exactly* k voters (including herself) receive an approval recommendation. When $k < n$, the sender can recommend most of the time that at least $k + 1$ voters approve, without jeopardizing the IC constraints of the voters. This observation explains the discontinuity between the unanimous and nonunanimous rules.

Proposition 6.3 (Certain approval under general policies). [31] *Suppose that the sender needs $k \leq n - 1$ approvals. Under general persuasion, the sender's payoff is one.*

θ \ $\hat{d}$	(0,0)	(1,0)	(0,1)	(1,1)
HH	ε^2	ε	ε	$1 - 2\varepsilon - \varepsilon^2$
HL	ε^2	ε^2	ε^2	$1 - 3\varepsilon^2$
LH	ε^2	ε^2	ε^2	$1 - 3\varepsilon^2$
LL	ε	ε^2	ε^2	$1 - \varepsilon - 2\varepsilon^2$

Table 2: General policy for $n = 2, k = 1$

To illustrate the proof idea, we construct a full-support policy for $n = 2$ and $k = 1$ in table 2. Each row summarizes the recommendation distribution for a fixed state profile. Upon receiving an approval recommendation, R_1 is pivotal only if recommendation $(1, 0)$ has realized. This recommendation is sent much more frequently in state profile HH than in other state profiles; therefore, R_1 is confident that her state is H when conditioning on being pivotal. The reasoning for R_2 is similar. Therefore, the constructed policy guarantees compliance with approval recommendations. On the other hand, the only recommendation for which R_i is pivotal upon being recommended to reject is $(0, 0)$. Since this recommendation

[30] Suppose that the sender does not provide any information and voters do not use weakly dominated strategies. For any voting rule $1 \leq k \leq n$, all voters reject in the unique correlated equilibrium.

[31] The proofs for propositions 6.3 and 6.4 can be found in online appendix B.1. Chan et al. (2016) independently reach a similar result to proposition 6.3 for the case of perfectly correlated states.

profile is suggestive of the unfavorable state profile LL more than of any other state profile, both voters are willing to follow a rejection recommendation. Therefore, this sequence of policies (indexed by ε) are obedient and attain a payoff for the sender that is arbitrarily close to one.

The certain-approval result is further strengthened by the observation that the sender achieves a certain approval even when constrained to independent general persuasion.

Proposition 6.4 (Certain approval under independent general policies). *Suppose that f has full support and the sender needs $k \leq n-1$ approvals. Under independent general persuasion, the sender's payoff is one.*

To illustrate the proof idea, we revisit the two-voter example with $k = 1$. The left-hand side of table 6.2 demonstrates the independent policy for each voter. The right-hand side shows the corresponding probability of each recommendation profile for each state profile.[32]

θ \ π_i	$\pi_1(\cdot)$	$\pi_2(\cdot)$
HH	$1-\varepsilon_1$	$1-\varepsilon_1$
HL	$1-\varepsilon_2$	ε_3
LH	ε_3	$1-\varepsilon_2$
LL	$1-\varepsilon_4$	$1-\varepsilon_4$

θ \ $\hat{d}$	(0,0)	(1,0)	(0,1)
HH	ε_1^2	$\varepsilon_1(1-\varepsilon_1)$	$\varepsilon_1(1-\varepsilon_1)$
HL	$\varepsilon_2(1-\varepsilon_3)$	$(1-\varepsilon_2)(1-\varepsilon_3)$	$\varepsilon_2\varepsilon_3$
LH	$\varepsilon_2(1-\varepsilon_3)$	$\varepsilon_2\varepsilon_3$	$(1-\varepsilon_2)(1-\varepsilon_3)$
LL	ε_4^2	$\varepsilon_4(1-\varepsilon_4)$	$\varepsilon_4(1-\varepsilon_4)$

Table 3: Independent general policy for $n = 2, k = 1$

In state profiles with exactly one high-state voter, this voter is recommended to approve with very high probability, while the other low-state voter is recommended to reject with very high probability. Hence, when R_1 is recommended to approve and she conditions on $(1, 0)$ being sent, the state profile is most likely to be HL. Because HL is favorable for R_1, she is willing to approve. By the same logic, upon being recommended to reject, R_1 conditions on $(0, 0)$ being sent. If ε_4 is chosen so as to shrink to zero at a much slower rate than ε_1 and ε_2, R_1 is sufficiently assured that the state profile is the unfavorable LL. Hence, she obeys the rejection recommendation. The reasoning for R_2 is similar. Therefore, the policies in the sequence are obedient and attain an arbitrarily high payoff for the sender.

More generally under independent general persuasion, whenever a state profile with exactly k high-state voters realizes, the high-state voters are very likely to be recommended to approve and the low-state voters very likely to be recommended to reject. Therefore, when R_i is recommended to approve, the pivotality of her decision

[32] The distribution for recommendation profile $(1, 1)$ has been omitted for ease of exposition.

suggests that the realized state profile very probably admits exactly k high-state voters, with R_i being among them. This construction has the flavor of "targeted persuasion" as in Alonso and Câmara (2016b), since the sender targets the voters who benefit from the project. By making the state profiles with exactly k high-state approvers the most salient event for each voter when her vote is pivotal, the sender is able to achieve a guaranteed approval.

Our construction relies on the existence of state profiles with $k > 0$ high-state voters and $(n - k) > 0$ low-state voters. So such state profiles have to occur with positive probability in order for the argument to hold. Lemma B.1 in online appendix B.1 shows that for any affiliated and exchangeable state distribution f, either (i) f has full support on Θ, or (ii) the states are perfectly correlated. If the latter holds, the problem of the sender reduces to one of individual persuasion, under which a certain approval is never attained, as shown in the next proposition. Hence, a full-support state distribution is both necessary and sufficient for certain approval under independent general persuasion.

Whenever each voter's recommendation can depend on the entire profile θ, a requirement for nonunanimous consent effectively imposes no check on the adoption of the project, as approval is guaranteed. We next show that when the sender can condition individual recommendations to a voter only on her state, the project is never approved for sure.

Proposition 6.5 (No certain approval under individual policies)**.** *Suppose that the sender needs $k \leq n - 1$ approvals. Under individual persuasion, the sender's payoff is strictly below one. The payoff of each voter is strictly higher than that under general and independent general persuasion.*

Under individual persuasion, each voter approves more frequently under a high state than a low state. If the project is approved with certainty, it must be approved with certainty under all possible state profiles. There must then exist a coalition of at least k voters who approve the project regardless of their individual states. However, any voter in this coalition does not become more optimistic about her state being high when she receives a recommendation to approve and considers her decision to be pivotal. Due to assumption 1, she is not willing to follow the approval recommendation. Hence, the project cannot be approved with certainty.

We further establish that there exists at least one voter who approves strictly more frequently under a high state than under a low state and she is pivotal with positive probability. Due to affiliation, all other voters benefit from such a voter's decision. Therefore, for any nonunanimous voting rule, all voters are strictly better off under individual persuasion than under either general or independent general persuasion.

The role of communication. For a fixed information structure π or $(\pi_i)_{i=1}^n$, a stage of pre-voting communication can be added to this voting game of incomplete information. The set of equilibria depends on the communication protocol. For instance, we can add one round of public cheap-talk communication: all voters first observe their own private signal and then simultaneously announce a public cheap-talk message. Afterwards they vote. For this particular communication protocol, there is always an equilibrium in which each voter votes informatively based on her private signal generated according to π or $(\pi_i)_{i=1}^n$.[33] This is clearly the optimal equilibrium for the sender among all possible communication equilibria.[34] Moreover, if all voters *share a consensus* given π or $(\pi_i)_{i=1}^n$, this communication protocol admits also an equilibrium in which all voters announce their signals truthfully and vote according to the aggregated information.[35] This might lead to a lower payoff for the sender than the previous equilibrium. In general, when voters do not share such a consensus, there might not exist such a full-revelation equilibrium.

To sum up, for any information structure π or $(\pi_i)_{i=1}^n$, the equilibrium outcome when the voters can communicate depends not only on the communication protocol but also on the equilibrium selection criterion. We view our result as an important and useful benchmark when the sender-optimal equilibrium is selected.

7 Concluding remarks

This paper analyzed the problem faced by an uninformed sender aiming to persuade a unanimity-seeking group of heterogeneous voters through the design of informative experiments. We characterized the optimal policies for two main modes of persuasion, —general and individual persuasion—and explored their implications for the players' welfare. Returning to our motivating examples, our results clarify who benefits from particular forms of evidence: the most demanding regulators prefer targeted evidence, while comprehensive evidence benefits, besides the sender,

[33]There is always a babbling equilibrium in which each voter sends an uninformative message and then votes according to her private signal.

[34]In fact, for any direct obedient π or $(\pi_i)_{i=1}^n$, there exists a communication equilibrium in which voters follow their private signals. See Gerardi and Yariv (2007) for details.

[35]As defined in Coughlan (2000) and Austen-Smith and Feddersen (2006), a consensus exists if, given the public revelation of all private information, all voters always agree on whether to approve. If voters do not share such a consensus, a full-revelation equilibrium might not exist. Austen-Smith and Feddersen (2006) present an example in which voters have the same state but different thresholds. Given the information structure in that example, there is no full-revelation equilibrium.

only the least demanding ones. Moreover, providing evidence that is comprehensive yet independent across regulators does not constrain the sender at all in his ability to persuade. When restricted to targeted evidence, —for instance, in highly specialized regulatory evaluations, — the sender might find it optimal to accurately inform the strictest regulators of the quality of their aspect of interest. An institutional arrangement that is based on targeted evidence designates a strong informational role to the most demanding regulators: the sender leverages the fully revealing recommendations provided to them to persuade others more easily. When the voters' states are affiliated and exchangeable, the possibility of full revelation is exclusive to targeted evidence.

Under unanimity rule, the characterized optimal policies remain optimal even when the voters publicly observe the recommendation profile and/or take turns to cast their votes. For nonunanimous rules, the sender obtains a sure approval if allowed to offer a general policy or an independent general policy to the group. In the context of regulation, this is effectively as if there were no regulatory check on the proposals of the industry as long as it can provide comprehensive evidence regarding all payoff-relevant aspects. We interpret this result as advocating for the institutionalization of unanimity rule whenever the sender (i) communicates with voters in private, and (ii) offers comprehensive evidence. Moreover, the fact that all voters prefer individual to general persuasion under any nonunanimous rule supports the institutionalization of targeted evidence in environments governed by such rules.

Our analysis of unanimity rule invites future work on the full characterization of optimal individual persuasion for nonunanimous rules. We have taken a first step by analyzing some features of individual persuasion under such rules. The effect of communication among voters on the persuasion effort of the sender also remains largely unexplored. Another natural question concerns the implications of sequential persuasion in settings with and without payoff externalities among receivers. For sequential collective decision-making under different persuasion modes, a departure from the unanimity rule complicates the analysis substantially. Moreover, we leave for future work a systematic examination of differences among persuasion modes when the receivers vary in their informational importance for the group, that is, when the assumption of exchangeability is dropped. The exploration of these questions promises to shed further light on the dynamics of group persuasion.

A Appendix: Proofs for sections 2 to 4

Proof for proposition 2.1. The policy only specifies $\pi(\hat{d}^a|\theta^H)$ and $\pi(\hat{d}^a|\theta^L)$. First, $\pi(\hat{d}^a|\theta^H) = 1$. Otherwise the objective can be improved without hurting any IC^a constraints. Secondly, in order for $\pi(\hat{d}^a|\theta^L)$ to be as high as possible without hurting any IC^a constraints, the sender optimally sets it so as to make IC^a-1 bind (because $\ell_1 = \max_i \ell_i$). Hence, $f(\theta^H) - f(\theta^L)\pi(\hat{d}^a|\theta^L)\ell_1 = 0$. □

Proof for lemma 2.1. Suppose that $\pi(\hat{d}^a|\theta^H) < 1$. Increasing $\pi(\hat{d}^a|\theta^H)$ relaxes IC^a-i for all i as $\theta^H \in \Theta_i^H$ for all i. It also strictly improves the payoff of the sender. Hence $\pi(\hat{d}^a|\theta^H) < 1$ cannot be optimal. □

Proof for lemma 2.2. Suppose to the contrary that there exists i such that $\pi(\hat{d}^a|\theta) = 1$ for all $\theta \in \Theta_i^L$. Then IC^a-i requires that $\left(\sum_{\theta\in\Theta_i^H} f(\theta)\pi(\hat{d}^a|\theta)\right) / \left(\sum_{\theta\in\Theta_i^L} f(\theta)\right) \geq \ell_i$. But assumption 1 tells us that

$$\ell_i > \frac{\sum_{\theta\in\Theta_i^H} f(\theta)}{\sum_{\theta\in\Theta_i^L} f(\theta)} \geq \frac{\sum_{\theta\in\Theta_i^H} f(\theta)\pi(\hat{d}^a|\theta)}{\sum_{\theta\in\Theta_i^L} f(\theta)}$$

for any specification of $\pi(\hat{d}^a|\theta)$ for $\theta \in \Theta_i^H$. We have thus reached a contradiction.

Suppose to the contrary that for some i, $\pi(\hat{d}^a|\theta) = 0$ for all $\theta \in \Theta_i^L$. Then IC^a-i is slack. Consider $\theta' \in \Theta_i^L$ such that $\theta'_j = H$ for all $j \neq i$. Consider increasing $\pi(\hat{d}^a|\theta')$ to some strictly positive value so as to still satisfy IC^a-i. This increase: i) strictly improves the payoff of the sender, ii) relaxes all IC^a-j for $j \neq i$. □

Proof for proposition 2.2. Slightly reformulated, the sender's problem is

$$\min_{\pi(\hat{d}^a|\theta)\geq 0} \; -\sum_{\theta\in\Theta} f(\theta)\pi(\hat{d}^a|\theta)$$

$$\text{s.t. } \pi(\hat{d}^a|\theta) - 1 \leq 0, \; \forall\theta, \quad \text{and} \sum_{\theta\in\Theta_i^L} f(\theta)\pi(\hat{d}^a|\theta)\ell_i - \sum_{\theta\in\Theta_i^H} f(\theta)\pi(\hat{d}^a|\theta) \leq 0, \; \forall i.$$

Let γ_θ be the dual variable associated with $\pi(\hat{d}^a|\theta) - 1 \leq 0$ and μ_i the dual variable associated with IC^a-i. Since the constraints of the primal are inequalities, for the dual we have the constraints that $\gamma_\theta \geq 0$ for all θ and $\mu_i \geq 0$ for all i. The dual is:

$$\min_{\gamma_\theta\geq 0, \mu_i \geq 0} \sum_{\theta\in\Theta} \gamma_\theta, \; \text{ s.t. } \; \gamma_\theta \geq f(\theta)\left(1 + \sum_{i:\theta_i = H} \mu_i - \sum_{i:\theta_i = L} \mu_i \ell_i\right), \; \forall \theta \in \Theta.$$

In the dual problem, there is an inequality constraint for each state profile θ. The associated primal variable for each inequality constraint is $\pi(\hat{d}^a|\theta)$.

We first show that it cannot be that $\mu_i = 0$ for all i. Suppose that was indeed the case. Then $\sum_{\theta\in\Theta} \gamma_\theta = \sum_{\theta\in\Theta} f(\theta) = 1$. Consider increasing μ_1 by a small amount $\varepsilon > 0$. The dual objective changes by $\varepsilon \sum_{\theta \in \Theta_1^H} f(\theta) - \varepsilon \ell_1 \sum_{\theta \in \Theta_1^L} f(\theta) < 0$ by assumption 1. Hence the dual objective can be improved.

We next show that there exists an optimal solution to the dual such that if $\mu_j = 0$, then $\mu_{j'} = 0$ for any $j' > j$. Suppose that $\mu_j = 0$ and $\mu_{j'} > 0$ for some $j' > j$. We can rewrite the constraint associated with θ as:

$$\gamma_\theta \geq f(\theta) \left(1 + \sum_{\substack{i \neq j,j' \\ \theta_i = H}} \mu_i - \sum_{\substack{i \neq j,j' \\ \theta_i = L}} \mu_i \ell_i + \sum_{i \in \{j,j'\}} \mu_i \left(\mathbf{1}_{\theta_i = H} - \mathbf{1}_{\theta_i = L} \ell_i \right) \right).$$

We construct a new set of $\tilde{\mu}_i$ and $\tilde{\gamma}_\theta$ such the the dual objective stays constant and the dual constraints are satisfied. We let $\tilde{\mu}_{j'} = \mu_j = 0$, $\tilde{\mu}_j = \mu_{j'} > 0$ and $\tilde{\mu}_i = \mu_i$ for $i \neq j, j'$. To construct $\tilde{\gamma}_\theta$, we consider two cases:

1. If the state profile θ is such that $\theta_j = \theta_{j'}$, then we let $\tilde{\gamma}_\theta = \gamma_\theta$. Since $\ell_j > \ell_{j'}$ and $\tilde{\mu}_j > \tilde{\mu}_{j'} = 0$, the RHS of the inequality associated with $\tilde{\gamma}_\theta$ is weakly lower than before, so the constraint is satisfied.

2. If the state profile θ is such that $\theta_j \neq \theta_{j'}$, then there exists another state profile θ' such that θ' is the same as θ except for θ'_j and $\theta'_{j'}$. In this case, we let $\tilde{\gamma}_\theta = \gamma_{\theta'}$ and $\tilde{\gamma}_{\theta'} = \gamma_\theta$. The inequalities associated with $\tilde{\gamma}_\theta$ and $\tilde{\gamma}_{\theta'}$ are both satisfied. This is easily verified given that $\ell_j > \ell_{j'}$, $f(\theta) = f(\theta')$, and the presumption that $\mu_{j'} > \mu_j = 0$ (or equivalently, $\tilde{\mu}_j > \tilde{\mu}_{j'} = 0$).

This shows that there exists an optimal solution to the dual such that μ_i's for the most demanding regulators are positive. Moreover, based on case (ii), if $\mu_{j'} > \mu_j = 0$ and $j' > j$, it must be true that $\gamma_\theta = 0$ for any θ such that $\theta_j = H, \theta_{j'} = L$. Otherwise, after the exchange operation in case (ii), we can lower γ_θ (or equivalently, $\tilde{\gamma}_{\theta'}$) without violating any constraint.

We next show that in *any* optimal solution of the dual, if $\mu_j = 0$, then $\mu_{j'} = 0$ for any $j' > j$. Suppose not. Suppose that $\mu_j = 0$ and $\mu_{j'} > 0$ for some $j' > j$. This implies that $\gamma_\theta = 0$ for any θ such that $\theta_j = H, \theta_{j'} = L$. In particular, $\gamma_\theta = 0$ when $\theta_j = H, \theta_{j'} = L$ and all other states are H. The constraint with respect to this state profile is $\gamma_\theta = 0 \geq f(\theta) \left(1 + \sum_{i \neq j,j'} \mu_i - \mu_{j'} \ell_{j'}\right)$. This imposes a lower bound on $\mu_{j'}$: $\mu_{j'} \geq \left(1 + \sum_{i \neq j,j'} \mu_i\right) / \ell_{j'}$. We next show that it is uniquely optimal to set $\mu_{j'}$ to be this lower bound. Firstly, this lower bound of $\mu_{j'}$ ensures that for any θ such that $\theta_{j'} = L$, we can set γ_θ to be zero. For any θ such that $\theta_{j'} = H$, lowering $\mu_{j'}$ makes the constraint associated with θ easier to be satisfied. In particular, for θ such that all states are H, lowering $\mu_{j'}$ strictly decreases the lower bound on γ_θ and thus strictly lower the dual objective. This shows that it is uniquely optimal to set $\mu_{j'}$ to be the lower bound $(1 + \sum_{i \neq j,j'} \mu_i)/\ell_{j'}$.

Lastly, we show that we can strictly improve by constructing a new set of $\tilde{\mu}_i$ and $\tilde{\gamma}_\theta$. We let $\tilde{\mu}_{j'} = 0$, $\tilde{\mu}_j = (1 + \sum_{i \neq j,j'} \mu_i)/\ell_j < (1 + \sum_{i \neq j,j'} \mu_i)/\ell_{j'}$ and $\tilde{\mu}_i = \mu_i$ for $i \neq j, j'$. For any θ such that $\theta_j = L$, we can set $\tilde{\gamma}_\theta$ to be zero. For any θ such that $\theta_j = H$, there exists a state profile θ' such that $\theta'_{j'} = \theta_j$, $\theta'_j = \theta_{j'}$, and $\theta'_i = \theta_i$ for $i \neq j, j'$. We let $\tilde{\gamma}_\theta = \gamma_{\theta'}$. Since $\tilde{\mu}_j$ is strictly smaller than $\mu_{j'}$, the constraint associated with $\tilde{\gamma}_\theta$ is satisfied. Moreover, the lower $\tilde{\mu}_j$ ensures that we can further lower $\tilde{\gamma}_{\theta^H}$ (which was equal to γ_{θ^H}). This shows that we can strictly decrease the objective. We have thus reached a contradiction.

Therefore, it cannot be true that $\mu_j = 0$ and $\mu_{j'} > 0$ for some $j' > j$. Moreover, at least one μ_i is strictly positive. This shows that there exists an $i' \geq 1$ such that $\mu_i > 0$ for $i \leq i'$. Therefore, in any optimal policy, the strictest regulators' IC^a constraints bind. □

Proof for proposition 3.1. Consider a general policy $\pi = \left(\pi(\hat{d}|\theta)\right)_{\hat{d},\theta}$. We want to construct an independent general policy $(\pi_i(\theta)_{i,\theta}$ that implements the same payoff for the sender. We construct $(\pi_i(\theta))_{i,\theta}$ such that $\prod_{i=1}^n \pi_i(\theta) = \pi(\hat{d^a}|\theta)$ for any $\theta \in \Theta$. IC^a constraints (2) are satisfied automatically since the right-hand side of (2) depends only on the product $\prod_{i=1}^n \pi_i(\theta)$. We next show that we can choose $(\pi_i(\theta))_{i,\theta}$ to satisfy IC^r constraints (3) as well. First, if $\pi(\hat{d^a}|\theta) = 0$, we set $\pi_i(\theta)$ to be zero for all i. Such state profiles contribute zero to the left-hand side of (3). Secondly, if $\pi(\hat{d^a}|\theta) = 1$, we have to set $\pi_i(\theta) = 1$ for all i. Again, such state profiles contribute nothing to the left-hand side of (3). Lastly, if $\pi(\hat{d^a}|\theta) \in (0, 1)$, we set $\pi_i(\theta)$ to be one if $\theta_i = H$ and $\pi_i(\theta)$ to be interior if $\theta_i = L$. Such state profiles contribute a weakly negative term to the left-hand side of (3). Thus, we have shown that (3) is satisfied. □

Proof for proposition 4.1. Under perfect correlation, IC^a constraint can be rewritten compactly as: $f(\theta^H) \prod_{j=1}^n \pi_j(H) \geq f(\theta^L)\ell_i \prod_{j=1}^n \pi_j(L)$, while the objective is: $f(\theta^H) \prod_{j=1}^n \pi_j(H) + f(\theta^L) \prod_{j=1}^n \pi_j(L)$. Increasing $\prod_j \pi_j(H)$ relaxes all approval IC constraints and also benefits the sender, therefore the sender sets $\pi_j(H) = 1$ for all j. The sender's payoff increases in $\prod_j \pi_j(L)$, so he sets it to be the highest possible value allowed by the IC^a constraints: $\prod_{j=1}^n \pi_j(L) = f(\theta^H)/\left(f(\theta^L)\ell_1\right)$. This policy satisfies IC^r constraints as well. If R_i receives a rejection recommendation, she knows for sure that her state is L because $\pi_i(H) = 1$, hence she strictly prefers to reject the project.

If states are independent, R_i receives no additional information from conditioning on the approval of R_{-i}. IC^a-i is simply: $\sum_{\theta \in \Theta_i^H} f(\theta)\pi_i(H) - \sum_{\theta \in \Theta_i^L} f(\theta)\pi_i(L)\ell_i \geq 0$. The policy offered to one voter does not affect the sender's ability to persuade other voters. So the problem of the sender is separable across voters. Therefore, the following single-voter policy is optimal: $\pi_i(H) = 1$, and

$$\pi_i(L) = \frac{\sum_{\theta \in \Theta_i^H} f(\theta)}{\ell_i \sum_{\theta \in \Theta_i^L} f(\theta)}.$$

□

Proof for lemma 4.1. Suppose $\pi_j(H) < 1$ for some j. It is easily verified that (OBJ) increases in $\pi_j(H)$, and increasing $\pi_j(H)$ weakly relaxes IC^a-j. It remains to be shown that increasing $\pi_j(H)$ weakly relaxes IC^a-i for any $i \neq j$. Consider IC^a-i rewritten in the following form:

$$\frac{\pi_j(H) \sum_{\Theta_i^H \cap \Theta_j^H} f(\theta) \prod_{k \neq i,j} \pi_k(\theta_k) + \pi_j(L) \sum_{\Theta_i^H \cap \Theta_j^L} f(\theta) \prod_{k \neq i,j} \pi_k(\theta_k)}{\pi_j(H) \sum_{\Theta_i^L \cap \Theta_j^H} f(\theta) \prod_{k \neq i,j} \pi_k(\theta_k) + \pi_j(L) \sum_{\Theta_i^L \cap \Theta_j^L} f(\theta) \prod_{k \neq i,j} \pi_k(\theta_k)} \geq \ell_i \frac{\pi_i(L)}{\pi_i(H)}.$$

For this step, note that if the denominator of the LHS equals zero, then IC^a-i holds for any $\pi_j(H)$, and if the denominator of the RHS equals zero, then $\pi_i(L)$ has to be zero so IC^a-i holds for any $\pi_j(H)$ as well. Therefore, we focus on the case in which neither denominator equals zero.

The derivative of the LHS with respect to $\pi_j(H)$ is positive if:

$$\left(\sum_{\Theta_i^H \cap \Theta_j^H} f(\theta) \prod_{k \neq i,j} \pi_k(\theta_k) \right) \left(\sum_{\Theta_i^L \cap \Theta_j^L} f(\theta) \prod_{k \neq i,j} \pi_k(\theta_k) \right) - \left(\sum_{\Theta_i^H \cap \Theta_j^L} f(\theta) \prod_{k \neq i,j} \pi_k(\theta_k) \right) \left(\sum_{\Theta_i^L \cap \Theta_j^H} f(\theta) \prod_{k \neq i,j} \pi_k(\theta_k) \right) \geq 0. \quad (4)$$

To prove this inequality, we will use the Ahlswede-Daykin inequality (Ahlswede and Daykin, 1978): Suppose $(\Gamma, \succ)$ is a finite distributive lattice and functions f_1, f_2, f_3, $f_4 : \Gamma \to \mathbb{R}^+$ satisfy the relation that $f_1(a) f_2(b) \leq f_3(a \wedge b) f_4(a \vee b)$, $\forall a, b \in \Gamma$. Then $f_1(A) f_2(B) \leq f_3(A \wedge B) f_4(A \vee B)$, $\forall A, B \subset \Gamma$, where $f_k(A) = \sum_{a \in A} f_k(a)$ for all $A \subset \Gamma, k \in \{1, 2, 3, 4\}$, and $A \vee B = \{a \vee b : a \in A, b \in B\}$, $A \wedge B = \{a \wedge b : a \in A, b \in B\}$.

First, we let Γ be the set of state profiles excluding R_i and R_j. That is, $\Gamma = \Theta_{-ij}$. The lattice Γ is finite. Secondly, the lattice is distributive, i.e. for any $\theta, \theta', \theta'' \in \Theta$ $\theta_{-ij} \vee (\theta'_{-ij} \wedge \theta''_{-ij}) = (\theta_{-ij} \vee \theta'_{-ij}) \wedge (\theta_{-ij} \vee \theta''_{-ij})$. To see this, consider the state of $R_k \neq R_i, R_j$ in both sides. If $\theta_k = H$, then the state of R_k is H in both sides. If $\theta_k = L$ and $\theta'_k = \theta''_k$, then the state of R_k is θ'_k in both sides. If $\theta_k = L$, and $\theta'_k \neq \theta''_k$, the state of R_k is L in both sides. Therefore the lattice is distributive. We write $f(\theta) = f(\theta_i \theta_j, \theta_{-ij})$. Define functions:

$$f_1(\theta_{-ij}) = f(HL, \theta_{-ij}) \pi_{-ij}(\theta_{-ij}), \quad f_2(\theta_{-ij}) = f(LH, \theta_{-ij}) \pi_{-ij}(\theta_{-ij}),$$

$$f_3(\theta_{-ij}) = f(LL, \theta_{-ij})\pi_{-ij}(\theta_{-ij}), \quad f_4(\theta_{-ij}) = f(HH, \theta_{-ij})\pi_{-ij}(\theta_{-ij}),$$

where $\pi_{-ij}(\theta_{-ij}) := \prod_{k\neq i,j} \pi_k(\theta_k)$. Due to affiliation and the easily verified equality that $\pi_{-ij}(\theta_{-ij})\pi_{-ij}(\theta'_{-ij}) = \pi_{-ij}(\theta_{-ij} \wedge \theta'_{-ij})\pi_{-ij}(\theta_{-ij} \vee \theta'_{-ij})$, the premise of the theorem holds:

$$f(HL, \theta_{-ij})\pi_{-ij}(\theta_{-ij})f(LH, \theta'_{-ij})\pi_{-ij}(\theta'_{-ij}) \leq \\ f(LL, \theta_{-ij} \wedge \theta'_{-ij})\pi_{-ij}(\theta_{-ij} \wedge \theta'_{-ij})f(HH, \theta_{-ij} \vee \theta'_{-ij})\pi_{-ij}(\theta_{-ij} \vee \theta'_{-ij}).$$

Take subsets of the lattice $A = B = \Gamma$, so $A \wedge B = A \vee B = \Gamma$. It follows from Ahlswede-Daykin inequality that

$$\left(\sum_{\theta_{-ij}\in A} f(HL, \theta_{-ij})\pi_{-ij}(\theta_{-ij})\right)\left(\sum_{\theta_{-ij}\in B} f(LH, \theta_{-ij})\pi_{-ij}(\theta_{-ij})\right) \leq \\ \left(\sum_{\theta_{-ij}\in A\wedge B} f(LL, \theta_{-ij})\pi_{-ij}(\theta_{-ij})\right)\left(\sum_{\theta_{-ij}\in A\vee B} f(HH, \theta_{-ij})\pi_{-ij}(\theta_{-ij})\right).$$

This is precisely the inequality we wanted to show. This concludes the proof. □

Proof for lemma 4.2. We first argue that there exists a voter R_i such that $\pi_i(L) < 1$. Suppose that for all i, $\pi_i(L) = 1$. Then R_i's belief of her state being H if she conditions on the approval of all other voters is equal to the prior belief $\sum_{\theta\in\Theta_i^H} f(\theta)$. But given assumption 1, $(\pi_i(H), \pi_i(L)) = (1, 1)$ is not incentive compatible under the prior. Suppose now that $\pi_i(L) < 1$ for some i and IC^a-j is slack for all j. Then increasing $\pi_i(L)$ by a small amount strictly increases the probability that the project is approved without violating any IC^a constraint.

We then show that $\pi_i(L) > 0$ for some i. Suppose $\pi_i(L) = 0$ for all i. The sender fully reveals the payoff state to each voter. Hence, all IC^a constraints are slack. This contradicts the first part of this proof. □

Proof for proposition 4.2. Suppose there exist R_i and R_j such that $\ell_j > \ell_i$ and $\pi_j(L) > \pi_i(L)$. We can rewrite IC^a-i and IC^a-j as:

$$\ell_i\pi_i(L) \leq \frac{a_1 + a_2\pi_j(L)}{a_2 + a_3\pi_j(L)}, \quad \ell_j\pi_j(L) \leq \frac{a_1 + a_2\pi_i(L)}{a_2 + a_3\pi_i(L)},$$

where

$$a_1 := \sum_{\Theta_i^H\cap\Theta_j^H} f(\theta) \prod_{k\neq i,j} \pi_k(\theta_k), \quad a_3 := \sum_{\Theta_i^L\cap\Theta_j^L} f(\theta) \prod_{k\neq i,j} \pi_k(\theta_k),$$

$$a_2 := \sum_{\Theta_i^H\cap\Theta_j^L} f(\theta) \prod_{k\neq i,j} \pi_k(\theta_k) = \sum_{\Theta_i^L\cap\Theta_j^H} f(\theta) \prod_{k\neq i,j} \pi_k(\theta_k).$$

It is easily verified that $a_1, a_2, a_3 \geq 0$. Moreover, using an argument similar to the proof for lemma 4.1, we can show that $a_2^2 \leq a_1 a_3$. Therefore, $(a_1 + a_2 x)/(a_2 + a_3 x)$ decreases in x.

Multiplying the LHS of IC^a-j by $\ell_i/\ell_j \in (0,1)$, we obtain the following inequality:

$$\ell_i \pi_j(L) < \frac{a_1 + a_2 \pi_i(L)}{a_2 + a_3 \pi_i(L)}. \tag{5}$$

Moreover, it is easy to show that $\frac{a_1+a_2 x}{a_2+a_3 x}x$ increases in x. Given the presumption that $\pi_i(L) < \pi_j(L)$, we thus have

$$\frac{\pi_i(L)}{\pi_j(L)} \leq \frac{\frac{a_1+a_2\pi_j(L)}{a_2+a_3\pi_j(L)}}{\frac{a_1+a_2\pi_i(L)}{a_2+a_3\pi_i(L)}}.$$

Multiplying the LHS of IC^a-j by $\pi_i(L)/\pi_j(L)$ and the RHS of IC^a-j by

$$\left(\frac{a_1 + a_2\pi_j(L)}{a_2 + a_3\pi_j(L)}\right) \Big/ \left(\frac{a_1 + a_2\pi_i(L)}{a_2 + a_3\pi_i(L)}\right),$$

we then obtain the following inequality:

$$\ell_j \pi_i(L) \leq \frac{a_1 + a_2\pi_j(L)}{a_2 + a_3\pi_j(L)}. \tag{6}$$

Based on (5) and (6), we can assign the lower $\pi_i(L)$ to R_j and the higher $\pi_j(L)$ to R_i without violating their IC^a constraints. Moreover, this switch does not affect the objective or any other voter's IC^a constraint. This shows that the sender is weakly better off after the switch. We have proved that there exists an optimal policy with the monotonicity property.

We next show that if IC^a-i binds, then $\pi_j(L) < \pi_i(L)$ if $\ell_j > \ell_i$. Suppose not: there exists R_j such that $\ell_j > \ell_i$ and $\pi_j(L) \geq \pi_i(L)$. Since IC^a-i binds, $\pi_i(L)$ is strictly positive. So is $\pi_j(L)$. Combining IC^a-j and the binding IC^a-i, we have

$$\frac{\ell_i \pi_i(L)}{\ell_j \pi_j(L)} \geq \frac{\frac{a_1+a_2\pi_j(L)}{a_2+a_3\pi_j(L)}}{\frac{a_1+a_2\pi_i(L)}{a_2+a_3\pi_i(L)}}.$$

On the other hand, given the presumption that $\pi_j(L) \geq \pi_i(L)$, it is easy to show that

$$\frac{\pi_i(L)}{\pi_j(L)} \leq \left(\frac{a_1 + a_2\pi_j(L)}{a_2 + a_3\pi_j(L)}\right) \Big/ \left(\frac{a_1 + a_2\pi_i(L)}{a_2 + a_3\pi_i(L)}\right).$$

This leads to the relation that $\pi_i(L)/\pi_j(L) \leq \ell_i\pi_i(L)/(\ell_j\pi_j(L))$, which contradicts the presumption that $\ell_j > \ell_i$. □

Proof for proposition 4.3. We first prove that among those voters who have an interior $\pi_i(L)$, at most one voter has a slack IC^a constraint. Suppose not. Suppose that $\ell_j > \ell_i$ and $0 < \pi_j(L) \leq \pi_i(L) < 1$. (From proposition 4.2, it is without loss to assume that the more lenient voter R_i has a higher $\pi_i(L)$.) Suppose that both IC^a-i

and IC^a-j are slack. Then we can find a pair of small positive numbers $(\varepsilon_1, \varepsilon_2)$ such that replacing $(\pi_i(L), \pi_j(L))$ with $(\pi_i(L)+\varepsilon_1, \pi_j(L)-\varepsilon_2)$ makes the sender strictly better off without violating any IC^a constraint. First, when $\varepsilon_1, \varepsilon_2$ are sufficiently small, IC^a-i and IC^a-j are satisfied since they were slack.

In order to leave IC^a-k intact for any $k \neq i, j$, we need the following ratio to be weakly higher:

$$\frac{\mathcal{S}(HHH)+\mathcal{S}(HLH)\pi_j(L)+\mathcal{S}(LHH)\pi_i(L)+\mathcal{S}(LLH)\pi_i(L)\pi_j(L)}{\mathcal{S}(HHL)+\mathcal{S}(HLL)\pi_j(L)+\mathcal{S}(LHL)\pi_i(L)+\mathcal{S}(LLL)\pi_i(L)\pi_j(L)}, \tag{7}$$

where $\mathcal{S}(\theta_i\theta_j\theta_k) := \sum_{\theta_{-ijk}\in\Theta_{-ijk}} f(\theta_i\theta_j\theta_k, \theta_{-ijk})\pi_{-ijk}(\theta_{-ijk})$. Given the exchangeability assumption, we have $\mathcal{S}(HHL) = \mathcal{S}(HLH) = \mathcal{S}(LHH)$ and $\mathcal{S}(HLL) = \mathcal{S}(LHL) = \mathcal{S}(LLH)$.

We next show that $\mathcal{S}(HHH)\mathcal{S}(LLL) > \mathcal{S}(HHL)\mathcal{S}(HLL)$ by applying the strict Ahlswede-Daykin inequality introduced and proved in part B.6 of the online appendix. We define f_1, f_2, f_3, f_4 as follows:

$$f_1(\theta_{-ijk}) = f(HHL, \theta_{-ijk})\pi_{-ijk}(\theta_{-ijk}), \quad f_2(\theta_{-ijk}) = f(HLL, \theta_{-ijk})\pi_{-ijk}(\theta_{-ijk}),$$
$$f_3(\theta_{-ijk}) = f(HHH, \theta_{-ijk})\pi_{-ijk}(\theta_{-ijk}), \quad f_4(\theta_{-ijk}) = f(LLL, \theta_{-ijk})\pi_{-ijk}(\theta_{-ijk}).$$

Suppose that among all voters other than i, j, k, only voters in $C \subseteq \{1, ..., n\} \setminus \{i, j, k\}$ have a fully revealing individual policy, i.e. $\pi_m^L(\theta_{-ijk}) = 0$ for any $m \in C$. Then, $\pi_{-ijk}(\theta_{-ijk}) \neq 0$ only for $\{\theta_{-ijk} \in \Theta_{-ijk} : \theta_m = H \text{ for all } m \in C\}] \equiv \mathcal{L}$. The set $\mathcal{L}$ forms a sublattice because: 1) it is nonempty, as $\theta^H_{-ijk} \in \mathcal{L}$, 2) for any $\theta_{-ijk}, \theta'_{-ijk} \in \mathcal{L}$, it is easily verified that $\theta_{-ijk} \vee \theta'_{-ijk} \in \mathcal{L}$ and $\theta_{-ijk} \wedge \theta'_{-ijk} \in \mathcal{L}$. Given that for any $\theta_{-ijk}, \theta'_{-ijk} \in \mathcal{L}$, the premise of the Ahlswede-Daykin inequality is satisfied strictly due to our assumption of strict affiliation for any three voters' states, i.e. $f_1(\theta_{-ijk})f_2(\theta'_{-ijk}) < f_3(\theta_{-ijk} \wedge \theta'_{-ijk})f_4(\theta_{-ijk} \vee \theta'_{-ijk}), \;$ for all $\theta_{-ijk}, \theta'_{-ijk} \in \mathcal{L}$, we can apply the strict Ahlswede-Daykin inequality with $A = B = \mathcal{L}$ to conclude that $f_1(\mathcal{L})f_2(\mathcal{L}) < f_3(\mathcal{L})f_4(\mathcal{L})$. But, because $\pi_{-ijk}(\theta_{-ijk} = 0$ for all $\theta_{-ijk} \in \Theta_{-ijk} \backslash \mathcal{L}$,

$$\begin{aligned}\mathcal{S}(\theta_i\theta_j\theta_k) &= \sum_{\theta_{-ijk}\in L} f(\theta_i\theta_j\theta_k, \theta_{-ijk})\pi_{-ijk}(\theta_{-ijk}) \\ &+ \sum_{\theta_{-ijk}\in\Theta_{-ijk}\backslash L} f(\theta_i\theta_j\theta_k, \theta_{-ijk})\pi_{-ijk}(\theta_{-ijk}) \\ &= \sum_{\theta_{-ijk}\in L} f(\theta_i\theta_j\theta_k, \theta_{-ijk})\pi_{-ijk}(\theta_{-ijk}).\end{aligned}$$

Therefore, $\mathcal{S}(HHH)\mathcal{S}(LLL) > \mathcal{S}(HHL)\mathcal{S}(HLL)$. Using the same method, we can prove that $\mathcal{S}(HHL)\mathcal{S}(LLL) > \mathcal{S}(HLL)^2$ and $\mathcal{S}(HHH)\mathcal{S}(HLL) > \mathcal{S}(HHL)^2$. We omit the details here. These inequalities imply that the ratio (7) strictly decreases in $\pi_i(L), \pi_j(L)$.

To make sure that the ratio (7) is higher after we replace $(\pi_i(L), \pi_j(L))$ with

$(\pi_i(L)+\varepsilon_1, \pi_j(L)-\varepsilon_2)$, we have to put a lower bound on ε_2 in terms of ε_1. After we substitute $\mathcal{S}(HLH)$, $\mathcal{S}(LHH)$ with $\mathcal{S}(HHL)$ and $\mathcal{S}(LHL)$, $\mathcal{S}(LLH)$ with $\mathcal{S}(HLL)$, the lower bound on ε_2 in terms of ε_1 can be written as:

$$\varepsilon_2 \geq \frac{\pi_j(L)(s_1s_4 - s_2s_3) + s_1s_3 - s_2^2 + (\pi_j(L))^2 (s_2s_4 - s_3^2)}{\pi_i(L)(s_1s_4 - s_2s_3) + s_1s_3 - s_2^2 + (\pi_i(L))^2 (s_2s_4 - s_3^2)}\varepsilon_1, \tag{8}$$

where $s_1 = \mathcal{S}(HHH),\ s_2 = \mathcal{S}(HHL),\ s_3 = \mathcal{S}(HLL),\ s_4 = \mathcal{S}(LLL)$.

Note that the lower bound on the RHS of (8) varies as we vary $k \neq i, j$. There exists a maximum lower bound. For the rest of the proof, we let R_k be the voter corresponding to this maximum lower bound.

The sender's payoff can be written as follows:

$$\{\mathcal{S}(HHH), \mathcal{S}(HHL), \mathcal{S}(HLH), \mathcal{S}(LHH), \mathcal{S}(HLL), \mathcal{S}(LHL), \mathcal{S}(LLH), \mathcal{S}(LLL)\}.$$
$$\{1, \pi_k(L), \pi_j(L), \pi_i(L), \pi_j(L)\pi_k(L), \pi_i(L)\pi_k(L), \pi_i(L)\pi_j(L), \pi_i(L)\pi_j(L)\pi_k(L)\}$$

To make sure that the sender's payoff is weakly higher after we replace $(\pi_i(L), \pi_j(L))$ with $(\pi_i(L)+\varepsilon_1, \pi_j(L)-\varepsilon_2)$, we have to impose an upper bound on ε_2 in terms of ε_1:

$$\varepsilon_2 \leq \frac{\pi_j(L)\left(\pi_k(L)s_4 + s_3\right) + \pi_k(L)s_3 + s_2}{\pi_i(L)\left(\pi_k(L)s_4 + s_3\right) + \pi_k(L)s_3 + s_2}\varepsilon_1.$$

In this expression, we have substituted $\mathcal{S}(HLH)$ with $\mathcal{S}(HHL)$, $\mathcal{S}(LHH)$ with $\mathcal{S}(LHL)$, and $\mathcal{S}(LLH)$ with $\mathcal{S}(HLL)$. The RHS decreases in $\pi_k(L)$, so this upper bound is the tightest when $\pi_k(L) = 1$:

$$\varepsilon_2 \leq \frac{(s_3 + s_4)\pi_j(L) + s_2 + s_3}{(s_3 + s_4)\pi_i(L) + s_2 + s_3}\varepsilon_1. \tag{9}$$

The upper bound in (9) is higher than the lower bound in (8), given that $s_1s_4 > s_2s_3$, $s_2s_4 > s_3^2$, and $s_1s_3 > s_2^2$. Therefore, we can find a pair $(\varepsilon_1, \varepsilon_2)$ such that the sender is better off after the change, and all the IC^a constraints are satisfied.

We then prove that among those voters who have interior $\pi_i(L)$, only the strictest voter might have a slack IC^a constraint. Suppose that $\ell_j > \ell_i$ and $0 < \pi_j(L) \leq \pi_i(L) < 1$. Suppose further that IC^a-i is slack and IC^a-j is binding. Then we can replace $(\pi_i(L), \pi_j(L))$ with $(\pi_i(L)+\varepsilon_1, \pi_j(L)-\varepsilon_2)$ where $(\varepsilon_1, \varepsilon_2)$ satisfy the two inequalities (8) and (9). Based on the argument above, this change makes the sender strictly better off without violating the IC^a constraints of the voters other than R_i and R_j. Moreover, since IC^a-i was slack, IC^a-i is satisfied after the change. IC^a-j is satisfied as well since we replace $\pi_j(L)$ with a more informative policy $\pi_j(L) - \varepsilon_2$ and the voters other than R_i and R_j are willing to be obedient if they condition on R_i's and R_j's approvals. It is easily verified that the decrease ε_2 required for IC^a-j to be satisfied is smaller than the decrease ε_2 that keeps the sender's payoff intact. □

Proof for proposition 4.4. Consider the auxiliary problem in which the sender needs only R_1's approval. The optimal policy is to set $\pi_1(L)$ as high as R_1's IC^a constraint allows: $\pi_1(L) = \sum_{\theta\in\Theta_i^H} f(\theta)/\left(\sum_{\theta\in\Theta_i^L} f(\theta)\ell_1\right)$. Conditioning on receiving an approval recommendation, R_1's posterior belief that her state is H is $\ell_1/(1+\ell_1)$. Let f' denote the state distribution such that

$$f'(\theta^H) = \sum_{\theta\in\Theta_i^H} f(\theta), \quad f'(\theta^L) = 1 - f'(\theta^H).$$

Given the single-voter policy above, the belief of any other voter $R_j \neq R_1$ about $\theta_j = H$ approaches R_1's posterior belief $\ell_1/(1+\ell_1)$, if $\|f - f'\|$ is small enough.[36] On the other hand, any $R_j \neq R_1$ is willing to approve the project if her belief of being H is above $\ell_j/(1+\ell_j)$ which is strictly smaller than $\ell_1/(1+\ell_1)$. This means that there exists $\varepsilon > 0$ such that every other $R_j \neq R_1$ is willing to rubber-stamp R_1's approval decision if $\|f - f'\| \leq \varepsilon$.

We then argue that when the above policy is incentive compatible, it is never optimal to fully reveal her payoff state to some voter. To the contrary, suppose that in the optimal policy $R_i \neq R_1$ learns about her state fully: $\pi_i(L) = 0$. Then the project is never approved when R_i's state is L. The sender's payoff is at most $\sum_{\theta\in\Theta_i^H} f(\theta)$, which is strictly below the payoff under the single-voter policy specified in the first paragraph. □

Proof for proposition 4.5. Let f' denote the state distribution such that voters' states are independent and each voter's state is H with probability $\sum_{\theta\in\Theta_i^H} f(\theta)$. We want to argue that if $\|f - f'\|$ is sufficiently small, all IC^a constraints bind. Suppose not. Suppose there exists R_i for which IC^a-i is slack. Let R_j be another voter whose IC^a-j constraint binds. Then,

$$\ell_i\pi_i(L) < \frac{a_1 + a_2\pi_j(L)}{a_2 + a_3\pi_j(L)}, \qquad \text{and } \ell_j\pi_j(L) = \frac{a_1 + a_2\pi_i(L)}{a_2 + a_3\pi_i(L)}.$$

The coefficients a_1, a_2, a_3 are as defined in the proof of proposition 4.2. When $\|f - f'\|$ is sufficiently small, the LHS of each voters' IC^a constraints becomes independent of the information of the other voters:

$$\lim_{\|f-f'\|\to 0} \frac{a_1 + a_2\pi_i(L)}{a_2 + a_3\pi_i(L)} = \frac{\sum_{\theta\in\Theta_i^H} f'(\theta)}{\sum_{\theta\in\Theta_i^L} f'(\theta)}.$$

If the sender increases $\pi_i(L)$, the LHS of IC^a-j will be only slightly affected if $\|f - f'\|$ is sufficiently small. Hence, $\pi_j(L)$ has to decrease only slightly so that IC^a-j continues to bind. For sufficiently independent states, i.e. for $\|f(\cdot) - f'(\cdot)\| < \varepsilon$

[36]Here, $\|f - f'\|$ denotes the Euclidean distance between f and f'.

for some small ε, setting all IC^a constraints binding strictly improves the payoff of the sender. $\square$

B Online appendix

B.1 Proofs for sections 5 and 6

Proof for lemma 5.1. (i) General persuasion: Let $(\pi(\hat{d}^a|\theta))_\theta$ be an optimal policy with n voters. Suppose a voter R_i is removed from the group, so there are only $(n-1)$ voters left. The distribution of states for the remaining voters is given by $\tilde{f} : \{H, L\}^{n-1} \to [0, 1]$ such that $\tilde{f}(\theta_{-i}) = f(\theta_{-i}, H) + f(\theta_{-i}, L)$. We construct $\tilde{\pi}(\hat{d}^a|\cdot) : \{H, L\}^{n-1} \to [0, 1]$ such that for any $\theta \in \{H, L\}^{n-1}$ and $\theta', \theta'' \in \{H, L\}^n$ with $\theta'_j = \theta''_j = \theta_j$ for $j \neq i$ and $\theta'_i = H$, $\theta''_i = L$:

$$\tilde{f}(\theta)\tilde{\pi}(\hat{d}^a|\theta) = f(\theta')\pi(\hat{d}^a|\theta') + f(\theta'')\pi(\hat{d}^a|\theta'').$$

Notice that the IC^a constraints of the remaining voters are satisfied, as

$$\frac{\sum_{\theta\in\{H,L\}^{n-1}:\theta_j=H} \tilde{f}(\theta)\tilde{\pi}(\hat{d}^a|\theta)}{\sum_{\theta\in\{H,L\}^{n-1}:\theta_j=L} \tilde{f}(\theta)\tilde{\pi}(\hat{d}^a|\theta)} = \frac{\sum_{\theta\in\{H,L\}^{n}:\theta_j=H} f(\theta)\pi(\hat{d}^a|\theta)}{\sum_{\theta\in\{H,L\}^{n}:\theta_j=L} f(\theta)\pi(\hat{d}^a|\theta)} \geq \ell_j.$$

By the same reasoning, the sender attains the same payoff under $\tilde{\pi}$ as under π. Hence, he is weakly better off with $(n-1)$ voters.

(ii) Individual persuasion: Consider a group of n voters, characterized by a threshold profile $\{\ell_1, ..., \ell_n\}$ with $\ell_i > \ell_{i+1}$ for any $i \leq n-1$. By proposition 4.2 there exists a monotone optimal policy. We let $(\pi_i(L))_{i=1}^n$ denote this policy. Suppose that the sender does not need some voter's approval, so we are left with $n-1$ voters. We relabel the remaining voters monotonically, so that R_1 is the strictest and R_{n-1} is the most lenient voter. In the first step, we assign the original $\pi_i(L)$ to the ith strictest voter among the remaining $n-1$ voters. Therefore R_n's original policy will be removed from the group.

Note that the remaining $n-1$ voters might not be willing to obey the approval recommendation given the policy $\{\pi_1(L), ..., \pi_{n-1}(L)\}$ if we remove R_n's original policy $\pi_n(L)$ from the group. This is because voter $R_j \neq R_n$ might be willing to approve only when she conditions the approvals by the other voters including R_n. We need to construct a new policy for the smaller group $\{R_1, ..., R_{n-1}\}$. The main step is to show that the informativeness of $\pi_n(L)$ can be loaded into the policies of the remaining voters so as to satisfy the remaining voters' IC^a constraints when $\pi_n(L)$ is removed from the policy profile.

Our first observation is that if $\pi_n(L) = 1$, then the policy $\{\pi_1(L), ..., \pi_{n-1}(L)\}$ is incentive compatible since each voter in the smaller group $\{R_1, ..., R_{n-1}\}$ is assigned a (weakly) more informative policy than before. Moreover, removing R_n's policy has no impact on the other voters' incentive constraints since she always approves. We focus on the case in which $\pi_n(L) < 1$. We want to show that we can adjust the policy $\{\pi_1(L), ..., \pi_n(L)\}$ in an incentive compatible way so as to increase $\pi_n(L)$ until it reaches one. At that point, removing $\pi_n(L)$ from the policy profile will not affect the IC^a constraints of $\{R_1, ..., R_{n-1}\}$.

Since R_n will eventually be removed, this is as if R_n's IC^a constraint is always slack. In proposition 4.3, we show that among those voters who have interior $\pi_i(L)$, only the strictest voter might have a slacking IC^a constraint. In our current setting, R_n's policy corresponds to the most lenient voter's policy and R_n's IC^a constraint is slack. Therefore, we follow the reasoning in the proof of proposition 4.3 to show that we can always increase $\pi_n(L)$ and decrease another voter's policy in an incentive-compatible way without decreasing the sender's payoff. In particular, if $\pi_{n-1}(L)$ is also interior, we can replace $(\pi_{n-1}(L), \pi_n(L))$ with $(\pi_{n-1}(L) - \varepsilon_2, \pi_n(L) + \varepsilon_1)$ so that the sender is better off and the IC^a constraints of $\{R_1, ..., R_{n-1}\}$ are satisfied. Since R_n's IC^a constraint is constantly slack, we can make this adjustment on the pair $(\pi_{n-1}(L), \pi_n(L))$ until either $\pi_n(L)$ reaches one or $\pi_{n-1}(L)$ reaches zero. If $\pi_n(L)$ reaches one before $\pi_{n-1}(L)$ drops to zero, we are done with the construction, because removing R_n from the group now will not affect the remaining voters' IC^a constraints. If $\pi_{n-1}(L)$ drops to zero before $\pi_n(L)$ reaches one, we can then adjust the pair $(\pi_{n-2}(L), \pi_n(L))$ by increasing $\pi_n(L)$ and decreasing $\pi_{n-2}(L)$ in a similar manner.

We keep making this adjustment until $\pi_n(L)$ reaches one: this is always possible if $\pi_i(L) > 0$ for some $i \in \{1, ..., n-1\}$. In other words, it cannot be that the policies of all other voters become fully revealing before $\pi_n(L) = 1$, as it takes an infinite amount of information from R_n's policy to push all other voters' policies to fully revealing ones. When $\pi_n(L)$ reaches one, removing R_n will not affect the remaining voters' IC^a constraints. If $\pi_j(L) = 0$ for $j \in \{1, ..., n-1\}$ in the original policy to begin with, removing R_n's policy will not affect the remaining voters' IC^a constraints since they all learn their states fully. This completes the construction. □

Proof for lemma 5.2. From the proof of proposition 4.4, we know that given a perfectly correlated distribution f', for sufficiently correlated states i.e. for some f within $\varepsilon > 0$ of f', in any optimal individual policy IC^a-1 is the only binding IC^a-i constraint. The sender achieves the same payoff as if he needed only R_1's approval.

We next argue that this is also the optimal policy under general persuasion.

From lemma 5.1, we know that the sender's payoff weakly decreases in the number of approvals he needs for a fixed threshold profile. Therefore, the sender's payoff cannot exceed the payoff from persuading only R_1. On the other hand, the sender can use the individual policy specified in the previous paragraph to achieve this payoff when the sender is allowed to use any general policy. Therefore, the probability of approval is equal for each θ across the two modes. □

Proof for propositions 6.1. Suppose that $\hat{d}$ is observed publicly; we want to show that each R_i continues to comply with $\hat{d}_i$. Consider first $\hat{d} \in \{\hat{d}^a, \{\hat{d}^{r,i}\}_i\}$. By incentive compatibility of the optimal policy π, it follows immediately that each R_i follows her own recommendation, i.e. she complies with $\hat{d}_i$. Suppose now that $\hat{d} \notin \{\hat{d}^a, \{\hat{d}^{r,i}\}_i\}$. If all other voters R_{-i} follow the recommendation, R_i's vote is not pivotal. For any such $\hat{d}$ with two or more rejections, the project is not approved. Hence, R_i is indifferent between $\hat{d}_i$ and the other available action, as both yield an ex-post payoff of zero. □

Proof for proposition 6.2. Consider first the case of individual persuasion. Take $(\pi_i)_i$ to be an optimal policy under simultaneous voting. Our first claim is that any policy that is incentive-compatible under simultaneous voting is incentive-compatible also under sequential voting, for a fixed voting order $\{1, ..., n\}$. Conversely, any policy that is incentive-compatible under sequential voting is incentive-compatible under simultaneous voting as well. It is sufficient to show that IC^a_{sim}-i, IC^a_{seq}-i, IC^r_{sim}-i, and IC^r_{seq}-i, are pairwise equivalent:

$$\left(\sum_{\Theta_i^H} f(\theta) \prod_{j \neq i} \pi_j(\theta_j)\pi_i(H)\right) \geq \ell_i \left(\sum_{\Theta_i^L} f(\theta) \prod_{j \neq i} \pi_j(\theta_j)\pi_i(L)\right), \qquad (IC^a_{sim}\text{-}i)$$

$$\left(\sum_{\Theta_i^H} f(\theta) \prod_{j=1}^{i-1} \pi_j(\theta_j)\pi_i(H) \prod_{j=i+1}^{n} \pi_j(\theta_j)\right) \geq \ell_i \left(\sum_{\Theta_i^L} f(\theta) \prod_{j=1}^{k-1} \pi_j(\theta_j)\pi_i(L) \prod_{j=k+1}^{n} \pi_j(\theta_j)\right), \qquad (IC^a_{seq}\text{-}i)$$

$$\left(\sum_{\Theta_i^H} f(\theta) \prod_{j \neq i} \pi_j(\theta_j)(1-\pi_i(H))\right) \geq \ell_i \left(\sum_{\Theta_i^L} f(\theta) \prod_{j \neq i} \pi_j(\theta_j)(1-\pi_i(L))\right), \qquad (IC^r_{sim}\text{-}i)$$

$$\left(\sum_{\Theta_i^H} f(\theta)\prod_{j=1}^{i-1}\pi_j(\theta_j)(1-\pi_i(H))\prod_{j=i+1}^{n}\pi_j(\theta_j)\right) \geq$$

$$\ell_i\left(\sum_{\Theta_i^L} f(\theta)\prod_{j=1}^{k-1}\pi_j(\theta_j)(1-\pi_i(L))\prod_{j=k+1}^{n}\pi_j(\theta_j)\right). \quad (IC^r_{seq}\text{-}i)$$

The constraints are pairwise equivalent for simultaneous and sequential voting. Therefore, the optimal policy under simultaneous voting is also optimal under sequential voting with order $\{1, ..., n\}$. We next argue that optimal policy in sequential voting is order-independent. Let $\{1, ..., n\}$ and δ such that $\delta(i) \neq i$ be two voting orders. Let us rewrite IC^a-i and IC^a-$\delta(i)$ in the following form respectively:

$$\left(\sum_{\Theta_i^H} f(\theta)\prod_{j=1}^{i-1}\pi_j(\theta_j)\pi_i(H)\prod_{j=i+1}^{n}\pi_j(\theta_j)\right) \geq$$

$$\ell_i\left(\sum_{\Theta_i^L} f(\theta)\prod_{j=1}^{k-1}\pi_j(\theta_j)\pi_i(L)\prod_{j=k+1}^{n}\pi_j(\theta_j)\right). \quad (\text{IC}^a\text{-}i)$$

$$\left(\sum_{\Theta_i^H} f(\theta)\prod_{j=1}^{\delta(i)-1}\pi_j(\theta_j)\pi_i(H)\prod_{j=\delta(i)+1}^{n}\pi_j(\theta_j)\right) \geq$$

$$\ell_i\left(\sum_{\Theta_i^L} f(\theta)\prod_{j=1}^{\delta(i)-1}\pi_j(\theta_j)\pi_i(L)\prod_{j=\delta(i)+1}^{n}\pi_j(\theta_j)\right). \quad (\text{IC}^a\text{-}\delta(i))$$

The set of other voters R_{-i} is the same despite the order of play for R_i. If R_i is offered the same policy in both sequences, IC^a-i and IC^a-$\delta(i)$ become equivalent to each-other and to:

$$\left(\sum_{\Theta_i^H} f(\theta)\pi_i(H)\Pr\left(R_{-i}\text{ approve}|\theta\right)\right) - \ell_i\left(\sum_{\Theta_i^L} f(\theta)\pi_i(L) f\left(R_{-i}\text{ approve}|\theta\right)\right) \geq 0.$$

Therefore the IC^a constraint for a voter R_i is order-independent. The objective of the sender is

$$\max_{\pi_i(H),\pi_i(L)}\sum_{\Theta} f(\theta)\prod_i \pi_i(\theta_i)$$

This objective is also order-independent (the product of approval probabilities is commutative). Therefore, if $((\pi_i(H), \pi_i(L)))_i$ is a solution to the original individual persuasion problem with order $\{1, ..., n\}$, $\{(\pi_{\delta(i)}(H), \pi_{\delta(i)}(L))\}_i$ is also a solution to the new problem with order δ.

The reasoning for general persuasion is very similar. First, it is straightforward that IC$^a_{sim}$-i is equivalent to IC$^a_{seq}$-i and IC$^r_{sim}$-i to IC$^r_{seq}$-i. Therefore the optimal simultaneous policy remains optimal under sequential voting with order $\{1, ..., n\}$. Secondly, each voter, despite her rank in the sequence, only cares about $\pi(\hat{d}^a|\theta)$ and $\pi(\hat{d}^{i,r}|\theta)$. Upon observing $(i-1)$ preceding approvals, R_i is sure that the generated recommendation is in $D^{i-1,a} = \{\hat{d} : \hat{d}_j = 1 \text{ for } j = 1, ..., i-1\}$. Yet she only cares about those elements in $D^{i-1,a}$ for which $\hat{d}_j = 1$ for voters in $\{R_{i+1}, ..., R_n\}$ as well. Hence, her IC^a is order-independent: $\left(\sum_{\Theta_i^H} f(\theta)\pi(\hat{d}^a|\theta)\right) \geq \ell_i \left(\sum_{\Theta_i^L} f(\theta)\pi(\hat{d}^a|\theta)\right)$. So is IC^r-$i$ as well. The objective of the sender is also order-independent:

$$\max_{\pi(\cdot|\theta)} \sum_{\Theta} f(\theta) \prod_i \pi(\hat{d}^a|\theta).$$

For any recommendation $\hat{d}$, let $\delta(\hat{d})$ be the permuted recommendation such that $\hat{d}_i = \hat{d}_{\delta(i)}$; in particular, $\delta(\hat{d}^a) = \hat{d}^a$. Therefore, if $(\pi(\hat{d}|\theta))_{\hat{d},\theta}$ is a solution to the original general persuasion problem with order $\{1, ..., n\}$, π' such that $\pi'(\delta(\hat{d})|\theta) = \pi(\hat{d}|\theta)$ for any $\hat{d}$ and θ is a solution to the general persuasion problem under δ. The sender's payoff is the same from both π and π'. □

Proof for proposition 6.3. Let $\hat{D}_i^a$ denote the set of recommendation profiles under which exactly k voters are recommended to approve and R_i is among them, and $\hat{D}_i^r$ the set of recommendation profiles under which exactly $k-1$ voters are recommended to approve and R_i is not among them. We design a full support policy as follows. For any i and any $\hat{d} \in \hat{D}_i^a$, we let

$$\Pr(\hat{d} \mid \theta) = \begin{cases} \varepsilon^2, & \text{if } \theta \neq \theta^H \\ \varepsilon, & \text{if } \theta = \theta^H. \end{cases}$$

Here, ε is a small positive number. This ensures that whenever R_i is recommended to approve, her belief of being H conditional on being pivotal (i.e., conditional on $\hat{d} \in \hat{D}_i^a$) is sufficiently high. So IC^a-i is satisfied. For any i and any $\hat{d} \in \hat{D}_i^r$, we let

$$\Pr(\hat{d} \mid \theta) = \begin{cases} \varepsilon^2, & \text{if } \theta \neq \theta^L \\ \varepsilon, & \text{if } \theta = \theta^L. \end{cases}$$

This ensures that whenever R_i is recommended to reject, her belief of being L conditional on being pivotal is sufficiently high. So IC^r-i is satisfied as well. For any recommendation profile $\hat{d} \neq \hat{d}^a$ such that $\hat{d} \notin \hat{D}_i^a \cup \hat{D}_i^r$ for any i, we let $\Pr(\hat{d} \mid \theta) = \varepsilon^2$. For each state profile θ, once we deduct the probabilities of the recommendation profiles specified above, the remaining probability is assigned to the unanimous approval recommendation $\hat{d}^a$. This construction ensures that the policy has full support. As ε goes to zero, the probability that the project is approved approaches one. □

Lemma B.1. *If f is exchangeable and affiliated, then either $f(\theta^H) + f(\theta^L) = 1$ or*

f has full support.

Proof. Since f is exchangeable, $f(\theta)$ depends only on the number of high-state voters in θ. For any $k \in \{0, ..., n\}$, we let p_k denote $f(\theta)$ when exactly k voters' states are high in θ. Due to affiliation, for any $2 \leq k \leq n$ we have $p_k p_{k-2} \geq p_{k-1}^2$.

We first show that if $f\left(\theta^H\right) = 0$ (that is $p_n = 0$), then $p_k = 0$ for any $k \geq 1$. This is because p_k being zero implies that p_{k-1} being zero for any $2 \leq k \leq n$. Given the presumption that $p_n = 0$, it must be true that $p_k = 0$ for $k \geq 1$. The only possibility is that $f(\theta^L) = 1$.

We then show that if $f\left(\theta^L\right) = 0$ (that is $p_0 = 0$), then $p_k = 0$ for any $k \leq n-1$. This is because p_{k-2} being zero implies that p_{k-1} being zero for any $2 \leq k \leq n$. The only possibility is that $f(\theta^H) = 1$.

Suppose that both $f\left(\theta^H\right)$ and $f\left(\theta^L\right)$ are strictly positive. We next show that either $p_k = 0$ for all $1 \leq k \leq n-1$ or $p_k > 0$ for all $1 \leq k \leq n-1$. Suppose there exists some k' such that $p_{k'} = 0$. Then, applying the inequality that $p_k p_{k-2} \geq p_{k-1}^2$, we conclude that p_k must be zero for any $1 \leq k \leq n-1$. Therefore, either f has full support or $f\left(\theta^H\right) + f\left(\theta^L\right) = 1$. □

Proof for proposition 6.4. Suppose that the sender needs k approvals. We divide the set of the state profiles Θ into three subsets. The first subset is denoted by Θ^k, which contains all the state profiles such that exactly k voters' states are H. The second subset is $\{\theta^L\}$, which contains a unique state profile such that all voters' states are L. The third subset includes the rest of the state profiles, i.e., $\Theta \setminus \left(\Theta^k \cup \{\theta^L\}\right)$. For any $\theta \in \Theta^k$, we let

$$\pi_i(\theta) = \begin{cases} 1 - \varepsilon_1, & \text{if } \theta_i = H \\ \varepsilon_2, & \text{if } \theta_i = L. \end{cases}$$

For θ^L, we let $\pi_i(\theta^L) = 1 - \varepsilon_3$. For any $\theta \in \Theta \setminus \left(\Theta^k \cup \{\theta^L\}\right)$, we let $\pi_i(\theta) = 1 - \varepsilon_4$. We first show that IC^a-i is satisfied. For any $\theta \in \Theta^k \cap \Theta_i^H$, the probability that exactly k voters including R_i are recommended to approve is $(1-\varepsilon_1)^k(1-\varepsilon_2)^{n-k} + \mathcal{O}\left((1-\varepsilon_1)^{k-1}(1-\varepsilon_2)^{n-k-1}\varepsilon_1\varepsilon_2\right)$. For any $\theta \in \Theta^k \cap \Theta_i^L$, the probability that exactly k voters including R_i are recommended to approve is $\mathcal{O}\left((1-\varepsilon_1)^{k-1}(1-\varepsilon_2)^{n-k-1}\varepsilon_1\varepsilon_2\right)$. For θ^L, the probability that exactly k voters including R_i are recommended to approve is $\mathcal{O}\left((1-\varepsilon_3)^k\varepsilon_3^{n-k}\right)$. Similarly, for any $\theta \in \Theta \setminus \left(\Theta^k \cup \{\theta^L\}\right)$, the probability that exactly k voters including R_i are recommended to approve is $\mathcal{O}\left((1-\varepsilon_4)^k\varepsilon_4^{n-k}\right)$. When $\varepsilon_1, \varepsilon_2, \varepsilon_3, \varepsilon_4$ are small enough, R_i puts most of the weight on the event that $\theta \in \Theta^k \cap \Theta_i^H$ when she is recommended to approve and she conditions on being pivotal. It is obvious that R_i is willing to approve since her state is H.

We next show that IC^r-i is satisfied as well. For any $\theta \in \Theta^k \cap \Theta_i^H$, the probability that exactly $k-1$ voters excluding R_i are recommended to approve is $\mathcal{O}\left((1-\varepsilon_1)^{k-1}\varepsilon_1(1-\varepsilon_2)^{n-k}\right)$. For any $\theta \in \Theta^k \cap \Theta_i^L$, the probability that exactly $k-1$ voters excluding R_i are recommended to approve is $\mathcal{O}\left((1-\varepsilon_1)^{k-1}\varepsilon_1(1-\varepsilon_2)^{n-k}\right)$. For θ^L, the probability that exactly $k-1$ voters excluding R_i are recommended to approve is $\mathcal{O}\left((1-\varepsilon_3)^{k-1}\varepsilon_3^{n-k+1}\right)$. Similarly, for any $\theta \in \Theta \setminus \left(\Theta^k \cup \{\theta^L\}\right)$, the probability that exactly $k-1$ voters excluding R_i are recommended to approve is $\mathcal{O}\left((1-\varepsilon_4)^{k-1}\varepsilon_4^{n-k+1}\right)$. If we let $\varepsilon_1, \varepsilon_4$ approach zero at a much faster rate than ε_3, R_i puts most of the weight on the event that $\theta = \theta^L$ when she is recommended to reject and she conditions on being pivotal. It is obvious that R_i is willing to reject since her state is L given θ^L. This completes the construction. □

Proof for proposition 6.5. Suppose there exists an individual policy $(\pi_i(H), \pi_i(L))_{i=1}^n$ which is the limit of a sequence of full-support incentive-compatible policies and ensures that the project is approved for sure. We first argue that $\pi_i(H) \geq \pi_i(L)$. Pick any full-support policy $(\tilde{\pi}_i(H), \tilde{\pi}_i(L))_{i=1}^n$ along the sequence. The following IC^a and IC^r constraints must hold for each R_i:

$$\Pr(\theta_i = H | k-1 \text{ approve})\tilde{\pi}_i(H) \geq \Pr(\theta_i = L | k-1 \text{ approve})\ell_i \tilde{\pi}_i(L),$$
$$\Pr(\theta_i = H | k-1 \text{ approve})(1-\tilde{\pi}_i(H)) \leq \Pr(\theta_i = L | k-1 \text{ approve})\ell_i (1-\tilde{\pi}_i(L)),$$

we obtain that $\tilde{\pi}_i(H)/\tilde{\pi}_i(L) \geq (1-\tilde{\pi}_i(H))/(1-\tilde{\pi}_i(L))$. This implies that $\tilde{\pi}_i(H) \geq \tilde{\pi}_i(L)$ for each R_i. This must hold for all policies along the sequence, so $\pi_i(H) \geq \pi_i(L)$ for each R_i.

If the project is approved for sure. It is approved for sure when $\theta = \theta^L$. Therefore, there exist k voters who approve with certainty when their states are L. Given that $\pi_i(H) \geq \pi_i(L)$ for each R_i, these voters also approve with certainty when their states are H. Thus, at least k voters approve the project for sure in both states. Let R^a be the set of voters who approve for sure: suppose there is exactly k such voters. We want to show that any voter $R_i \in R^a$ prefers to reject the project when she is recommended to approve. Conditional on being pivotal, R_i knows that all the voters not in R^a have rejected while all the voters in $R^a \setminus \{R_i\}$ have approved. R_i does not get more optimistic about her state being H from the approvals by $R^a \setminus \{R_i\}$ since these voters approve regardless of their states. R_i become more pessimistic about her state being H from the rejections by voters not in R^a. Therefore, R_i's posterior belief of being H conditional on being pivotal is lower than her prior belief. R_i strictly prefers to reject. Contradiction.

The case in which more than k voters are in R^a can be analyzed in a similar manner. Suppose that there are $k' > k$ voters in R^a. This policy is the limit of a sequence of full-support incentive-compatible policies. There exists $M \in (0,1)$ such that voters not in R^a approve with probability less than M in either state H or L, i.e., $\min\{\pi_i(H), \pi_i(L)\} = \pi_i(L) \leq M$. For any small $\varepsilon > 0$, there is a full-support policy such that voters in R^a approve with probability above $1-\varepsilon$ in both states, i.e., $\min\{\pi_i(H), \pi_i(L)\} = \pi_i(L) \geq 1-\varepsilon$. Pick any $R_i \in R^a$. When R_i is recommended to approve, there are two types of events in which she is pivotal: (i) $k-1$ voters in R^a approve and the rest reject; (ii) $k'' < k-1$ voters in R^a approve, $k-1-k''$ voters which are not in R^a approve, and the rest reject. As ε converges to zero, for any event of type (ii), there exists an event of type (i) which is much more likely to occur. This is because voters outside R^a are much more likely to reject than those in R^a. Therefore, the belief of R_i about her state being H is mainly driven by events of type (i). Note that in these events only voters in R^a approve. R_i does not get more optimistic about her state being H from the approvals by voters in $R^a \setminus \{R_i\}$ since these voters approve regardless of their states. R_i become more pessimistic about her state being H from the disapprovals by voters not in R^a. Therefore, R_i's posterior belief of being H conditional on being pivotal is either arbitrarily close to or smaller than her prior, so she strictly prefers to reject. Therefore R_i does not obey her approval recommendation. This shows that the project cannot be approved for sure.

We next show that each voter's payoff under individual persuasion is higher than that under general persuasion. We have shown that for each R_i, $\pi_i(H) \geq \pi_i(L)$. We next show that there exists at least one voter such that the above inequality is strict. Suppose not. Then each voter approves with the same probability in both state H and L. Therefore, each voter's posterior belief, when she conditions on being pivotal, is the same as her prior belief. It is not incentive compatible to obey the approval recommendation since each voter's own policy is uninformative as well. Therefore, at least one voter approves strictly more frequently in state H than in state L. We index this voter by i.

We next show that R_i is pivotal with strictly positive probability. Suppose not. Then for each state profile, either (i) at least k voters out of the rest $n-1$ voters approve for sure, or (ii) at least $n-k+1$ voters out of the rest $n-1$ voters reject for sure. We first argue that, if f has full support, it is not possible that case (i) holds for some state profiles, and case (ii) holds for the others. Suppose not. Suppose that there exists a state profile such that exactly $k' \geq k$ voters out of the rest $n-1$ voters approve for sure. There also exists another state profile such that exactly $k'' \geq n-k+1$ out of the rest $n-1$ voters disapprove for sure. The intersection of these two sets of voters who approve or disapprove for sure include at least $k' + k'' - (n-1)$ voters. These voters approve under one state and disapprove under the other. It is easily verified that $k' + k'' - (n-1) \geq k' - k + 2$. If we begin with the previous state profile under which exactly k' voters approve, we can flip the states and the decisions of $k' - k + 1$ such voters so that exactly $k - 1$ voters approve for sure. This makes R_i's decision pivotal, contradicting the presumption that R_i is never pivotal. Therefore, it has to be true that either case (i) holds for all state profiles or case (ii) holds for all state profiles. If f does not have full support (or equivalently the voters' states are perfectly correlated), then it is possible that case (i) holds for one state profile and case (ii) holds for the other. Given that $\pi_i(H) \geq \pi_i(L)$ for each R_i, the only possibility is that case (i) holds for θ^H and case (ii) holds for θ^L. Note that each voter obtains the highest possible payoff which is strictly positive. The statement of the proposition holds. Therefore, we can focus on the case in which either case (i) holds for both θ^H, θ^L or case (ii) holds for both θ^H, θ^L.

If case (ii) holds for all θ, this is clearly sub-optimal for the sender since he obtains a payoff of zero. Therefore, we assume that case (i) holds for all θ. However, we have argued previously that the project is not approved for sure. Contradiction. Therefore, R_i is pivotal with strictly positive probability. For R_i, a high state project is more likely to be approved than a low one. Due to affiliation, other voters benefit from the selection effect of R_i.

This shows that for any voter, a high state project is more likely to be approved than a low state one. Individual persuasion is better than general persuasion. □

B.2 Optimal policy when assumption 1 fails

Our results remain intact when we add voters who prefer to approve ex ante. Once the sender designs the optimal policy for those voters who are reluctant to approve, those who prefer to approve ex ante become more optimistic about their own states. The sender simply recommends that they approve all the time. This is the case under both general and individual persuasion.

Proposition B.1. *Suppose that voter $n+1$ prefers to approve ex ante. Under both general and individual persuasion, given the optimal policy for the first n voters, voter $n+1$ is willing to rubber-stamp the first n voters' approval decision.*

Proof of proposition B.1. This result holds for individual persuasion since each voter is weakly more optimistic about her state conditional on the others' approval. Therefore, voter $n+1$ is more optimistic about her state when she conditions on the others' approvals. We thus focus on general persuasion. There are $n+1$ voters.

Let $\tilde{f}$ be the distribution of these voters' states. The last voter is *convinced*, i.e., she approves given the prior belief $\tilde{f}$:

$$\ell_{n+1} \leq \frac{\sum_{\theta \in \{H,L\}^n} \tilde{f}(\theta, H)}{\sum_{\theta \in \{H,L\}^n} \tilde{f}(\theta, L)}. \tag{10}$$

Let f denote the distribution of the first n voters' states. We thus have $f(\theta) = \tilde{f}(\theta, H) + \tilde{f}(\theta, L)$. Let $(\pi(\hat{d}^a|\theta))_{\theta \in \{H,L\}^n}$ be an optimal policy with the first n voters. We want to construct a policy for $n+1$ voters $(\tilde{\pi}(\hat{d}^a|\tilde{\theta}))_{\tilde{\theta} \in \{H,L\}^{n+1}}$ such that the sender's payoff stays the same as when he faces only n voters. To ease notation, we write $\pi(\theta)$ for $\pi(\hat{d}^a|\theta)$ and $\tilde{\pi}(\tilde{\theta})$ for $\tilde{\pi}(\hat{d}^a|\tilde{\theta})$.

We let $\tilde{\pi}(\theta, H) = \tilde{\pi}(\theta, L) = \pi(\theta)$. Under this policy, the convinced voter approves whenever the first n voters approve. Then, (i) the sender's payoff stays the same; (ii) the first n voters' IC constraints are still satisfied. We only need to show that the convinced voter's IC constraint is satisfied, i.e.,

$$\ell_{n+1} \leq \frac{\sum_{\theta \in \{H,L\}^n} \tilde{f}(\theta, H)\tilde{\pi}(\theta, H)}{\sum_{\theta \in \{H,L\}^n} \tilde{f}(\theta, L)\tilde{\pi}(\theta, L)} = \frac{\sum_{\theta \in \{H,L\}^n} \tilde{f}(\theta, H)\pi(\theta)}{\sum_{\theta \in \{H,L\}^n} \tilde{f}(\theta, L)\pi(\theta)}. \tag{11}$$

Next, we want to show that the RHS of (11) is larger than that of (10). This will complete the proof.

Given the optimal policy $\pi(\cdot)$ for n voters, we let π^k be the average probability of unanimous approval for state profiles with k high states:

$$\pi^k := \frac{\sum_{|\theta|=k} f(\theta)\pi(\theta)}{\sum_{|\theta|=k} f(\theta)} = \frac{\sum_{|\theta|=k} \pi(\theta)}{\#\{\theta : |\theta| = k\}}.$$

Here $|\theta|$ is the number of high states in θ. We want to show that $\pi^k \geq \pi^{k-1}$ for $k \in \{1, ..., n\}$, that is, a profile with more high states is *on average* more likely to be approved than that with fewer high states. This, combined with the affiliation assumption, implies that the RHS of (11) is higher than that of (10). Intuitively, if $\pi^k \geq \pi^{k-1}$ for all $k \in \{1, ..., n\}$, then the convinced voter is more optimistic about her state when she conditions on the approval decision by the first n voters.

We illustrate the argument for $\pi^k \geq \pi^{k-1}$ through an example. Suppose that $n = 3$, $k = 2$. We first argue that

$$\begin{aligned} &\pi(HHL) \geq \pi(HLL), \quad \pi(HLH) \geq \pi(HLL), \quad \pi(LHH) \geq \pi(LHL), \\ &\pi(HHL) \geq \pi(LHL), \quad \pi(HLH) \geq \pi(LLH), \quad \pi(LHH) \geq \pi(LLH). \end{aligned} \tag{12}$$

Suppose not. Suppose that, for instance, $\pi(HHL) < \pi(HLL)$. We can increase $\pi(HHL)$ by $\epsilon > 0$ and decrease $\pi(HLL)$ by $\epsilon' > 0$ such that

$$f(HHL)\epsilon = f(HLL)\epsilon'.$$

This change will not affect the sender's payoff. It will not affect the incentive of any voter who has the same state in HHL and HLL, i.e., R_1's and R_3's IC constraints are still satisfied. The change will only make R_2's IC constraint easier to satisfy. Summing up inequalities in (12) and simplifying, we have shown that $\pi^k \geq \pi^{k-1}$ for $k = 2$. The argument for any n and $k \in \{1, ..., n\}$ is similar. This completes the proof. □

B.3 Full-support direct obedient policies

Direct obedient policies

Let $\pi : \Theta \to \Delta\left(\prod_{i=1}^n S_i\right)$ be an arbitrary information policy. A mixed strategy for R_i in the induced voting game is given by $\sigma_i : \Pi \times S_i \to \Delta(\{0,1\})$. Let $\sigma_i(s_i) := \sigma_i(\pi, s_i)$ denote the probability that $d_i = 1$ upon observing s_i. We claim that given any policy π and any profile of equilibrium strategies $(\sigma_i)_{i=1}^n$ we can construct a direct obedient policy that implements the same outcome as the original policy. Consider the direct policy $\tilde{\pi} : \Theta \to \Delta(\{0,1\}^n)$ such that for any $\hat{d}$ and any $\theta \in \Theta$,

$$\tilde{\pi}(\hat{d}|\theta) = \sum_s \pi(s|\theta) \prod_{i=1}^n \left(\hat{d}_i \sigma_i(s_i) + (1 - \hat{d}_i)(1 - \sigma_i(s_i))\right).$$

Claim 1. *The direct policy π satisfies IC^a-i and IC^r-i for any i.*

Proof. Let us first show that voter R_i obeys an approval recommendation:

$$\begin{aligned}
&\sum_{\Theta_i^H} f(\theta)\tilde{\pi}(\hat{d}^a|\theta) - \ell_i \sum_{\Theta_i^L} f(\theta)\tilde{\pi}(\hat{d}^a|\theta) \\
=& \sum_{\Theta_i^H} f(\theta)\Big(\sum_s \pi(s|\theta)\prod_{i=1}^n \sigma_i(s_i)\Big) - \ell_i \sum_{\Theta_i^L} f(\theta)\Big(\sum_s \pi(s|\theta)\prod_{i=1}^n \sigma_i(s_i)\Big) \\
=& \sum_s \prod_i \sigma_i(s_i) \left\{\sum_{\Theta_i^H} f(\theta)\pi(s|\theta) - \ell_i \sum_{\Theta_i^L} f(\theta)\pi(s|\theta)\right\} \\
=& \sum_{s_i \in S_i} \sigma_i(s_i) \left(\sum_{s_{-i}} \prod_{j \neq i} \sigma_j(s_j) \left\{\sum_{\Theta_i^H} f(\theta)\pi(s_i, s_{-i}|\theta) - \ell_i \sum_{\Theta_i^L} f(\theta)\pi(s_i, s_{-i}|\theta)\right\}\right) \geq 0.
\end{aligned}$$

The last inequality follows from the fact that if $\sigma_i(s_i) \in (0,1)$ (resp., $\sigma_i(s_i) = 1$) for some s_i then the term in brackets is zero (resp., strictly positive). Hence IC^a-i is satisfied. Similar reasoning shows that IC^r constraints are satisfied as well by $\tilde{\pi}$,

$$\begin{aligned}
&\sum_{\Theta_i^H} f(\theta)\tilde{\pi}(\hat{d}^{r,i}|\theta) - \ell_i \sum_{\Theta_i^L} f(\theta)\tilde{\pi}(\hat{d}^{r,i}|\theta) \\
=& \sum_{s_i \in S_i} (1 - \sigma_i(s_i)) \left(\sum_{s_{-i}} \prod_{j \neq i} \sigma_j(s_j) \left(\sum_{\Theta_i^H} f(\theta)\pi(s_i, s_{-i}|\theta) - \ell_i \sum_{\Theta_i^L} f(\theta)\pi(s_i, s_{-i}|\theta)\right)\right).
\end{aligned}$$

If $\sigma_i(s_i) \in (0,1)$ (resp., $\sigma_i(s_i) = 0$) for some s_i, the term in brackets is zero (resp., strictly negative). Therefore, $\tilde{\pi}$ is obedient. □

Full-support policies

Our analysis focuses on policies that are the limit of some sequence of full-support obedient policies. For this purpose, let us define a full-support policy for each mode of persuasion. A general policy π is a full-support general policy if $\pi(\hat{d}|\theta) \in (0,1)$ for any $\hat{d} \in \{0,1\}^n$ and for any $\theta \in \Theta$. This means that for any given state profile, all recommendation profiles are sent with strictly positive probability. Similarly, a full-support independent general policy $(\pi_i(\cdot))_i$ is such that for each voter R_i, $\pi_i(\theta) \in (0,1)$ for any $\theta \in \Theta$. That is, in each state profile, each voter is recommended to approve and to reject with strictly positive probability. An individual policy $(\pi_i))_i$ has full support if for each R_i, the probability of R_i receiving each recommendation for each individual state is bounded away from zero and one, i.e. for each R_i, $\pi_i(\theta_i) \in (0,1)$ for each $\theta_i \in \{H, L\}$. That is, for each individual state θ_i, each voter is recommended to approve and to reject with strictly positive probability.

B.4 Individual persuasion with two voters

The optimal policy when the states are perfectly correlated or independent is given by proposition 4.1, so here we focus on imperfectly correlated states. Without loss, it is assumed that $\ell_1 > \ell_2$. Based on lemma 4.1 and lemma 4.2, the optimal policy features $\pi_1(H) = \pi_2(H) = 1$ and at least one voter has a binding IC^a constraint.

We first want to argue that the stricter voter's IC^a-1 constraint must bind. Suppose not. Then IC^a-2 must bind. We can solve for $\pi_2(L)$ from this binding constraint:

$$\pi_2(L) = \frac{f(LH)\pi_1(L) + f(HH)\pi_1(H)}{\ell_2(f(HL)\pi_1(H) + f(LL)\pi_1(L))}.$$

Substituting $\pi_1(H) = \pi_2(H) = 1$ and $\pi_2(L)$ into the objective of the sender, we obtain the following objective:

$$\frac{f(HH)(1+\ell_2) + f(LH)\pi_1(L)(1+\ell_2)}{\ell_2},$$

which strictly increases in $\pi_1(L)$. Therefore, the sender finds it optimal to set $\pi_1(L)$ as high as possible. This means that either IC^a-1 constraint binds or $\pi_1(L) = 1$. Given the presumption that IC^a-1 constraint does not bind, it has to be the case that $\pi_1(L) = 1$. This policy essentially provides the more lenient voter with an informative policy and asks the stricter voter to approve with probability 1. We argue that this policy cannot be incentive-compatible, because IC^a-2 binds but R_1 is asked to rubber-stamp. Given that R_1 is stricter than R_2, R_1 learns about her state indirectly from R_2's approval, and whenever R_2 is indifferent between approving and not, R_1 must strictly prefer to reject. This shows that in the optimal policy IC^a-1 must bind.

Given that IC^a-1 always binds, we can solve for $\pi_1(L)$ as a function of $\pi_2(L)$. Substituting $\pi_1(H) = \pi_2(H) = 1$ and $\pi_1(L)$ into the sender's objective, we obtain that the objective strictly increases in $\pi_2(L)$. Therefore, the sender sets $\pi_2(L)$ as high as possible. Either IC^a-2 binds or $\pi_2(L) = 1$. We summarize the discussion above in the following lemma:

Lemma B.2. *Given two voters $\ell_1 > \ell_2$ whose states are imperfectly correlated, the optimal policy is unique. The stricter voter's IC^a-1 binds. The more lenient voter's $\pi_2(L)$ is as high as IC^a-2 allows.*

B.5 Individual persuasion with homogeneous thresholds

Consider individual persuasion under the unanimous rule when n voters have the same thresholds ℓ. Without loss, we assume that $\pi_i(L) \leq \pi_{i+1}(L)$ for $1 \leq i \leq n-1$. In any optimal policy, the voter(s) with the highest $\pi_i(L)$ must have binding IC^a constraints. The rest have slack IC^a constraints. We also assume that $\ell > 1$.

We focus on the following class of distributions f. Nature first draws a grand state which can be either G or B. If the state is G, each voter's state is H with probability $\lambda_1 \in (1/2, 1)$. If the state is B, each voter's state is L with probability $(1-\lambda_1) \in (0, 1/2)$.[37] Conditional on the grand state, the voters' states are drawn independently. We let f_k denote the probability of a state profile with k low-state voters. We thus have $f_k = p_0\lambda_1^{n-k}(1-\lambda_1)^k + (1-p_0)(1-\lambda_1)^{n-k}\lambda_1^k$.

We first argue that the support of $(\pi_i(L))_{i=1}^n$ has at most three elements. Suppose not. Then there exist i and j such that $\pi_i(L), \pi_j(L) \in (0, 1)$ and both IC^a-i and IC^a-j are slack. This violates proposition 4.3.[38] This argument also shows that the support of $(\pi_i(L))_{i=1}^n$ has at most two interior-valued elements.

Therefore, any optimal policy can be characterized by three numbers (n_0, y, x) with $n_0 \geq 0$ and $0 < y \leq x \leq 1$ such that: (i) $\pi_i(L) = 0$ for $i \in \{1, ..., n_0\}$; (ii) $\pi_{n_0+1}(L) = y$; (iii) $\pi_i(L) = x$ for $i \in \{n_0 + 2, ..., n\}$.[39] Only the IC^a constraints of

[37] If $\lambda_1 = 1$, voters' states are perfectly correlated. If $\lambda_1 = 1/2$, voters' states are independent. Both these polar cases have been addressed in proposition 4.1.

[38] It is easy to verify that the states of any three voters are strictly affiliated for this class of distributions.

[39] If $y \neq x$, only one voter has the policy y. If instead two or more voters had y, at least two voters would have interior $\pi_i(L)$ and slack IC^a constraints, contradicting proposition 4.3.

voters with $\pi_i(L) = x$ bind:

$$\ell = \frac{\sum_{k=0}^{n-n_0-2} C_{n-n_0-2}^k f_k x^k + \sum_{k=0}^{n-n_0-2} C_{n-n_0-2}^k f_{k+1} x^k y}{\sum_{k=0}^{n-n_0-2} C_{n-n_0-2}^k f_{k+1} x^{k+1} + \sum_{k=0}^{n-n_0-2} C_{n-n_0-2}^k f_{k+2} x^{k+1} y}. \tag{IC^a}$$

The sender's payoff is the sum of the numerator and the denominator of the right-hand side.

We let λ_k^* denote the value of λ_1 such that conditional on k voters' states being high, a voter is just willing to rubber-stamp. In other words, conditional on k voters' states being high, a voter's belief of being high is exactly $\ell/(1+\ell)$. Therefore, $\lambda_1 = \lambda_k^*$ solves the following equation:

$$\frac{p_0\lambda_1^k}{p_0\lambda_1^k + (1-p_0)(1-\lambda_1)^k}\lambda_1 + \frac{(1-p_0)(1-\lambda_1)^k}{p_0\lambda_1^k + (1-p_0)(1-\lambda_1)^k}(1-\lambda_1) = \frac{\ell}{1+\ell}.$$

We let p_h denote the ex ante probability of being H, that is, $p_h = p_0\lambda_1 + (1-p_0)(1-\lambda_1)$. The domain of p_h is $(1-\lambda_1, \lambda_1)$. Assumption 1 that no voter prefers to approve ex ante is equivalent to $p_h - \ell(1-p_h) < 0$. Substituting $p_0 = \frac{p_h-(1-\lambda_1)}{2\lambda_1-1}$ into the equation above, we obtain the following equation that defines λ_k^*:

$$\frac{\lambda_1^{k+1}(p_h-(1-\lambda_1)) + (1-\lambda_1)^{k+1}(\lambda_1-p_h)}{\lambda_1^k(p_h-(1-\lambda_1)) + (1-\lambda_1)^k(\lambda_1-p_h)} = \frac{\ell}{1+\ell}.$$

Note that λ_k^* depends on p_h, ℓ, k but not on n. The left-hand side increases in p_h, k, an λ_1.[40] The right-hand side increases in ℓ. Therefore, λ_k^* decreases in p_h and k, and it increases in ℓ.

The sequence $(\lambda_k^*)_{k=1}^\infty$ is a decreasing sequence which converges to $\ell/(1+\ell)$. Each voter learns about the grand state from the information regarding other voters' states. If $\lambda_1 < \ell/(1+\ell)$, even if a voter is certain that the grand state is G, this voter is not willing to rubber-stamp the project. Therefore, if $\lambda_1 \leq \ell/(1+\ell)$, no matter how many other voters' states are high, a voter is unwilling to rubber-stamp. On the other hand, for any $\lambda_1 > \ell/(1+\ell)$, there exists $k \geq 1$ such that $\lambda_1 > \lambda_k^*$. We next argue that, if $\lambda_1 > \lambda_k^*$, $n_0 < k$ in any optimal policy. Because a voter is willing to rubber-stamp if $k-1$ voters have $\pi_i(L) = 0$ and another voter has $\pi_i(L) \in (0,1)$, the sender can strictly improve his payoff if at least k voters learn their states fully. For instance, if $\lambda_1 > \lambda_1^*$, the sender is able to persuade R_2 to R_n to rubber-stamp even if he only partially reveals the state to R_1. No voter will learn her state fully, so $n_0 = 0$. For $\lambda_1 > \lambda_1^*$, we are left with two possible cases: either $y < x$ or $y = x$. In the former case, R_1's IC^a is slack. The sender provides more precise information to R_1 in order to persuade the other voters more effectively. In the latter case, all voters' IC^a constraints bind. In general, if $\lambda_1 \in (\lambda_k^*, \lambda_{k-1}^*]$, we must have $n_0 < k$.

[40] We fix the probability p_h that a voter's state is high, so varying λ_1 only varies the correlation of the states across voters.

We next show that it is not possible that both y and x are interior and they are not equal to each other.

Lemma B.3. *The support of* $(\pi_i(L))_{i=1}^n$ *does not have two interior values.*

Proof. The first $n_0 \geq 0$ voters learn their states fully. Voter R_{n_0+1}'s policy is $\pi_{n_0+1}(L) = y$. Subsequent voters from R_{n_0+2} to R_n have the policy $\pi_i(L) = x$. Suppose that both y and x are interior and $y < x$. We want to show that the sender can strictly improve his payoff.

The sender's payoff can be written as

$$\frac{c_0(x)(1-\lambda_1)^{n_0}(\lambda_1 - p_h)(1-\lambda_1+\lambda_1 x)(1-\lambda_1+\lambda_1 y)}{2\lambda_1 - 1}$$
$$+\frac{c_1(x)\lambda_1^{n_0}(p_h - 1 + \lambda_1)(\lambda_1 + (1-\lambda_1)x)(\lambda_1 + (1-\lambda_1)y)}{2\lambda_1 - 1},$$

where

$$c_0(x) = (1-\lambda_1+\lambda_1 x)^{n-n_0-2}, \quad c_1(x) = (\lambda_1 + (1-\lambda_1)x)^{n-n_0-2}.$$

The binding IC^a constraint can be written as:

$$\frac{c_0(x)}{c_1(x)} = \frac{\lambda_1^{n_0}(\lambda_1 + p_h - 1)(\lambda_1 + (1-\lambda_1)y)(\ell(\lambda_1 - 1)x + \lambda_1)}{(1-\lambda_1)^{n_0}(\lambda_1 - p_h)(\lambda_1(y-1)+1)(\ell\lambda_1 x + \lambda_1 - 1)}.$$

Given $(\ell, p_h, \lambda_1, n, n_0)$, the binding IC^a constraint implicitly defines y as a function of x. The domain of x depends on the values of $(\ell, p_h, \lambda_1, n, n_0)$. The lowest value that x can take is denoted $\underline{x}$, which is obtained when we set y to be equal to x. The highest value of x is denoted by $\overline{x}$, which is either equal to 1 or obtained by setting y to be zero. The domain of x is $[\underline{x}, \overline{x}]$.

Suppose that the pair (x, y) solves IC^a. We can replace (x, y) with $(x+\varepsilon_x, y+\varepsilon_{y,ic})$ so that the IC^a constraint still holds as an equality. Similarly, we can replace (x, y) with $(x+\varepsilon_x, y+\varepsilon_{y,obj})$ so that the sender's payoff remains constant. We define $\varepsilon'_{y,ic}(x)$ and $\varepsilon'_{y,obj}(x)$ as follows:

$$\varepsilon'_{y,ic}(x) = \lim_{\varepsilon_x \to 0} \frac{y+\varepsilon_{y,ic} - y}{x+\varepsilon_x - x}, \quad \varepsilon'_{y,obj}(x) = \lim_{\varepsilon_x \to 0} \frac{y+\varepsilon_{y,obj} - y}{x+\varepsilon_x - x}.$$

It is easy to show that both derivatives $\varepsilon'_{y,ic}(x), \varepsilon'_{y,obj}(x)$ are negative over the domain of x. It is easily verified that $\varepsilon'_{y,ic}(x)$ is negative, since a voter with the policy x must become more optimistic about her state based on policy y after we increase x by a small amount. Therefore, the binding IC^a defines y as a decreasing function of x. It is also easily verified that $\varepsilon'_{y,obj}(x)$ is negative as well: In order to keep the sender's payoff constant, we must decrease y as we increase x.

If $\varepsilon'_{y,obj}(x) < \varepsilon'_{y,ic}(x)$ over some region of x, then the decrease in y required for IC^a to hold is smaller than the decrease in y required for the sender's payoff to stay constant. In this case, the sender can improve by increasing x. Analogously, if

$\varepsilon'_{y,obj}(x) > \varepsilon'_{y,ic}(x)$ over some region of x, then the sender can improve by decreasing x. We want to argue that in the domain of x, one of the following three cases occurs: (i) $\varepsilon'_{y,obj}(x) > \varepsilon'_{y,ic}(x)$ for any $x \in [\underline{x}, \overline{x}]$; (ii) $\varepsilon'_{y,obj}(x) < \varepsilon'_{y,ic}(x)$ for any $x \in [\underline{x}, \overline{x}]$; (iii) there exists x' such that $\varepsilon'_{y,obj}(x) < \varepsilon'_{y,ic}(x)$ if $x > x'$ and $\varepsilon'_{y,obj}(x) > \varepsilon'_{y,ic}(x)$ if $x < x'$. In all three cases, the sender finds it optimal to set x either as high as possible or as low as possible. Therefore, the support of any optimal policy cannot have two interior values $x, y \in (0,1)$ with $x \neq y$.

We next show that one of three cases occurs by examining the sign of $\varepsilon'_{y,obj}(x) - \varepsilon'_{y,ic}(x)$. The term $\varepsilon'_{y,obj}(x) - \varepsilon'_{y,ic}(x)$ is positive if and only if:

$$\left(\frac{\lambda_1 + (1-\lambda_1)x}{1-\lambda_1+\lambda_1 x}\right)^{n-n_0-1} > \frac{\left(\frac{1-\lambda_1}{\lambda_1}\right)^{n_0} (\lambda_1 - p_h)\left(\ell\lambda_1 x^2(-n+n_0+2) - (\lambda_1-1)x(n-n_0-1)+\lambda_1\right)}{(p_h - 1 + \lambda_1)\left(\ell(1-\lambda_1)x^2(n-n_0-2) + \lambda_1 x(-n+n_0+1) + \lambda_1 - 1\right)}. \tag{13}$$

Let us first show that the second term in the denominator is negative:

$$\ell(1-\lambda_1)x^2(n-n_0-2) + \lambda_1 x(-n+n_0+1) + \lambda_1 - 1 < 0.$$

This inequality holds when $n = n_0 + 2$, which is the lowest value that n can take. The derivative of the left-hand side with respect to n is $-x(\lambda_1 - \ell(1-\lambda_1)x)$. We want to argue that $\lambda_1 - \ell(1-\lambda_1)x$ is weakly positive. If $\lambda_1 \geq \ell/(1+\ell)$, $\lambda_1 - \ell(1-\lambda_1)x$ is positive for any x. If $\lambda_1 < \ell/(1+\ell)$, no voter will ever rubber-stamp. The highest value of x is obtained when we set y to be zero. It is easily verified that $x < \lambda_1/(\ell(1-\lambda_1))$ in this case. Therefore, the derivative of the left-hand side with respect to n is weakly negative, so the inequality holds for any $n \geq n_0 + 2$.

The left-hand side of (13) decreases in x whereas the right-hand side increases in x. Therefore, either the above inequality holds for any x in the domain, or it does not hold for any x in the domain, or it holds only when x is below a threshold x'. This completes the proof. □

Based on lemma B.3, we are left with four possible cases:

(i) $n_0 = 0$, $y = x \in (0,1)$. In this case, all voters have the same policy. All IC^a constraints bind. This is the only symmetric policy.

(ii) $n_0 = 0$, $y \in (0,1)$, $x = 1$. R_1's information is more precise than her IC^a constraint requires. The other voters are willing to rubber-stamp given that R_1's policy is partially informative. This case is possible if and only if $\lambda_1 > \lambda_1^*$.

(iii) $n_0 > 0$, $y = x \in (0,1]$. The sender provides fully revealing policies to R_1 through R_{n_0}.

(iv) $n_0 > 0$, $y \in (0,1)$, $x = 1$. The sender provides fully revealing policies to R_1 through R_{n_0} and partial information to R_{n_0+1}.

The theme that the sender provides more precise information to some voters in order to persuade the others more effectively is reflected in the latter three cases. Moreover, when $\lambda_1 \leq \ell/(1+\ell)$, no voter ever rubber-stamps. The optimal policy must take the form of either case (i) or case (iii). If it is optimal to set n_0 to be zero, the optimal policy is the symmetric one. For $\lambda_1 > \ell/(1+\ell)$, the policy is either $(n_0, y, 1)$ with $n_0 \geq 0$ or (n_0, y, x) with $n_0 \geq 0$ and $y = x$. In the rest of this section, we analyze the parameter regions $(\lambda_1^*, 1)$, $(\ell/(1+\ell), \lambda_1^*]$, and $(1/2, \ell/(1+\ell)]$ separately.

When $\lambda_1 \in (\lambda_1^*, 1)$, only case (i) and case (ii) are possible. The threshold λ_1^* is given by:

$$\lambda_1^* = \frac{1}{2}\left(1 + \sqrt{\frac{1+\ell-4p_h}{1+\ell}}\right).$$

We are interested in characterizing how the optimal policy varies as we increase the number of voters n. If the sender chooses a policy as in case (ii), the sender's payoff is

$$\frac{(\ell+1)(\lambda_1 - p_h)(\lambda_1 + p_h - 1)}{\ell((\lambda_1 - 1)\lambda_1 - p_h + 1) + (\lambda_1 - 1)\lambda_1}. \tag{14}$$

The sender chooses y so that the other voters are willing to rubber-stamp. The payoff of the sender is given by offering the policy $\pi_1(L) = y$ to R_1 since the other voters approve for sure. If $\lambda_1 = \lambda_1^*$, the sender's payoff is equal to p_h, since at this correlation level the other voters are willing to rubber-stamp only if the sender fully reveals θ_1 to R_1. As λ_1 increases, the sender's payoff increases as well. When states are perfectly correlated, i.e., $\lambda_1 = 1$, the sender's payoff is $(1 + 1/\ell)p_h$, which is the same as if he were facing R_1 alone. In case (ii) the number of voters has no impact on the sender's payoff.

If the sender chooses a policy as in case (i), the IC^a constraint can be written as

$$\frac{(\lambda_1 + p_h - 1)(\ell(\lambda_1 - 1)x + \lambda_1)}{(\lambda_1 - p_h)(\ell\lambda_1 x + \lambda_1 - 1)} = \left(\frac{\lambda_1 x + (1-\lambda_1)}{\lambda_1 + (1-\lambda_1)x}\right)^{n-1}. \tag{15}$$

This equation implicitly defines x. The sender's payoff is given by

$$\frac{(\ell+1)x(\lambda_1 + p_h - 1)}{\ell\lambda_1 x + \lambda_1 - 1}(\lambda_1 + x(1-\lambda_1))^{n-1}. \tag{16}$$

We will show that x, the probability that a low-state voter approves, increases in n. Moreover, the sender's payoff decreases in n. The limit of the sender's payoff as

n approaches infinity is given by

$$\frac{(\ell+1)(\lambda_1+p_h-1)\left(\frac{(\ell(\lambda_1-1)+\lambda_1)(\lambda_1+p_h-1)}{(\ell\lambda_1+\lambda_1-1)(\lambda_1-p_h)}\right)^{\frac{1-\lambda_1}{2\lambda_1-1}}}{\ell\lambda_1+\lambda_1-1}. \tag{17}$$

This limiting payoff as $n \to \infty$ is lower than the payoff in case (ii). Moreover, when n equals 2, (16) is strictly higher than (14), so the sender obtains a strictly higher payoff in case (i) than in case (ii). Therefore, for each ℓ, p_h, and $\lambda_1 \in (\lambda_1^*, 1)$, there exists $n' \geq 3$ such that the sender is strictly better off in case (ii) than in case (i) if and only if $n \geq n'$.

Proposition B.2. *Suppose $\lambda_1 \in [\lambda_1^*, 1)$. For each ℓ, p_h and λ_1, there exists $n' \geq 3$ such that for $n \geq n'$, the sender is strictly better off in case (ii) than in case (i), so case (ii) policy is uniquely optimal.*

Proof. We first show that (14) is strictly higher than (17). Then we show that (16) strictly decreases in n. This completes the proof.

The ratio of (14) over (17) is given by

$$\frac{(\ell\lambda_1+\lambda_1-1)(\lambda_1-p_h)\left(\frac{(\ell(\lambda_1-1)+\lambda_1)(\lambda_1+p_h-1)}{(\ell\lambda_1+\lambda_1-1)(\lambda_1-p_h)}\right)^{\frac{\lambda_1-1}{2\lambda_1-1}}}{\ell((\lambda_1-1)\lambda_1-p_h+1)+(\lambda_1-1)\lambda_1}. \tag{18}$$

This ratio equals one when $\lambda_1 = 1$. We want to show that this ratio is strictly above one for $\lambda_1 \in [\lambda_1^*, 1)$. The derivative of this ratio with respect to p_h is negative if and only if

$$\left(\frac{(\ell(\lambda_1-1)+\lambda_1)(\lambda_1+p_h-1)}{(\ell\lambda_1+\lambda_1-1)(\lambda_1-p_h)}\right)^{\frac{\lambda_1+1}{2\lambda_1-1}} > 0.$$

Since this inequality holds, the ratio (18) decreases in p_h. Assumption 1 ensures that $p_h < \ell/(1+\ell)$. Substituting $p_h = \ell/(1+\ell)$ into (18), the ratio equals one. Therefore, for any $p_h < \ell/(1+\ell)$, the ratio (18) is above one. This completes the proof that (14) is strictly higher than (17) for $\lambda_1 \in [\lambda_1^*, 1)$.

Now we restrict attention to the policy of case (i). We first show that the solution x to (15) increases in n. The right-hand side of (15) increases in x and decreases in n. The left-hand side decreases in x. Therefore, as n increases, x must increase as well.

We next show that (16) decreases in n. It is easily verified that $\frac{(\ell+1)x(\lambda_1+p_h-1)}{\ell\lambda_1 x+\lambda_1-1}$ decreases in x. So it also decreases in n, since x increases in n. Therefore, if we can show that $(\lambda_1 + x(n)(1-\lambda_1))^{n-1}$ decreases in n, then (16) must decrease in n. For the rest of this proof, we use $x(n)$ instead of x to highlight the dependence of x on n. From the analysis of the previous paragraph, we know that the left-hand side of (15) decreases in n. Therefore, the total derivative of the right-hand side of

(15) with respect to n is negative. This puts an upper bound on $x'(n)$:

$$x'(n) < \frac{(\lambda_1 x(n) - \lambda_1 + 1)(\lambda_1 x(n) - \lambda_1 - x(n)) \log\left(\frac{\lambda_1(x(n)-1)+1}{\lambda_1(-x(n))+\lambda_1+x(n)}\right)}{(2\lambda_1 - 1)(n-1)}.$$

For $(\lambda_1 + x(n)(1-\lambda_1))^{n-1}$ to be increasing in n, the derivative $x'(n)$ must be at least:

$$x'(n) > \frac{(\lambda_1 + x(n) - \lambda_1 x(n)) \log(\lambda_1(-x(n)) + \lambda_1 + x(n))}{(\lambda_1 - 1)(n-1)}.$$

We want to show that this is impossible. The lower bound on $x'(n)$ is higher than the upper bound if and only if

$$\frac{(\lambda_1 x(n) - \lambda_1 + 1) \log\left(\frac{\lambda_1(x(n)-1)+1}{\lambda_1(-x(n))+\lambda_1+x(n)}\right)}{2\lambda_1 - 1} - \frac{\log(\lambda_1(-x(n)) + \lambda_1 + x(n))}{1 - \lambda_1} > 0.$$

The left-hand side of the above inequality decreases in $x(n)$. Moreover, the left-hand side equals zero when $x(n)$ equals one. Therefore, the above inequality always holds. This shows that the lower bound on $x'(n)$ is indeed higher than the upper bound. Therefore, $(\lambda_1 + x(n)(1-\lambda_1))^{n-1}$ and (16) decreases in n. □

We next focus on the parameter region $\lambda_1 \in \left(\frac{\ell}{(1+\ell)}, \lambda_1^*\right]$. The optimal policy might take the form of case (i), (iii), or (iv). We show that case (iv) dominates case (i) and case (iii) for n large enough. This implies that some voters learn their states fully for n large enough.

Proposition B.3. *Suppose $\lambda_1 \in (\ell/(1+\ell), \lambda_1^*]$. For each ℓ, p_h and λ_1, there exists $n' \geq 3$ such that for $n \geq n'$, the sender is strictly better off in case (iv) than in case (i) and (iii), so case (iv) policy is uniquely optimal.*

Proof. For any $\lambda_1 \in (\ell/(1+\ell), \lambda_1^*]$, there exists $k \geq 1$ such that $\lambda_1 \in (\lambda_{k+1}^*, \lambda_k^*]$. We first argue that the policy in case (iv) with n_0 being k leads to a higher payoff than the symmetric policy in case (i) when n is large enough.

If the sender uses the policy in case (iv) with n_0 being k, the IC^a constraint can be written as:

$$\ell = \frac{\lambda_1^{k+1}(\lambda_1 + p_h - 1)(\lambda_1 + (1-\lambda_1)y) + (1-\lambda_1)^{k+1}(\lambda_1 - p_h)(\lambda_1 y + (1-\lambda_1))}{(1-\lambda_1)\lambda_1^k(\lambda_1 + p_h - 1)(\lambda_1 + (1-\lambda_1)y) + \lambda_1(1-\lambda_1)^k(\lambda_1 - p_h)(\lambda_1 y + (1-\lambda_1))}.$$

This allows us to solve for y. Substituting this value of y into the sender's payoff, we obtain the sender's payoff as

$$\frac{(\ell+1)(2\lambda_1 - 1)((1-\lambda_1)\lambda_1)^k(\lambda_1 - p_h)(\lambda_1 + p_h - 1)}{(\lambda_1 - 1)(\ell(\lambda_1 - 1) + \lambda_1)\lambda_1^k(\lambda_1 + p_h - 1) + \lambda_1(\ell\lambda_1 + \lambda_1 - 1)(1-\lambda_1)^k(\lambda_1 - p_h)}. \tag{19}$$

The IC^a constraint holds as an equality for some y in $[0,1]$. The right-hand side of the IC^a constraint decreases in y. After substituting $y = 1$ into the IC^a constraint, we obtain the following inequality:

$$\ell \geq \frac{\lambda_1^{k+1}(\lambda_1 + p_h - 1) + (1-\lambda_1)^{k+1}(\lambda_1 - p_h)}{\lambda_1(1-\lambda_1)^k(\lambda_1 - p_h) - (\lambda_1 - 1)\lambda_1^k(\lambda_1 + p_h - 1)}.$$

This inequality imposes an upper bound on p_h:

$$p_h \leq \frac{(1-\lambda_1)(\ell(\lambda_1 - 1) + \lambda_1)\lambda_1^k + \lambda_1(\ell\lambda_1 + \lambda_1 - 1)(1-\lambda_1)^k}{(\ell(\lambda_1 - 1) + \lambda_1)\lambda_1^k + (\ell\lambda_1 + \lambda_1 - 1)(1-\lambda_1)^k}. \tag{20}$$

The payoff from the symmetric policy in case (i) approaches (17) as n goes to infinity. The ratio of (19) over (17) is given by:

$$\frac{(2\lambda_1 - 1)(\ell\lambda_1 + \lambda_1 - 1)((1-\lambda_1)\lambda_1)^k(\lambda_1 - p_h)\left(\frac{(\ell(\lambda_1-1)+\lambda_1)(\lambda_1+p_h-1)}{(\ell\lambda_1+\lambda_1-1)(\lambda_1-p_h)}\right)^{\frac{\lambda_1-1}{2\lambda_1-1}}}{(\lambda_1 - 1)(\ell(\lambda_1 - 1) + \lambda_1)\lambda_1^k(\lambda_1 + p_h - 1) + \lambda_1(\ell\lambda_1 + \lambda_1 - 1)(1-\lambda_1)^k(\lambda_1 - p_h)}.$$

This ratio decreases in p_h given the inequality (20). Moreover, if we substitute the upper bound on p_h as in (20) into the ratio above, this ratio equals:

$$\lambda_1^k\left((1-\lambda_1)^k\lambda_1^{-k}\right)^{\frac{\lambda_1}{1-2\lambda_1}+1}. \tag{21}$$

The term above strictly decreases in λ_1 if $\lambda_1 \in (1/2, 1)$. Moreover, the limit of the term above when λ_1 goes to one is equal to one. This shows that the ratio of (19) over (17) is strictly greater than one. Therefore, the symmetric policy in case (i) is dominated by the asymmetric policy in case (iv) when n is large enough.

We next argue that the policy in case (iii) with $n_0 \leq k$ leads to a lower payoff than the policy in case (iv) when n is large enough. If the sender uses the policy in case (iii), the IC^a constraint can be written as:

$$\frac{\lambda_1^{n_0}}{(1-\lambda_1)^{n_0}}\frac{(\lambda_1 + p_h - 1)(\ell(\lambda_1 - 1)x + \lambda_1)}{(\lambda_1 - p_h)(\ell\lambda_1 x + \lambda_1 - 1)} = \left(\frac{\lambda_1 x + (1-\lambda_1)}{\lambda_1 + (1-\lambda_1)x}\right)^{n-n_0-1}.$$

This equation implicitly defines x. The limit of x as n goes to infinity is 1. The limit of the sender's payoff as n approaches infinity is given by

$$\frac{(\ell+1)(\lambda_1 + p_h - 1)\lambda_1^{n_0}\left(\frac{\lambda_1^{n_0}}{(1-\lambda_1)^{n_0}}\frac{(\ell(\lambda_1-1)+\lambda_1)(\lambda_1+p_h-1)}{(\ell\lambda_1+\lambda_1-1)(\lambda_1-p_h)}\right)^{\frac{1-\lambda_1}{2\lambda_1-1}}}{\ell\lambda_1 + \lambda_1 - 1}. \tag{22}$$

The ratio of the limit payoff in case (iii) over the limit payoff in case (i) is given by the ratio of (22) over (17):

$$\left(\frac{1}{\lambda_1} - 1\right)^{\frac{(\lambda_1-1)n_0}{2\lambda_1-1}}\lambda_1^{n_0}.$$

This ratio is strictly smaller than (21) if $n_0 < k$. This ratio is equal to (21) If $n_0 = k$. Note that (21) is strictly higher than the ratio of the payoff in case (iv) over the limit payoff in case (i). This shows that the payoff in case (iv) is strictly higher than the limit payoff in case (iii). □

Let us now consider the case of a low λ_1: let $\lambda_1 \in \left(\frac{1}{2}, \frac{\ell}{(1+\ell)}\right]$. For this parameter region, even if a voter is certain that the realized grand state is G, she is still not willing to approve the project if $\lambda_1 - \ell(1-\lambda_1) < 0$. Due to this, no voter is willing to rubber-stamp the project, hence there is no voter for which $\pi_i(L) = 1$. A symmetric policy assigns the same $\pi_i(L) \in (0,1)$ for any i. An asymmetric policy assigns a fully revealing policy to a subgroup of the voters and a symmetric interior policy to the remaining voters, due to lemma B.3. Hence, any asymmetric policy is indexed by n_0, the number of voters who receive fully revealing recommendations. Let $P(n_0, n)$ denote the payoff from a policy with exactly n_0 voters with fully revealing recommendations among n voters in total.

For a policy with $n_0 \geq 0$ and $x \in (0,1)$ denoting the recommendation probability $\pi_i(L)$ for the partially informed voters, the binding IC^a is:

$$\left(\frac{\lambda_1 + (1-\lambda_1)x}{1-\lambda_1+\lambda_1 x}\right)^{n-n_0-1} = \frac{\lambda_1^{n_0}(p_h - 1 + \lambda_1)(\lambda_1 - \ell(1-\lambda_1)x)}{(1-\lambda_1)^{n_0}(\lambda_1 - p_h)(\lambda_1 - 1 + \ell\lambda_1 x)} \tag{23}$$

The corresponding payoff to the sender is:

$$P(n_0, n) = \frac{(\ell+1)\lambda_1^{n_0}(p_h + \lambda_1 - 1)x(\lambda_1 + (1-\lambda_1)x)^{n-n_0-1}}{\lambda_1 - 1 + \ell\lambda_1 x}.$$

The approval probability x implicitly depends on n. For any fixed $n_0 \geq 0$, $x(n)$ is increasing in n. To see this, notice that the right-hand side of (23) increases in x and decreases in n, while the left-hand side decreases in x. Therefore, an increase in n makes the right-hand side smaller, while it does not directly affect the left-hand side: for equality to hold, $x(n)$ has to increase as well.

Moreover, it is straightforward to see from the IC^a constraint that $x(n)$ increases in n_0 for a fixed group size n. Naturally, if more voters are offered fully revealing recommendations, this allows the sender to recommend the partially informed voters to approve more frequently.

The following result establishes that as $n \to \infty$, the payoff of the sender decreases in n_0; hence, for an infinitely large group, it is optimal for the sender to assign the symmetric policy with $n_0 = 0$.

Proposition B.4. *Suppose that* $\lambda_1 \in \left(\frac{1}{2}, \frac{\ell}{(1+\ell)}\right]$. *For any* $n_0 \geq 0$,

$$\lim_{n\to\infty} \frac{P(n_0, n)}{P(n_0+1, n)} = \frac{\ell+1}{\ell} > 1,$$

that is, for sufficiently large n, *the sender's payoff is decreasing in* n_0.

Proof. Let $x_0(n)$ and $x_1(n)$ denote the probability of recommendation to a low-state voter corresponding to the policies with n_0 and $n_0 + 1$ fully revealing recommendations. The ratio of payoffs is:

$$\frac{P(n_0,n)}{P(n_0+1,n)}$$
$$=\frac{\lambda_1+(1-\lambda_1)x_0(n)}{\lambda_1}\left(\frac{\lambda_1+(1-\lambda_1)x_0(n)}{\lambda_1+(1-\lambda_1)x_1(n)}\right)^{n-n_0-1}\frac{x_1(n)}{x_0(n)}\frac{\ell x_0(n)\lambda_1+\lambda_1-1}{\ell x_1(n)\lambda_1+\lambda_1-1}.$$

It follows from a comparison of the two IC^a constraints that $x_1(n)>x_0(n)$ for any n. Moreover, $(x_1(n)-x_0(n))\to 0$ as $n\to\infty$. Both $x_1(n)$ and $x_0(n)$ tend to $\frac{\lambda_1}{(1-\lambda_1)\ell}$. Therefore,

$$\lim_{n\to\infty}\frac{x_1(n)}{x_0(n)}=\lim_{n\to\infty}\frac{\ell x_0(n)\lambda_1+\lambda_1-1}{\ell x_1(n)\lambda_1+\lambda_1-1}=1;\quad \lim_{n\to\infty}\frac{\lambda_1+(1-\lambda_1)x_0(n)}{\lambda_1}=\frac{\ell+1}{\ell}.$$

Hence, the limit ratio of the payoffs reduces to:

$$\lim_{n\to\infty}\frac{P(n_0,n)}{P(n_0+1,n)}=\frac{\ell+1}{\ell}\lim_{n\to\infty}\left(\frac{\lambda_1+(1-\lambda_1)x_0(n)}{\lambda_1+(1-\lambda_1)x_1(n)}\right)^{n-n_0-1}.$$

In order to evaluate the remaining limit term, we need to approximate the rate at which $x_1(n)$ and $x_0(n)$ converge to $\frac{\lambda_1}{\ell(1-\lambda_1)}$ as $n\to\infty$. From the IC^a constraint when n_0 voters receive fully revealing recommendations, we have:

$$\frac{\lambda_1}{\ell(1-\lambda_1)}-x_0(n)=\frac{(2\lambda_1-1)(1-\lambda_1)^{n_0-2}(\lambda_1-p_h)w_0(n)^{n-n_0-1}}{\ell\left(\lambda_1^{n_0}(\lambda_1+p_h-1)+\lambda_1(1-\lambda_1)^{n_0-1}(\lambda_1-p_h)w_0(n)^{n-n_0-1}\right)},$$

where $w_0(n)=(1-\lambda_1+\lambda_1x_0(n))/(\lambda_1+(1-\lambda_1)x_0(n))$. Since $x_0(n)$ converges to $\frac{\lambda_1}{\ell(1-\lambda_1)}$ as $n\to\infty$, there exist constants $\underline{w}_0<\overline{w}_0\in(0,1)$ such that $w_0(n)\in(\underline{w}_0,\overline{w}_0)$ when n is sufficiently large. Similarly, from the IC^a constraint when n_0+1 voters receive fully revealing recommendations, we have:

$$\frac{\lambda_1}{\ell(1-\lambda_1)}-x_1(n)=\frac{(2\lambda_1-1)(1-\lambda_1)^{n_0-1}(\lambda_1-p_h)w_1(n)^{n-n_0-2}}{\ell\left(\lambda_1^{n_0+1}(\lambda_1+p_h-1)+\lambda_1(1-\lambda_1)^{n_0}(\lambda_1-p_h)w_1(n)^{n-n_0-2}\right)},$$

where $w_1(n)=(1-\lambda_1+\lambda_1x_1(n))/(\lambda_1+(1-\lambda_1)x_1(n))$. Since $x_1(n)$ converges to $\frac{\lambda_1}{\ell(1-\lambda_1)}$ as $n\to\infty$, there exist constants $\underline{w}_1<\overline{w}_1\in(0,1)$ such that $w_1(n)\in(\underline{w}_1,\overline{w}_1)$ when n is sufficiently large.

Notice that for any $a_0,a_1\in(0,1)$ and any $\kappa_0,\kappa_1>0$,

$$\lim_{n\to\infty}\left(\frac{1-\kappa_0a_0^n}{1-\kappa_1a_1^n}\right)^{n-1}=1.$$

Therefore,

$$\lim_{n\to\infty}\left(\frac{\lambda_1+(1-\lambda_1)x_0(n)}{\lambda_1+(1-\lambda_1)x_1(n)}\right)^{n-n_0-1}=\lim_{n\to\infty}\left(\frac{1-\frac{\ell(1-\lambda_1)}{(1+\ell)\lambda_1}\left(\frac{\lambda_1}{\ell(1-\lambda_1)}-x_0(n)\right)}{1-\frac{\ell(1-\lambda_1)}{(1+\ell)\lambda_1}\left(\frac{\lambda_1}{\ell(1-\lambda_1)}-x_1(n)\right)}\right)^{n-n_0-1}$$
$$=1.$$

This concludes the proof. □

An immediate implication of proposition B.4 is the following corollary, which pins down the exact limiting ratio of the payoffs from the symmetric policy and any asymmetric policy with $n_0 > 0$. This ratio depends only on ℓ and n_0: the larger the threshold ℓ, the smaller the comparative benefit from a symmetric policy.

Corollary B.4. *For any asymmetric policy with $n_0 > 0$ fully informed voters,*

$$\lim_{n\to\infty} \frac{P(0,n)}{P(n_0,n)} = \left(\frac{\ell+1}{\ell}\right)^{n_0}.$$

To sum up, we analyzed optimal individual persuasion for the case of homogeneous thresholds when the group size n is sufficiently large. We have established that when λ_1 is above $\frac{\ell}{\ell+1}$, the sender finds it optimal to rely on an asymmetric policy, that assigns different recommendation probabilities $\pi_i(L)$ across voters. For instance, if $\lambda_1 > \lambda_1^*$, the optimal policy consists of a partially informed voter and all other voters rubber-stamping; if $\lambda_1 \in (\lambda_2^*, \lambda_1^*)$, the optimal policy consists of one fully informed voter, another partially informed voters, and all other voters as rubber-stampers, and so on. If on the other hand λ_1 is below $\frac{\ell}{\ell+1}$, it is optimal for the sender to offer the same probability of recommendation in state L to all voters when the size group is sufficiently high. No voter is fully revealed her state for such low λ_1.

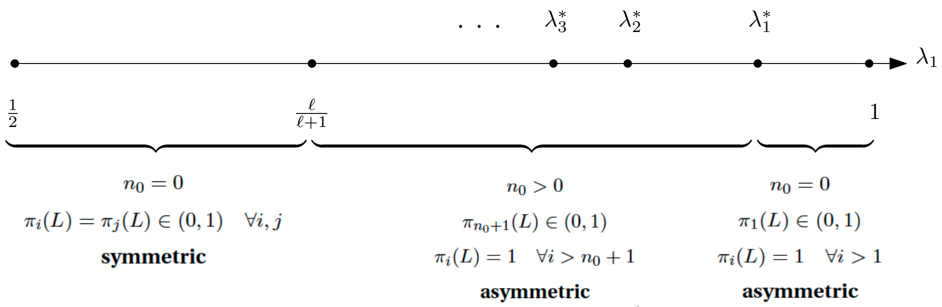

Figure 2: Optimal individual persuasion with homogeneous thresholds

B.6 Strict Ahlswede-Daykin theorem

Theorem B.5. *Suppose $(\Gamma, \succ)$ is a finite distributive lattice and functions f_1, f_2, f_3, $f_4 : \Gamma \to \mathbb{R}^+$ satisfy the relation that*

$$f_1(a)f_2(b) < f_3(a \wedge b)f_4(a \vee b),$$

$\forall a, b \in \Gamma$. Furthermore, suppose that $f_3(c) > 0$ and $f_4(c) > 0$ for any $c \in \Gamma$. Then

$$f_1(A)f_2(B) < f_3(A \wedge B)f_4(A \vee B),$$

$\forall A, B \subset \Gamma$, where $f_k(A) = \sum_{a\in A} f_k(a)$ for all $A \subset \Gamma, k \in \{1,2,3,4\}$, and $A \vee B = \{a \vee b : a \in A, b \in B\}$, $A \wedge B = \{a \wedge b : a \in A, b \in B\}$.

Proof. [41] Because Γ is a finite distributive lattice, it suffices[42] to prove that the result holds for $\Gamma = 2^N$, the lattice of subsets of the set $N = \{1, ..., n\}$ partially ordered by the inclusion relation. Then, $a \in \Gamma$ is a particular subset of N, and $A \subset \Gamma$ is a subset of subsets of N.

First, let us establish the result for $n = 1$, so $N = \{1\}$. Then $\Gamma = \{\emptyset, \{1\}\}$. Let f_k^0, f_k^1 denote the function f_k for $k = 1, 2, 3, 4$ evaluated at $\emptyset$ and $\{1\}$ respectively. By the premise, given that $f_3^0, f_3^1, f_4^0, f_4^1 \neq 0$,

$$f_1^0 f_2^0 < f_3^0 f_4^0, \quad f_1^1 f_2^0 < f_3^0 f_4^1, \quad f_1^0 f_2^1 < f_3^0 f_4^1, \quad f_1^1 f_2^1 < f_3^1 f_4^1.$$

It is straightforward to check that the result holds for any A and B that are singletons. This leaves only the case of $A = B = \{\emptyset, \{1\}\}$. We need to show that

$$(f_1^0 + f_1^1)(f_2^0 + f_2^1) < (f_3^0 + f_3^1)(f_4^0 + f_4^1).$$

If either f_1 or f_2 is zero, then the result follows trivially. So let us now consider the case in which they all nonzero. It is sufficient to consider the case for which $f_k^0 = 1$ for all k. It follows that

$$f_1^1 < f_4^1, \quad f_2^1 < f_4^1, \quad f_1^1 f_2^1 < f_3^1 f_4^1.$$

We would like to show that

$$(1 + f_1^1)(1 + f_2^1) < (1 + f_3^1)(1 + f_4^1). \tag{24}$$

If $f_4^1 = 0$, the result follows immediately. So let us consider the case for which $f_4^1 > 0$. The result we want to show becomes easier to satisfy as f_4^1: hence it is sufficient to establish it for a very low f_4^1. From the inequality $f_1^1 f_2^1 < f_3^1 f_4^1$, we know that

$$f_4^1 > \frac{f_1^1 f_2^1}{f_3^1}.$$

For some small fixed $\epsilon > 0$, let $f_4^1 = \frac{f_1^1 f_2^1}{f_3^1} + \epsilon$. From the fact that $f_1^1 < f_4^1$ and $f_2^1 < f_4^1$, it follows that

$$(f_4^1 - f_1^1)(f_4^1 - f_2^1) > 0. \tag{25}$$

We want to prove that

$$(1 + f_1^1)(1 + f_2^1) < (1 + f_3^1)\left(1 + \frac{f_1^1 f_2^1}{f_3^1} + \epsilon\right) \tag{26}$$

[41] This proof adapts the proof presented in Graham (1983). See: Graham, R. L. 1983. "Applications of the FKG Inequality and Its Relatives," in *Conference Proceedings, 12th Intern. Symp. Math. Programming.* Springer: 115-131.

[42] Every distributive lattice can be embedded in a powerset algebra so that all existing finite joins and meets are preserved.

which is equivalent to

$$f_4^{1^2} + f_1^1 f_2^1 - f_4^1 f_1^1 - f_4^1 f_2^1 > -f_4^1\epsilon - f_4^{1^2}\epsilon.$$

But from 25,

$$f_4^{1^2} + f_1^1 f_2^1 - f_4^1 f_1^1 - f_4^1 f_2^1 > 0 > -f_4^1\epsilon - f_4^{1^2}\epsilon$$

for any $\epsilon > 0$. This establishes 26 for any arbitrarily small ϵ, hence, inequality 24 is satisfied. So the proof for $n = 1$ is concluded.

Let us now assume that the result holds for $n = m$ for some $m \geq 1$, and we would like to show that it holds for $n = m + 1$ as well. Suppose f_k, $k = 1, 2, 3, 4$ satisfy the premise of the result for $n = m + 1$ for $\Gamma = 2^{\{1,\ldots,m+1\}}$.. Let A, B be two fixed subsets of the power set $2^{\{1,\ldots,m+1\}}$. Let us define $f'_k : 2^{\{1,\ldots,m\}} \to \mathbb{R}^+$ such that

$$f'_1(a') = \sum_{a\in A, a'=a\setminus\{m+1\}} f_1(a), \quad f'_2(b') = \sum_{b\in B, b'=b\setminus\{m+1\}} f_2(b),$$

$$f'_3(w') = \sum_{w\in A\cap B, w'=w\setminus\{m+1\}} f_3(w), \quad f'_4(v') = \sum_{v\in A\cup B, v'=v\setminus\{m+1\}} f_4(v).$$

For any $a' \in 2^{\{1,\ldots,m\}}$,

$$f'_1(a') = \begin{cases} f_1(a) + f_1(a \cup \{m+1\}) & \text{if } a' \in A, a' \cup \{m+1\} \in A \\ f_1(a) & \text{if } a' \in A, a' \cup \{m+1\} \notin A \\ f_1(a \cup \{m+1\}) & \text{if } a' \notin A, a' \cup \{m+1\} \in A \\ 0 & \text{if } a' \notin A, a' \cup \{m+1\} \notin A \end{cases}$$

With such a definition,

$$f_1(A) = f'_1(2^{\{1,\ldots,m\}}).$$

Similarly, $f_2(B) = f'_2(2^{\{1,\ldots,m\}})$, $f_3(A \wedge B) = f'_3(2^{\{1,\ldots,m\}})$, $f_4(A \vee B) = f'_4(2^{\{1,\ldots,m\}})$. Now, a similar argument to the one followed for $n = 1$ with a' corresponding to $\emptyset$ before and $a' \cup \{n\}$ corresponding to $\{1\}$, gives us that

$$f'_1(a')f'_2(b') < f'_3(a' \wedge b')f'_4(a' \vee b'),$$

$\forall a', b' \in 2^{\{1,\ldots,m\}}$. But by the induction hypothesis for $n = m$,

$$f'_1(2^{\{1,\ldots,m\}})f'_2(2^{\{1,\ldots,m\}}) < f'_3(2^{\{1,\ldots,m\}})f'_4(2^{\{1,\ldots,m\}})$$

since $2^{\{1,\ldots,m\}} \wedge 2^{\{1,\ldots,m\}} = 2^{\{1,\ldots,m\}}$ and $2^{\{1,\ldots,m\}} \vee 2^{\{1,\ldots,m\}} = 2^{\{1,\ldots,m\}}$. This implies the desired result for $n = m + 1$:

$$f_1(A)f_2(B) < f_3(A \wedge B)f_4(A \vee B).$$

This concludes the proof. □