The Fragility of Specialized Advice*

Patrick Lahr

University of Bonn

Justus Winkelmann

University of Bonn

June 17, 2021

Abstract

We consider a multi-sender cheap talk model, where the receiver faces uncertainty over whether senders have aligned or state-independent preferences. This uncertainty generates a trade-off between giving sufficient weight to the most informed aligned senders and minimizing the influence of the unaligned. We show that preference uncertainty diminishes the benefits from specialization, i.e., senders receiving signals with more dispersed accuracy. When preference uncertainty becomes large, it negates them entirely, causing qualified majority voting to become the optimal form of communication. Our results demonstrate how political polarization endangers the ability of society to reap the benefits of specialization in knowledge.

Keywords: cheap talk, information aggregation, specialization in knowledge, voting

*Corresponding author e-mail: patricklahr@uni-bonn.de. This manuscript has previously been circulated under the title "Fake Experts". We thank Ian Ball, Felix Bierbrauer, Jan Knöpfle, Daniel Krähmer, Stephan Lauermann, Benny Moldovanu, Martin Pollrich, Larry Samuelson, Dezsö Szalay, and Aleh Tsyvinski. Both authors wish to express their gratitude for the hospitality of the Yale Economics Department during their respective visits, amid which much of this research has been conducted. Funding by the Deutsche Forschungsgemeinschaft (DFG) through CRC TR 224 Project B1 is gratefully acknowledged.

1 Introduction

Consider a receiver who must decide on an action but who relies on multiple senders to obtain relevant information.[1] Given any particular question, the more specialized the senders are, the more information is concentrated among a few of them, because fewer and fewer will know anything about the subject at hand. Greater specialization can benefit decision making, because having one source that conveys perfect information is superior to having many sources whose majority advice may err. On the other hand, relying on only a few sources naturally grants each of them more influence, which increases the risk that some senders with diverging interests pose as experts. Therefore, the main questions of this article are as follows. What are optimal communication patterns, taking into account that some senders have privately known diverging interests? How do specialization and preference uncertainty affect communication, and how do they interact? Lastly, how can we explain that voting mechanisms are still prevalent in some situations, despite seemingly ignoring specialization?

To illustrate these questions, let us consider an example of a government agency asking for expert advice on a new regulatory decision. In these cases, experts hold information that is valuable in making the decision, yet they might also have private interests owing to financial ties to the regulated industry. Due to different specializations and experiences, it is likely that the experts in the advisory bodies are not equally well-informed. In such a situation, should the regulator give more weight to those experts who express high confidence in their positions? This allows her to account for the heterogeneity of information, but it also increases the possibility of experts with conflicts of interest exerting greater influence by falsely claiming high confidence. Or should she simply decide based on the numbers of individual votes for and against approval?

More generally, how should individuals adjust their learning when they might be lied to by interested parties? This concern extends beyond commercial interest and has recently been at the center of political debates, most notably that of "fake news." We therefore try to understand the impact of a voter struggling to distinguish between interest-led rhetoric and reputable news. Is in-depth reporting still heard, or does public discourse reduce to deciding which side has gathered sufficient numbers?

In this article, we examine such decision problems using a multi-sender cheap talk model in which the senders receive conditionally independent information. Instead, we assume that they have either aligned interests or state-independent preferences. We believe that this assumption is reasonable in situations like the ones above, in which a clear common interest objective that might be trumped by the private considerations of the senders. As is standard in the literature, we focus throughout on the most informative

[1] Throughout the article, we refer to the receiver with the female pronoun and to the senders with the male pronoun.

equilibrium. Essential to our analysis is the interaction between the senders' information structures and the receiver's uncertainty over their preferences. To make things concrete, suppose the regulator from our first example has a prior of 1/2 that the approval of a new regulation is the right decision. She is advised by 5 experts. All these experts are informed conditionally iid, which can take either of the two following forms:

1. Each expert receives a binary signal that is known to match the true state of the world with a 62% chance.
2. Each expert has a 15% chance of learning the state perfectly and otherwise learns nothing.

In the absence of preference uncertainty, the regulator bases her decision under information structure 1 on a majority vote. This gives her a 72% chance of being correct. Under Information structure 2, she follows perfectly informed experts, of whom there is at least one with a 56% chance. Otherwise, she has to toss a coin, which, in total, gives her a 78% chance of being correct.[2] However, as we show below, because gains from very precise signals are particularly vulnerable to preference uncertainty, the advantage of Information Structure 2 shrinks as preference uncertainty is gradually introduced. At around a 10% chance that a expert is partisan in either direction, both information structures allow the regulator to make the right decision with the same probability, i.e., only 68% of the time. If the chance of partisanship rises further, Information Structure 1 becomes even superior to Information Structure 2.

To conceptualize the difference between such information structures, we introduce two concepts. Any information a expert receives has both a direction, i.e., it favors either the approval or rejection of the regulation, and an intensity, i.e., how much it moves the expert's belief away from the prior. We call the mean intensity of a expert's signal his *average informativeness*. In our example, average informativeness is higher under information structure 1 because each expert's posterior is moved 0.12 away from his prior, although in the second structure this distance on average is only 0.075.

In our example, a specialized expert learns whether the question is in his field of specialization and hence updates his posterior more or less than his generalist colleague. To illustrate the concept of specialization, consider the following Information Structure 3:

3. Each expert receives a signal that matches the true state of the world with a 60% chance, but also has a 5% chance of learning the state perfectly.[3]

[2]The two precision numbers can be derived from the following two expressions: $0.62^5 + \binom{5}{4}0.62^4 0.38 + \binom{5}{3}0.62^3 0.38^2$ and $1 - 0.85^5 + \frac{1}{2}0.85^5$ respectively.

[3]As $0.05 * 0.5 + 0.95 * 1 = 0.12$ average informativeness is the same as for information structure 1.

Information Structure 3 can be interpreted as a expert now recognizing, whether a given question is in his field of specialization in which case he is more secure in his judgment or not in which case he is less secure. For a generalist the same information would have no effect. Information Structure 3 is always weakly better than 1, but completely loses its edge when there is at least a 1 in 12 chance that a expert with either information structure is a partisan in either direction an effect we formalize in Lemma 2. Thus, preference uncertainty destroys possible gains from specialization.

As noted above, in these comparisons, we focus on the most informative equilibrium. Behavior in this equilibrium in the presence of preference uncertainty is characterized by senders with aligned interest stating their true beliefs, whereas partisans send messages independently of their information, mimicking senders who receive the most informative signals in their preferred direction. The receiver acts on the central trade-off to use as much information from aligned senders as possible although limiting the influence of partisans and putting caps on the influence any one message can have on her decision. However, between these caps, communication between the aligned senders and the receiver remains perfect. This is in contrast to the typical coarsening of messages in cheap talk games with known bias in which there is a coarsening affects all messages. Consequently, partisanship has two distinct effects on information transmission, which relate to the two concepts of informational content: average informativeness and specialization.

The first effect is that the information held by partisans is lost, because they send messages independent of the signal they receive. This leads to a proportional loss in average informativeness. Such a loss is due to the mere existence of partisans and is equally effective if the partisan senders preferences are known to the receiver; hence, partisan messages can be ignored. In that sense known partisanship is equivalent to having aligned but incompetent sender types.

The second effect is a loss in effective specialization caused by the indistinguishability of advisory and partisan senders. The receiver experiences uncertainty over the senders' preferences, which results in the uninformative messages of the partisans, being treated the same as the messages from senders with the most valuable signals. This leads to an effective loss in specialization, because the best signals are now diluted by signals that are, on average, uninformative, which makes the value of different messages more homogeneous overall.

If partisanship becomes sufficiently strong, all gains from specialization are wiped out. The senders reduce their messages to mere indications of the direction of their information, although the receiver bases her decision on whether the number of messages in favor of one alternative meets a fixed threshold. This binary communication between the senders and the receiver resembles a form of qualified majority voting. Consequently, even in situations in which some senders naturally have more to contribute than others, voting arises as an optimal way of communication. In these environments, the average informativeness of each sender becomes the decisive predictor of the receiver's ability to match the state with her decision, although specialization becomes worthless. More generally, in some configurations, a group of expert advisors with more heterogeneously distributed posteriors is preferable to a receiver, when partisanship is low, but performs worse when partisanship is high. We are thus worried that increasing political polarization might substantially diminish the gains society can reap from increases in the specialization of knowledge.

The article continues as follows. In the rest of this section, we review the literature. In Section 2, we introduce our model. Section 3 analyzes the special case in which all the senders have aligned interests. We use this natural benchmark to contrast our later findings. Section 4 introduces concepts to analyze both the senders' information structure and the information that is transmitted to the receiver. We apply these concepts to our general model in Section 5 and derive our main results. Section 6 concludes.

Related Literature

We place our work between the literatures on cheap talk and information aggregation in voting. The former builds on the seminal work of Crawford and Sobel [1982] and analyzes strategic communication between a better-informed sender and a receiver whose action determines the payoff of both. In their original setup, the sender has private and perfect information on a one-dimensional state of the world and a bias known to the receiver.

We depart from this classical model in three central ways, with multiple senders who are imperfectly informed and whose preferences are unknown to the receiver.

Gilligan and Krehbiel [1989] were the first to study a model with multiple senders. In their model, two privately and perfectly informed senders with publicly known biases communicate with a receiver. The focus of their analysis is the comparison of three communication protocols that comprise different forms of cheap talk. Similarly, Krishna and Morgan [2001] study a setting with two senders who sequentially send public messages to a receiver. The degree of information revelation depends on whether the senders have aligned or opposing biases.

Austen-Smith [1990] was the first to study a cheap talk problem in which the senders are imperfectly informed about a binary state of the world. Wolinsky [2002] identifies circumstances under which a cheap talk phase between senders alters the decision and solves for the most efficient communication structure.

Battaglini [2017] studies public protests as an informal means to aggregate dispersed information in democracies. As politicians generally do not commit to a reaction to protests ex-ante, such protests are, in fact, modeled by cheap talk. Information does not need to aggregate if the variation in optimal decision threshold between the politicians and the citizens is large relative to the information available to each citizen.

In Alonso et al. [2008] and Hummel et al. [2013], uncertainty about the senders' preferences arises endogenously, because each sender is interested in the decision matching his type, although the receiver wants to match her decision to the average of the senders' types. Hence, a sender's type contains both relevant information about the state, i.e., the average of types and information about the bias, i.e., the distance of the individual sender's type from the average.

In contrast to this approach and in line with our own, Morgan and Stocken [2003] and Li and Madarász [2008] analyze a single-sender game with private bias that does not enter the receiver's payoff. In Morgan and Stocken [2003] senders are either aligned or have a non-partisan bias. The authors show that any uncertainty on the receiver's side about such a bias yields to the bunching of messages and therefore information loss. If the potential conflict of interest is sufficiently large, a case they call semi-response equilibrium of size one, the communication of their single sender, is similar to the one of a every sender individually in our model, when partisanship can only occur in one direction. In extreme cases, only information that is opposed to the potential bias is truthfully reportable. Li and Madarász [2008] find that both players can benefit from the privacy of the sender's bias. The receiver's ignorance sender's preferences allows for conflict-hiding equilibria in which a more or oppositely biased type can transmit more information, because identical messages are also transmitted by the other type, changing the expected meaning of the message. In contrast in our model not knowing the senders' types is always detrimental to the receiver, because senders with state independent preferences message independently of the signals they received. Knowing a sender's bias therefore makes it possible to ignore such messages, instead of them being garbled with informative messages from aligned agents.

Another strand of literature has assumed that one type of sender is non-strategic and always communicates thruthfully. This strand includes Sobel [1985], Benabou and Laroque [1992], i Vidal [2006], and Glazer et al. [2019]. The last is an independent working paper modeling internet recommendation systems. Their work, like ours, studies a model with multiple imperfectly informed senders, in which some senders have state-independent preferences. because we find in our model that the most informative equilibrium is one in which aligned senders have a best response to be honest, the equilibria in both models are similar. However, given the different economic motivations of the two works, the analyses building on these equilibria are distinct. For a comparison of honest sender types, versus strategic yet aligned types see also Kim and Pogach [2014].

There is also a structural similarity of our model to models of voluntary disclosure, a literature started by Dye [1985] and in particular to the of work of Jung and Kwon [1988]. These similarities can be understood by focusing on the special case of our model with a single sender, that can only be a partisan towards state 0. As partisans will always send messages yielding to the lowest posterior, the resulting communication is similar to disclosure games where agents desire the receiver to have high believes about their type, yet some agents are unable to provide evidence and are hence treated equally to the the lowest types that opt not to disclose their information.

The second body of literature we relate to is on information aggregation in voting. It goes back to Condorcet [1785] and his famous jury theorem, stating that large groups of independently informed senders select the correct alternative with near certainty. He assumes that senders vote sincerely, although Feddersen and Pesendorfer [1997] establish a similar result for strategic senders. They show that when the number of voters grows large, privately held information leads to the same decision as public information.

Despite the effectiveness of voting for information aggregation in large populations, the same literature has revealed effects such as the swing voter's curse, first discussed by Feddersen and Pesendorfer [1996], which illustrates a loss of information in small populations. This loss is mainly due to the nature of the voting game, with its limited number of messages, usually two or three, and its fixed threshold. Any such voting rule can be interpreted in our model as a behavioral type of receiver to which strategic voters react optimally. We therefore believe that our model, with its strategic receiver, provides a natural benchmark for voting systems and helps distinguish which losses of information are necessary consequences of conflicts of interest and which are due to the specific features of real-world voting systems.

McMurray [2017] and Azrieli [2018] examine the limitations of elections with few available messages. McMurray [2017] studies a common interest election of ex-ante symmetric candidates by a fixed number of heterogeneously informed senders. In equilibrium, voters coordinate around specific candidates to transmit information. His model can be interpreted as a cheap talk game with a restricted number of messages. If the number of candidates becomes large, the model converges to our common interest setting. Azrieli [2018] analyzes the loss of anonymous voting rules if the senders are publicly known to be differently well-informed. The common-value analysis is also closely related to ours. However, we assume that signals are private information and focus on their interplay with private interests.

Lastly, Li et al. [2001] analyze two-player decision games with known conflicts of interest. Agents can choose from a set of integer weights, and an action is taken depending on whether the sum of the weights exceeds a predetermined threshold. These rules are reminiscent of the equilibrium play of the receiver in our communication games, despite the lack of commitment in our model.

2 The Model

There is a set of senders $\{1, ..., n\}$ and a receiver indexed by 0. Each sender i receives a signal about the unknown state of the world $\omega = \{0, 1\}$. Signals are identically distributed and independent, conditional on the true state of the world. There is a common prior $p_0 = \mathbb{P}[\omega = 1] \in (0, 1)$ that the state of the world is 1. As signals are conditionally independent, all information is contained in the resulting distribution over posteriors, and we shift the attention completely to the latter. Each sender draws his posterior from the probability mass function μ, which is consistent with p_0. We assume that the information structure is such that it leads to a finite number of possible posteriors $\mathcal{P} = \text{supp } \mu$. For some results, we assume that no signal is uninformative, i.e., $\mu(p_0) = 0$. We call a distribution μ that fulfills this assumption *never-ignorant*. The receiver shares the prior but observes no signal.[4]

In addition to different signals, players are heterogeneous with respect to their preferences, as described by a parameter $\lambda \in \{0, \lambda_0, 1\}$ with $\lambda_0 \in (0, 1)$. Each sender i independently draws a preference parameter λ_i that is independent of the posteriors and distributed according to probability mass function γ. The decision-maker has the commonly known preference parameter λ_0. Each sender also draws a posterior p_i independently from both other agents and his preference type according to probability mass function μ. We call the tuple (p_i, λ_i) the type of sender i, and denote with $\mu \times \gamma$ the distribution over types.

After observing the signal, each sender i simultaneously sends a cheap talk message $t_i \in [0, 1]$ to the receiver. We denote the potentially mixed strategy by $m_i : \mathcal{P} \times \{0, \lambda_0, 1\} \to \Delta[0, 1]$, where $\Delta[0, 1]$ denotes the set of all probability measures over $[0, 1]$. Whenever the strategy of the sender specifies a finite set of messages to be sent a.s. we denote the probability that sender i with type (p_i, λ_i) sends message t_i by $m_i(p_i, \lambda_i)(t_i)$. We call a strategy *truthful for preference type* λ_i if $m_i(p_i, \lambda_i)(p_i) = 1$ for all types (p_i, λ_i). The tuple of messages of all the senders is denoted by $t = (t_1, ..., t_n)$.

We denote the belief of the receiver accounting only for sender i's message t_i by $q(t_i) = E[p_i|t_i]$ and call it the virtual posterior of sender i.[5] The posterior of the receiver incorporating the messages t of all the senders is denoted by $q(t)$. After processing all messages, the receiver takes action $a \in \{0, 1\}$. Utilities for the senders and the receiver are given by

$$u(a, \omega, \lambda_i) = (1 - \lambda_i)\mathbb{1}\{a = 0\} + \lambda_i\mathbb{1}\{a = 1\} + \mathbb{1}\{a = \omega\},$$

where $\mathbb{1}$ is the indicator function, i.e., $\mathbb{1}\{A\}$ is 1 if event A is true and 0 otherwise.

A player i prefers action 1 if and only if his belief that the state of the world is 1 is larger than or equal to $1 - \lambda_i$. A higher preference parameter λ_i leads to a higher expected utility of player i given that the action is equal to 1. Senders with preference

[4] We discuss alternatives to some of the assumptions made in this model in Appendix C.

[5] Different strategies m_i induce different virtual posteriors $q_i(\cdot)$. Anticipating that the senders play symmetric strategies in an optimal equilibrium, we drop the subscript i of the virtual posterior $q_i(\cdot)$ to simplify the notation.

parameters 0 and 1 weakly prefer the action that matches their preference parameter irrespective of the posterior. We call senders with these preference parameters *partisans*. The remaining senders with $\lambda_i = \lambda_0$ have the same interests as the receiver. We call these senders *advisors*.

Before we proceed, we summarize the timing of the game. First, nature draws a state of the world ω. Second, every sender i randomly draws a type (p_i, λ_i) according to the conditional type distribution $\mu_\omega \times \gamma$. Third, each sender i sends a message t_i to the receiver. Last, the receiver takes an action a, and payoffs are realized. We assume that the receiver does not have commitment power, i.e., she can not credibly commit to a decision rule before getting the messages of the senders.[6] Consequently, we solve for Bayesian Nash equilibria.

In the following, we split the analysis into three parts. We start by studying the common interest case in Section 3. Here, all the senders have aligned preferences. The special case of our setting serves as a benchmark and allows us to become familiar with how the receiver processes the signals from the senders. In Section 4, we focus on the information structure of the senders, introduce the concept of specialization, and illustrate its significance in the common interest case. Lastly, in Section 5, we apply these concepts in our analysis of the general case, in which we allow for private interests.

3 Common Interest

In this section, we derive a benchmark equilibrium that maximizes the utility of the receiver when all the senders have aligned preferences, i.e., $\gamma(\lambda_0) = 1$ and $\gamma(0) = \gamma(1) = 0$. In the common interest case, such an equilibrium maximizes the utility of the senders as well. The general idea of this equilibrium is straightforward. The receiver needs to perform Bayesian updating given the senders' messages, and the senders, knowing that their information is aggregated in a statistically correct way, can state their posteriors, revealing all their information. It is as if the regulator could observe all the signals of the experts.

In the following description of this equilibrium, we focus on the statistical properties and interpretation of how the receiver updates the information and how she translates it into her decision.

Definition 1. *A receiver follows a weighted majority rule if her strategy* $a : [0,1]^n \to \{0,1\}$ *is of the form*

$$a(t) = \begin{cases} 1 & \text{if } \sum_{i=1}^{n} w(t_i) > \tau \\ 0 & \text{else} \end{cases}$$

[6] In particular, this excludes equilibria of the kind discussed in Gerardi et al. [2009].

for messages $t = (t_1, ..., t_n)$ *of the senders, a weighting function* $w : [0,1] \to \mathbb{R}$*, and a threshold* τ*.*

Under a weighted majority rule, the receiver transforms every message t_i into a weight $w(t_i)$ and takes decision 1 if the sum of weighted messages is larger than a threshold τ. One can interpret this as the receiver giving the senders free choice over the weights in the image of w and limiting herself to the application of a simple rule. When the size of the image is equal to 2, this comes down to proposing a decision by qualified majority voting. We come back to this analogy in Section 5. The next proposition translates the above-described equilibrium into this language.

Proposition 1. *The following describes a receiver-optimal Bayesian Nash equilibrium:*

- *Advisors message truthfully, i.e.,* $m_i(p_i, \lambda_0)(p_i) = 1$*.*
- *The receiver follows a weighted majority rule with weighting function*

$$w(x) = \ln \frac{x}{1-x} - \ln \frac{p_0}{1-p_0}$$

and threshold $\tau = -\left(\ln \frac{\lambda_0}{1-\lambda_0} + \ln \frac{p_0}{1-p_0}\right)$*.*

Proof. See Appendix A. □

In this optimal equilibrium, the senders play the truthful strategy to transmit their posterior to the receiver. The receiver has correct beliefs about this and can deduce from the posteriors the entirety of their information. She then translates it into the optimal decision via Bayesian updating, which we interpret as her applying a weighted majority rule, with log-likelyhood ratio weights.[7] Hence, an equilibrium with higher payoffs for the receiver cannot exist.[8]

Figure 1 illustrates the weighting function with prior $p_0 = \frac{3}{4}$ for the common interest case.

[7] See Nitzan and Paroush [1982] and Shapley and Grofman [1984] for two classical treatments of the role of such weighting roles for optimal information aggregation in groups.

[8] McLennan [1998] studies optimality of equilibria in common interest games more generally.

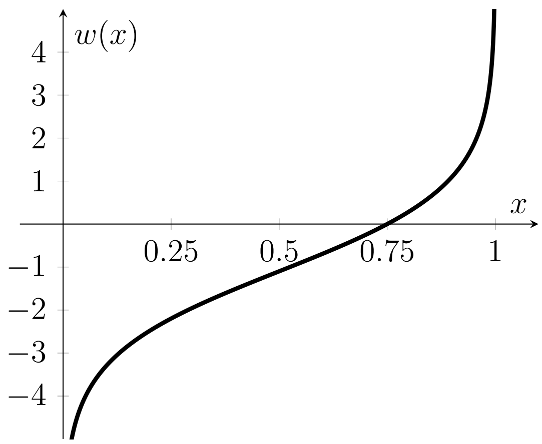


Figure 1: Weighting function $w(x) = \ln \frac{x}{1-x} - \ln \frac{p_0}{1-p_0}$ with prior $p_0 = \frac{3}{4}$ for the common interest case.

A posterior p_i of sender i that equals the prior p_0 gets weight 0 because it does not transmit any additional information. In contrast, a posterior $p_i \in \{0, 1\}$ means that sender i knows the state of the world perfectly. This sender's information is sufficient to make an optimal decision, and he should outweigh all other senders. Thus, as p_i goes to 1 (0), the corresponding weight tends to ∞ ($-\infty$). The unrestrictedness of the weighting function encodes the extraordinary value of perfect information.

In the next section, we refer to the receiver-optimal equilibrium when we assess different distributions of sender types. The expected utility of the receiver $u^\star(q(t))$ with the posterior $q(t)$ is given by

$$u^\star(q) = \begin{cases} \lambda_0 + q & \text{if } q > 1 - \lambda_0 \\ 2 - \lambda_0 - q & \text{else.} \end{cases}$$

We now turn to the analysis of the senders' information structure.

4 Specialization

The distribution over posteriors of the senders is a crucial object in our model. In this section, we develop the concepts that we use to describe them throughout. A classical concept in this regard is, Blackwell's informativeness order, introduced in, Blackwell [1950].[9]

Definition 2. *Let μ and ν be two distributions over posteriors with cdfs F and G,*

[9] The particular characterization used is related to the second stochastic order used in decision theory. A "riskier" distribution over posteriors contains more information.

respectively. We say that μ is more informative than ν, denoted by $\mu \succ \nu$, if

$$\int_0^y F(x)\mathrm{d}x \geq \int_0^y G(x)\mathrm{d}x \quad \textit{for all } y \in [0,1].$$

For convenience, we have summarized some facts about the informativeness order in Appendix B. An important interpretation of the above integral condition is that the more informative information structure is a mean-preserving spread of its less informative counterpart. Note that in our setting, both integrals are equal at 1 given equal priors. In this section, we build on this classical order to conceptualize specialization in knowledge:

Definition 3. *Let μ and ν be two distributions over posteriors with cdfs F and G, respectively. We say that μ is more specialized than ν, denoted by $\mu \succ_s \nu$, if $\mu \succ \nu$ and*

$$\int_0^{p_0} F(x)\mathrm{d}x = \int_0^{p_0} G(x)\mathrm{d}x. \tag{1}$$

We refer to the reverse order as one measure being more generalized than another.

Because of the additional equality condition, specialization is clearly a coarsening of the informativeness order. This condition is equivalent to requiring that no mass can be spread to the other side of or away from the prior. These one-sided spreads can be understood as learning about the quality of ones signal, in contrast to learning about the direction. The more specialized an agent is, the more the quality of his assessment depends on whether a given question is in his area of spezialization or not and hence the more he can learn about the quality of his judgment, based on observing the question.

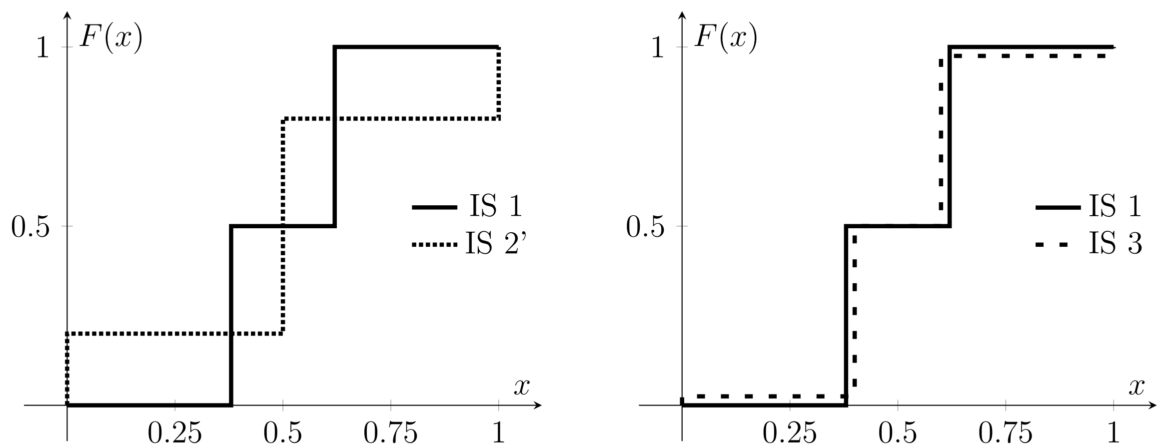
To illustrate the difference between specialization and informativeness order, let us get back to our example from the introduction with a slight adjustment. To recall, we have assumed a symmetric prior and discussed three possible signal structures for experts advising a regulator:

1. Each expert receives a signal that matches the true state of the world with a 62% chance.

2′. Each expert has a 40% chance to learn the state perfectly.

3. Each expert receives a signal that matches the true state of the world with a 60% chance, but also has a 5% chance to learn the state perfectly.

In contrast to the introduction, we have changed the probability that a expert receives a perfect signal in 2'. Figure 2 illustrates the cdfs corresponding to these information structures.



(a) Information Structure 2' is a mean-preserving spread of 1. However, 2' is not a mean-preserving spread on both sides of the prior separately. Hence, 1 and 2' are not comparable with the specialization order.

(b) Information Structure 3 can be constructed by applying mean-preserving spreads on both sides of the prior of Information Structure 1. Therefore, 3 is more informative and more specialized than 1.

Figure 2: Illustration of specialization. Distribution of virtual posteriors.

As we can see, 3 is a mean-preserving spread of 1, wheras 2' is a mean-preserving spread of both. However, the integrals are also equal at the prior only in the case of 3 and 1. We can interpret this spread as a expert under signal structure 3 receiving an additional signal that identifies 5% of his signals as fully revealing, without ever changing the direction of any of his signals.

Another possible way to understand specialization is to pose that between two experts that have equal information once one averages above their signal in either direction the specialist is more heterogeneously informed. We fix this idea in the following definition.[10]

Definition 4. *The average informativeness $\pi(\mu)$ of a sender's distribution over posteriors μ is*

$$\pi(\mu) = \mathbb{E}\left[|p_i - p_0|\right].$$

A distribution μ with $\pi(\mu) = 0$ has all mass at the prior and hence does not contain any information, whereas the maximal average informativeness is $2p_0(1 - p_0)$. Note that equal or higher average informativeness of μ compared to ν is a necessary but not a sufficient condition for μ to be more informative than ν.

In our example, average informativeness in 1 and 3 equals 0.12, although it is 0.2 in 2'. The following proposition clarifies the link between average informativeness and specialization:

[10] See Frankel and Kamenica [2019] for a critical discussion of the use of metrics, such as the euclidean metric used here, to quantify information. We should note however, that because we only use average informativeness in combination with the informativeness order, their criticisms generally do not apply to our setting.

Proposition 2. *Let μ and ν be two distributions over posteriors. Then $\mu \succ_s \nu$ if and only if $\mu \succ \nu$ and $\pi(\mu) = \pi(\nu)$.*

Proof. See Appendix A. □

By this characterization, a specialist is more informed not because his posterior is, on average, further away from the prior, but because his posteriors are more heterogeneous. The proof consists in arguing that the one-sided mean preserving spreads from the definition are exactly those that do not increase average informativeness.

In the remainder of this subsection, we connect the above insights to the receiver's utility. As more informativeness of independent individuals' distributions over posteriors produces more informativeness overall, as discussed in Blackwell and Girshick [1979] (see Proposition C in Appendix B), the utility of the receiver increases with the specialization of posteriors. Thus, the next corollary to Proposition 1 links our discussion in this section so far with the utility of the receiver.

Corollary 1. *Let μ and ν be distributions over posteriors with $\mu \succ \nu$. When comparing receiver optimal equilibria under common interests, the utility of the receiver is monotone with respect to the informativeness and hence the specialization order, i.e., weakly higher if she is facing senders with distribution over posteriors μ rather than ν.*

Proof. See Appendix A. □

As we learned previously that under common interests, all information held by the senders reaches the receiver and because more information is advantageous, the receiver directly benefits from more informative and therefore more specialized senders.

5 The Vulnerability of Specialized Advice - Private Interest Analysis

In this section, we turn to the case with private interests. In Subsection 5, we solve for the receiver-optimal equilibrium. In Subsection 5, we decompose the total loss of information into despecialization and a loss of average informativeness. Lastly, in Subsection 5, we present two consequences of despecialization. First, we show that voting is optimal if preferences are sufficiently heterogeneous. Second, average informativeness becomes more important and specialization less important as the number of partisans increases.

Receiver-Optimal Equilibrium

Similar to our treatment of the common-interest case we focus on a receiver optimal equilibrium. It turns out that an optimal equilibrium exists such that advisors play the truthful strategy, as in Proposition 1. Transmitting as much information as possible is in their best interest. Partisans interfere with this communication, by mixing over the most extreme messages in their preferred direction.

As the average posterior of a partisan sender equals the prior, their strategy shifts virtual posteriors towards the prior. They do not transmit any information to the receiver, but maximize their influence by imitating advisors with the most informative signals. Therefore, the receiver needs to discount these messages. This way, expertise bounds $\underline{b}$ and $\overline{b}$ arise. They constitute bounds on the highest (lowest) possible virtual posteriors associated with any message. The weights $w(\underline{b})$ and $(w(\overline{b}))$ are the lowest and highest weights used in the weighted majority rule of the receiver, respectively. They are endogenously determined by the sender type distribution.

Within the expertise bounds, communication between the advisors and the receiver is noise-free because partisans do not imitate advisors with imprecise signals. Thus, communication is perfect within these bounds, as in the equilibrium from Proposition 1. Off-equilibrium messages receive weight 0. Figure 3 depicts an example of a weighting function of virtual posteriors with upper and lower expertise bounds. We formalize the above discussion in our first theorem.

Theorem 1. *The following describes a receiver-optimal Bayesian Nash equilibrium. There exist unique expertise bounds $\underline{b}, \overline{b} \in [0,1]$, s.t.*

- *Advisors message truthfully, i.e., $m_i(p_i, \lambda_0)(p_i) = 1$.*
- *Partisans imitate and devalue expertise:*

$$m_i(p_i, 0)(t_i) = \begin{cases} \frac{\gamma(\lambda_0)\mu(t_i)(\underline{b}-t_i)}{\gamma(0)(p_0-\underline{b})} & \text{if } t_i \le \underline{b} \\ 0 & \text{else} \end{cases}$$

$$m_i(p_i, 1)(t_i) = \begin{cases} \frac{\gamma(\lambda_0)\mu(t_i)(t_i-\overline{b})}{\gamma(1)(\overline{b}-p_0)} & \text{if } t_i \ge \overline{b} \\ 0 & \text{else} \end{cases}$$

- *The receiver uses weighted majority rule with weight function*

$$w(x) = \begin{cases} \ln \frac{\underline{b}}{1-\underline{b}} - \ln \frac{p_0}{1-p_0} & \text{if } x < \underline{b} \\ \ln \frac{x}{1-x} - \ln \frac{p_0}{1-p_0} & \text{if } x \in [\underline{b}, \overline{b}] \\ \ln \frac{\overline{b}}{1-\overline{b}} - \ln \frac{p_0}{1-p_0} & \text{if } x > \overline{b} \\ 0 & \text{else} \end{cases}$$

and threshold $\tau = -\left(\ln \frac{\lambda_0}{1-\lambda_0} + \ln \frac{p_0}{1-p_0}\right)$.

Proof. See Appendix A. □

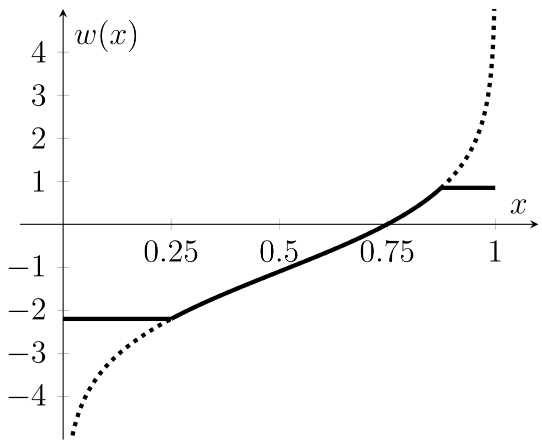


Figure 3: Weighting function $w(x) = \ln \frac{x}{1-x} - \ln \frac{p_0}{1-p_0}$ with prior $p_0 = \frac{3}{4}$, $\underline{b} = \frac{1}{4}$, and $\overline{b} = \frac{7}{8}$. The dashed line is the weighting function of the receiver in the absence of partisans.

We briefly sketch the proof, which consists of three steps. First, we provide an explicit formula for expertise bounds and show that they exist and are unique. This is formally captured in Lemma 1 in the Appenddix.

Second, we verify that the senders and the receiver play a best strategy. We show that the receiver uses weights corresponding to the expected posterior for every message. It is clear that partisans' play a best response by maximizing their weight in the preferred direction. Further, aligned senders ideally want their messages to be weighted according to the untruncated weighting function, yet given the receiver's strategy, they get as close as possible by stating their true posterior.

Lastly, we show that no other equilibrium is better for the receiver. For this, we first establish that the equilibrium utilities of all agents are characterized by the strategy played by senders conditioned on being advisors. We already saw in the common interest case that the receiver's on-path strategy is fully determined by the strategies used by the senders because he acts last.

This builds on the insight that whenever an advisory type strategy leads to more informative messaging in the absence of partisan types, the same holds true, when partisans are introduced.

This argument in lemma 2 in the Appendix. Optimality follows in combination with Proposition 1 that has established, that truthful messaging by advisors is optimal, in the absence of private interests.

We end the Subsection by briefly illustrating the theorem by means of the example from the introduction. In Information Structure 3, we assume that an expert gets a perfect signal with a 5% chance and otherwise a signal that is right 60% of the time. Under a symmetric prior, the senders' posterior is 0 or 1 and 0.4 or 0.6 with probabilities 2.5% and 47.5%, respectively. Now we introduce private interests in the form of a sender being a partisan type, with an equal chance of $\frac{1}{38} \approx 2.6\%$ on either side. Theorem 1 predicts the following effect on the information structure:

- Senders claim with probability $\frac{1}{38}+\frac{18}{19}\cdot 2.5\% = 5\%$ a posterior of 0 or 1, respectively. The receiver treats these messages as if everyone has truthfully reported a posterior of $\frac{5}{19}$ and $\frac{14}{19}$, respectively.
- Senders claim with probability $\frac{18}{19}\cdot 47.5\% = 45\%$ a posterior of 0.4 or 0.6, respectively. The receiver treats these messages as true statements.

The example illustrates how partisans undermine the messages of the most informed sender types, although not affecting the messaging of their less informative counterparts. We decompose the associated loss of information in the next Subsection into two parts that relate to the concepts introduced in Section 4.

Lack of Trust versus Lack of Competence

In this Subsection, we inspect the equilibrium derived in the last Subsection with respect to the notion of specialization introduced in Section 4. To do so, we introduce an intermediate regime between private interest and common interest. In this intermediate regime, there are non-advisory types, just as before but instead of having misaligned interests, these types are aligned but incompetent, i.e., they receive no signal.[11] A receiver-optimal equilibrium is characterized by full communication, just as in Proposition 1, with the only difference that incompetent agents always message uninformatively.

We make two principal observations. First, the possibility that a sender is incompetent has the same impact on average informativeness as the possibility that he is a partisan. Second, in contrast to this regime, partisanship also leads to an effective loss in specialization because the uninformative messages of partisans are treated the same as messages from senders with the most valuable signals. The most informative signals are thus diluted by signals that are, on average, uninformative. The distrust in messages with a reportedly high precision destroys the benefits a receiver can reap from the specialization of her senders.

Therefore, we can decompose the loss of information caused by partisanship in a loss of average informativeness and a loss in specialization, as formalized in the following theorem:

Theorem 2. *Let $\mu^{\gamma}_{partisan}$ and $\mu^{\gamma}_{incompetent}$ be the distribution over virtual posteriors in the equilibrium above and the most informative equilibrium with incompetent types, respectively. The loss of average informativeness in both cases is identical and equal to the probability that a sender is non-advisory:*

[11] The discussion below is equally valid for the case of a receiver who knows the identity of (informed) senders with partisanship. The receiver ignores any message from these sender types.

$$\pi(\mu^{\gamma}_{partisan}) = \pi(\mu^{\gamma}_{incompetent}) = \gamma(\lambda_0)\pi(\mu)$$

However, the virtual posteriors under partisanship are less specialized than under incompetence:

$$\mu^{\gamma}_{partisan} \prec_s \mu^{\gamma}_{incompetent}$$

Proof. See Appendix A. □

Let us briefly sketch the proof. First, note that average informativeness is reduced proportionally under ignorance almost axiomatically. Second, we show that the way partisans garble advisors' best signals with their own uninformative messages is a one-sided mean-preserving contraction of the distribution of virtual posterior under ignorance. Therefore, it is less specialized and, by Proposition 2, implies that both distributions have the same average informativeness.

Having studied the two distinct parts of information loss, we continue by pointing out two consequences of high degrees of partisanship. We demonstrate that sufficiently heterogeneous preferences can prevent any differentiated weighting of messages. Further, we show that the value of specialization vanishes and that average informativeness becomes more important as the share of partisans rises.

A Justification for Voting

In this Subsection, we analyze the effect of a rising share of partisans by building on the results from the previous two Subsections and the concepts from Section 4. We demonstrate that specialization vanishes as the share of partisans rises. This is reflected by an equilibrium in which the receiver weights all messages with the same direction equally. To see why such an equilibrium might evolve, consider the following reasoning. A partisan imitates senders with the most valuable signals. The receiver devalues these messages accordingly. If the share of partisans is high enough, the weight of the most informative and second most informative signals become equal. As the share of partisans increases further although assuming that no completely uninformed senders exist, only two distinct weights remain - one for each direction.

We interpret such an equilibrium as voting because only the number of senders messaging in each direction matters. Specifically, suppose that there are weights w^0 and w^1 for the senders' messages expressing that their posteriors are left and right from the prior, respectively. According to the equilibrium, the receiver takes action 1 if $n_0 \cdot w^0 + n_1 \cdot w^1 > \tau$, where n_0 and n_1 denote the number of senders with weight w_0 and w_1, respectively. This decision rule corresponds to a qualified majority rule that we consider as a form of voting. The formal statement of the result is as follows:

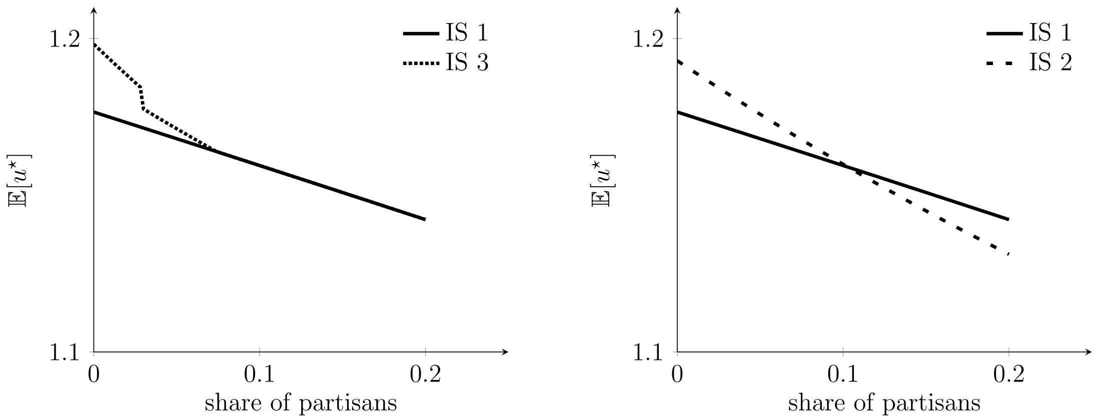
Theorem 3. *Let μ be never-ignorant. Then there exists $c_0, c_1 \in (0,1)$ with $c_0 + c_1 < 1$, s.t. for all γ with $\gamma(0) \geq c_0$ and $\gamma(1) \geq c_1$ the receiver forms only two virtual posteriors, i.e., (qualified majority) voting is the most informative equilibrium.*

Proof. See Appendix A. □

We find it instructive to illustrate this result by means of our example too. Consider Information Structure 3:

3. Each expert receives a signal that matches the true state of the world with 60%, but also has a 5% chance to learn the state perfectly.

If there are only a few partisans, all of them imitate senders with a perfect signal. But as the share of partisans rises, the weight that the receiver assigns to this message decreases and eventually equals to the weight of the initially less informative message. To be precise, equality is achieved if $\frac{1}{12}$ of the senders are partisan in either direction. Equilibrium play guarantees that the weight of both messages stays equal for any greater share of partisans. In fact, if at least $\frac{1}{12}$ of the senders are partisan on each side, Information Structure 3 collapses to Information Structure 1 in terms of the expected posteriors. See Panel (a) in Figure 4 for an illustration of the receiver's utility with Information Structure 1 and 3 in a setting with three senders.



(a) Theorem 3: Information Structure 1 vs. Information Structure 3

(b) Proposition 3: Information Structure 1 vs. Information Structure 2

Figure 4: Illustration of Theorem 3 and Proposition 3. Utility of the receiver under different Information Structures with an increasing share of partisans and $n = 3$ senders.

In such equilibria, the benefits of specialization are completely destroyed by partisans. The level of distrust is so high that effectively only the direction of a signal can be communicated.[12] At the same time, Theorem 3 underlines the robustness of voting.

[12]When the receiver is not indifferent at the prior, there always exist levels of partisanship s.t. the receiver completely ignores senders, because the virtual posterior distribution becomes to uninformative.

Thus, privately interested senders might prevent any communication of specialization and provide a justification for the prevalence of voting in many real-world information aggregation systems.

We demonstrated that specialization vanishes if there are sufficiently many partisans. Following Theorem 2, average informativeness decreases too. In the following we relate the information loss of both concepts to each other. Note that in an equilibrium where voting is the most efficient way to communicate, average informativeness becomes the most important statistic for the receiver because there is no more specialization. We capture this observation by the following Proposition:

Proposition 3. *Let μ and ν distributions over posteriors with $\pi(\mu) > \pi(\nu)$ and cdfs F and G, respectively. Then there exist $c_0, c_1 \in (0,1)$ with $c_0 + c_1 < 1$, s.t. for all γ with $\gamma(0) \geq c_0$ and $\gamma(1) \geq c_1$ and any number of senders n, we have that $\mu^\gamma \succ \nu^\gamma$ and hence the ex-ante expected utility of the receiver is weakly greater under distribution over posteriors μ than under ν.*

Proof. See Appendix A. □

Proposition 3 implies that the comparison of two information structures in terms of the receiver's utility depends on the number of partisans. See Panel (b) in Figure 4 for an illustration of the receiver's utility with Information Structure 1 and 2 in a setting with three senders. While Information Structure 2 leads to a higher utility of the receiver with only a few partisans, Information Structure 1 leads to a higher utility if the share of partisans is at least 10%. Information Structure 2 has a high degree of specialization that can only be materialized with low numbers of partisans. In contrast, Information Structure 1 has a higher average informativeness that becomes decisive if there are enough partisans. Taken together, specialization is particularly valuable with few partisans, and average informativeness is valuable with many partisans. Thus, even if partisans interfere, information aggregation guaranteeing that signals are good on average is an effective way to counteract.

6 Conclusion

Our goal in this research has been to understand the optimal communication of a decision maker with multiple advising experts when she faces uncertainty about experts' preferences. In particular, we have been interested in how these uncertainties affect changes among senders with different information structures.

This collapse of communication corresponds to a voting threshold below 1 or above n. Our interpretation of voting is therefor most natural, when either, the receiver is close or indifferent, the number of senders is high, or the least informative signals in either direction are relatively informative. The same point applies to Proposition 3

We have found that communication that discriminates between messages, indicating different degrees of confidence, is potentially very informative for the receiver, but also highly vulnerable to strategic manipulation by partisan experts. Consequently, such communication is not optimal in a case with high levels of partisanship. In contrast, binary communication protocols such as voting prove to be very robust, explaining their prevalence as a means for information aggregation.

Our research may also lead to a new approach towards questions regarding political lobbying. Much of the literature on the subject (see, for example, Buchanan et al. [1980] and Baye et al. [1993]) has focused on lobbying as a way in which special interest groups try to provide incentives for political actors, in order to sway them in their favored direction. It is, however, just as plausible for such groups to buy influence with advising experts to influence politicians' beliefs rather than offer direct incentives. Our work shows that this can be effective even if politicians are aware of it, as long as they remain ignorant about the exact identity of the experts who have been compromised. In particular, interest groups may seek to sometimes influence experts against their own favored decisions to create the justified belief that some experts advocating the other side are not trustworthy. When talk is cheap, trust is a valuable yet vulnerable asset.

A Appendix: Proofs

Proof of Proposition 1. In this proof, we follow closely standard arguments regarding the representation of Bayesian updating usings log-liklyhood ratios as they can be found, for example, in the proof of Theorem 1 in Nitzan and Paroush [1982], who derive the optimal non-strategic processing of signals with a symmetric prior $\lambda_0 = \frac{1}{2}$.

In the main text, we use the same notation for random variables and their realizations. For this proof, it is useful to introduce a separate notation. We use upper-case characters for random variables and lower-case characters for their realizations.

The receiver processes messages t to update her posterior. She prefers the action that yields the higher expected utility given her posterior $q(t)$. More precisely, an optimal decision rule selects action 1 if

$$
\begin{aligned}
& \lambda_0 + \mathbb{P}[\omega = 1 | T = t] > (1 - \lambda_0) + \mathbb{P}[\omega = 0 | T = t] \\
\Leftrightarrow \quad & \lambda_0 \mathbb{P}[\omega = 1 | T = t] > (1 - \lambda_0) \mathbb{P}[\omega = 0 | T = t] \\
\Leftrightarrow \quad & \lambda_0 \frac{\mathbb{P}[P = p | \omega = 1] \cdot \mathbb{P}[\omega = 1]}{\mathbb{P}[P = p]} > (1 - \lambda_0) \frac{\mathbb{P}[P = p | \omega = 0] \cdot \mathbb{P}[\omega = 0]}{\mathbb{P}[P = p]} \\
\Leftrightarrow \quad & \lambda_0 p_0 \prod_i \mathbb{P}[P_i = p_i | \omega = 1] > (1 - \lambda_0)(1 - p_0) \prod_i \mathbb{P}[P_i = p_i | \omega = 0] \\
\Leftrightarrow \quad & \lambda_0 p_0 \prod_i \frac{p_i}{p_0} > (1 - \lambda_0)(1 - p_0) \prod_i \frac{1 - p_i}{1 - p_0} \\
\Leftrightarrow \quad & \prod_i \left(\frac{p_i}{1 - p_i} \frac{1 - p_0}{p_0} \right) > \frac{1 - \lambda_0}{\lambda_0} \frac{1 - p_0}{p_0} \\
\Leftrightarrow \quad & \sum_i \left(\ln \frac{p_i}{1 - p_i} - \ln \frac{p_0}{1 - p_0} \right) > - \left(\ln \frac{\lambda_0}{1 - \lambda_0} + \ln \frac{p_0}{1 - p_0} \right).
\end{aligned}
$$

The first equivalence is a simple algebraic consequence of the fact that the first and second factors each add to one. For the second equivalence, we apply Bayes's rule and exploit the fact that senders play the truthful strategy. In the third step, we use the conditional independence of signals. We arrive at the fifth equation by applying Bayes's rule once again. The sixth equation is a simple reformulation of the fourth. Finally, we obtain the last equation by taking the logarithm on both sides. The resulting decision rule can be interpreted as a weighted majority rule with weighting function

$$
w(t_i) = \ln \frac{t_i}{1 - t_i} - \ln \frac{p_0}{1 - p_0}
$$

and threshold $\tau = - \left(\ln \frac{\lambda_0}{1 - \lambda_0} + \ln \frac{p_0}{1 - p_0} \right)$.

It is optimal for the senders to play the truthful strategy because senders and the receiver have the same utility function. With the truthful strategy, the senders can transmit all available information. Any beneficial transformation of messages can be done by the receiver. □

Proof of Proposition 2. We rewrite $\pi(\mu)$ until we arrive at an expression from which the result is immediate:

$$\begin{aligned}
\pi(\mu) &= \mathbb{E}\left[|p_i - p_0|\right] \\
&= \int_0^1 |x - p_0|\ \mu(x)\mathrm{d}x \\
&= \int_0^{p_0} (p_0 - x)\ \mu(x)\mathrm{d}x + \int_{p_0}^1 (x - p_0)\ \mu(x)\mathrm{d}x \\
&= p_0 \cdot F(p_0) - \int_0^{p_0} x\ \mu(x)\mathrm{d}x + \left(p_0 - \int_0^{p_0} x\ \mu(x)\mathrm{d}x\right) - p_0 \cdot (1 - F(p_0)) \\
&= 2\left(p_0 \cdot F(p_0) - \int_0^{p_0} x\ \mu(x)\mathrm{d}x\right) \\
&= 2\int_0^{p_0} F(x)\mathrm{d}x.
\end{aligned}$$

The fourth equation follows from the common prior $p_0 = \int_0^1 x\ \mu(x)\mathrm{d}x$ and the last equality from integration by parts. Therefore, two posterior distributions with the same prior have the same average informativeness if and only if they satisfy the integral condition for second-order stochastic dominance at the prior with equality. □

Proof of Corollary 1. In the receiver-optimal equilibrium from Proposition 1, we have seen that all the information reaches the receiver. When senders' signals are more informative and their information is independent, their joint information also becomes more informative by Proposition C in Appendix B. The receiver's utility then can only increase from more information reaching her.

□

Proof of Theorem 1, Part A: Expertise bounds:

Lemma 1. *For distribution over posteriors μ with cdf F and preference distribution γ, the lower expertise bound $\underline{b}$ in the receiver-optimal equilibrium is determined by*

$$\gamma(0)\,(p_0 - \underline{b}) = \int_0^{\underline{b}} (\underline{b} - x)\mathrm{d}\mu = \gamma(\lambda_0) \cdot \int_0^{\underline{b}} F(x)\ \mathrm{d}x,$$

and the upper expertise bound $\overline{b}$ is determined by

$$\gamma(1)\,(\overline{b} - p_0) = \gamma(\lambda_0) \int_{\overline{b}}^1 (x - \overline{b})\mathrm{d}\mu = \gamma(\lambda_0) \cdot \int_{\overline{b}}^1 1 - F(x)\ \mathrm{d}x.$$

Both equations have a unique solution for all μ and γ.

Proof of Lemma 1. As all messages $t_i \leq \underline{b}$ result in the same virtual posterior $\underline{b}$, we have

$$\frac{\gamma(\lambda_0) \int\limits_0^{\underline{b}} F(x)dx + \gamma(0)p_0}{\gamma(\lambda_0)F(\underline{b}) + \gamma(0))} = \underline{b}$$

An algebraic manipulation yields

$$\gamma(0)\left(p_0 - \underline{b}\right) = \gamma(\lambda_0) \int_0^{\underline{b}} (\underline{b} - x) \mathrm{d}\mu$$

as stated in the lemma. Note that the left side of the equation is strictly decreasing in $\underline{b} \in [0, p_0]$ and is 0 only if $\underline{b} = p_0$. The right side is weakly increasing in $\underline{b}$ and is 0 for $\underline{b} = 0$. Further, both sides are continuous in $\underline{b}$. Thus, there is a unique $\underline{b}$ that fulfills the equation.

The proof for the upper expertise bound is analogous.

□

Proof of Theorem 1, Part B: Strategies from Theorem 1 form an Equilibrium. As in the proof of Proposition 1, we use upper-case characters for random variables and lower-case characters for their realizations.

We start to calculate the virtual posterior $q(t_i)$. The only senders that send messages within the expertise bounds are advisors. Thus, $q(t_i) = t_i$ for $t_i \in (\underline{b}, \overline{b}) \cap \mathcal{P}$. For messages $t_i \leq \underline{b}$ with $t_i \in \mathcal{P}$, the virtual posterior of the receiver is

$$\begin{aligned}
q(t_i) &= \mathbb{P}\left[\omega = 1 | T_i = t_i\right] \\
&= \mathbb{P}\left[\omega = 1 | T_i = t_i \wedge \lambda_i = \lambda_0\right] \mathbb{P}\left[T_i = t_i | \lambda_i = \lambda_0\right] \\
&\quad + \mathbb{P}\left[\omega = 1 | T_i = t_i \wedge \lambda_i = 0\right] \mathbb{P}\left[T_i = t_i | \lambda_i = 0\right] \\
&= \frac{t_i \gamma(\lambda_0)\mu(t_i) + p_0\gamma(0)m(., 0)(t_i)}{\gamma(\lambda_0)\mu(t_i) + \gamma(0)m(., 0)(t_i)} \\
&= \frac{t_i + \dfrac{\underline{b} - t_i}{p_0 - \underline{b}} p_0}{1 + \dfrac{\underline{b} - t_i}{p_0 - \underline{b} p_0}} \\
&= \underline{b}.
\end{aligned}$$

The calculation for the virtual posterior of messages $t_i \geq \underline{b}$ with $t_i \in \mathcal{P}$ is $q(t_i) = \overline{b}$ by an analogous calculation. Thus, the receiver's on-path beliefs are consistent with Bayesian updating.

The technique of Nitzan and Paroush [1982] and the proof of Proposition 1 teach us how to process a set of (virtual) posteriors optimally. Again, the best response of the receiver can be interpreted as a weighted majority rule with weighting function

$$w(x) = \ln \frac{q(x)}{1-q(x)} - \ln \frac{p_0}{1-p_0} = \begin{cases} \ln \frac{\underline{b}}{1-\underline{b}} - \ln \frac{p_0}{1-p_0} & x \in \mathcal{P} \wedge x \leq \underline{b} \\ \ln \frac{x}{1-x} - \ln \frac{p_0}{1-p_0} & x \in \mathcal{P} \wedge \underline{b} \leq x \leq \overline{b} \\ \ln \frac{\overline{b}}{1-\overline{b}} - \ln \frac{p_0}{1-p_0} & x \in \mathcal{P} \wedge \overline{b} \leq x \\ 0 & \text{else} \end{cases}$$

and threshold $\tau = -\left(\ln \frac{\lambda_0}{1-\lambda_0} + \ln \frac{p_0}{1-p_0}\right)$.

We proceed by proving that senders play best responses. Partisans maximize the probability that the receiver takes the action that matches their preference parameter. Given the strategy of advisors and the receiver, they send a message with maximal weight in the preferred direction. In the equilibrium strategy 0- (1-) partisans mix over messages with weight $\ln \frac{\overline{b}}{1-\overline{b}} - \ln \frac{p_0}{1-p_0}$ $(\ln \frac{\underline{b}}{1-\underline{b}} - \ln \frac{p_0}{1-p_0})$ which is the highest (lowest) weight assigned by the receiver. Hence, these partisans play best responses.

The advisor and the receiver have the same utility function and prefer the same action when they have the same information. Thus, the best the advisor can do is to get as close to revealing all his information to the receiver as possible. Given any posterior the advisor holds he can either transmit his information noise-free if his posterior happens to be within the expertise bounds or their is a cap to what he can communicate, which he reaches, when sending his posterior as his message. He hence never faces the trade-off of having to over or undershoot in what he communicates and hence truthfully messaging is optimal.

Taken together, the strategies and beliefs in Theorem 1 are a weak perfect Bayesian equilibrium.

□

Proof of Theorem 1, Part C: Receiver Optimality . We show that the equilibrium in Theorem 1 is optimal for the receiver. We proceed in two steps. First, we introduce a technique that allows us to compare equilibria in the common interest case. This represents a more complete discussion than necessary to prove Proposition 1. Second, we show that the comparison carries over to the case with partisans. More concretely, we show that if an equilibrium in which advisors play the truthful strategy is more informative than another one in the case without partisans, it continues to be more informative than the other one in the presence of partisans.

The receiver bases her decision on the virtual posteriors $q(t_i)$, which she infers from messages t_i of senders $i = \{1, ..., n\}$. The same set of virtual posteriors leads to the same decision. The distribution of virtual posteriors $q(t_i)$ for sender i is determined by the distribution of posteriors μ and the sender i's strategy m_i.

Definition 5. *Let μ be a distribution of posteriors and m_i the strategy of sender i. We denote the distribution of virtual posteriors of sender i by $\mu^{\gamma}_{m_i}$ and define it by its cdf*

$$F^{\gamma}_{m_i}(x) = \mathbb{P}[q(t_i) \leq x],$$

where t_i is sender i's message. We suppress superscript γ in the common interest case, i.e., we write μ_{m_i} and F_{m_i} if $\gamma(\lambda_0) = 1$.

In the following, we compare the distribution over virtual posteriors of the equilibrium in which advisors play the truthful strategy with the distributions over virtual posteriors of other equilibria. We know from Proposition 1 that playing the truthful strategy is part of a receiver-optimal equilibrium for the common interest case. Using the concept of distributions over virtual posteriors helps us generalize this observation to the case with partisans.

The rest of the proof consists of three steps. First, we formalize that the virtual posterior distribution of a single sender is most informative in the common interest case, when advisors play the truthful strategy. Then second we show that even though partisans do not imitate and devalue expertise in all equilibria, whenever they do not, it never changes the action the receiver takes. We can hence ignore such equilibria when checking that our optimality candidate might be dominated. Lastly, we demonstrate that given that truth telling is optimal in the absence of partisans and given that partisans imitate and devalue expertise when they are added, truth telling remains optimal. These three statements are formalized in the three following Lemmas respectively.

Lemma 2. *Let μ be a distribution over posteriors, $m_i^\star$ a strategy in which advisors play truthfully, and m_i' any other strategy of sender i. Then, if $\mu_{m_i^\star}$ is more informative than $\mu_{m_i'}$, it follows that $\mu^{\gamma}_{m_i^\star}$ is more informative than $\mu^{\gamma}_{m_i'}$, i.e.*

$$\mu_{m_i^\star} \succ \mu_{m_i'} \Rightarrow \mu^{\gamma}_{m_i^\star} \succ \mu^{\gamma}_{m_i'}.$$

Lemma 3. *Let μ be a distribution over posteriors, $m_i^\star$ the truthful strategy, and m_i' any other strategy. Then it holds that $\mu_{m_i^\star}$ is more informative than $\mu_{m_i'}$, i.e., $\mu_{m_i^\star} \succ \mu_{m_i'}$.*

Lemma 4. *Take any weak Bayes Nash Equilibrium s.t. some senders' strategies conditioned on partisanship do not imitate and devalue expertise. Then the receiver's action is always the same as if they were to play this strategy.*

Given these three statements, the virtual posterior of sender i is most informative if types with $\lambda_i = \lambda_0$ play the truthful strategy. Again, under Theorem 12.3.2 in Blackwell and Girshick [1979] (see Proposition B in Appendix B), the sender-wise comparison carries over to the overall information structure. By Theorem 12.2.2 (4) in Blackwell and Girshick [1979] (see Proposition C in Appendix B), we conclude that no better equilibrium for the receiver than that described in Theorem 1 exists. We now proof the three lemmas we just applied.

Proof of Lemma 2. To simplify the notation, we denote $\mu_{m_i^\star}$ by μ, $\mu_{m_i'}$ by ν, $\mu^\gamma_{m_i^\star}$ by μ^γ, and $\mu^\gamma_{m_i'}$ by ν^γ. Further, we denote $F_{m_i^\star}$ by F, $F_{m_i'}$ by G, $F^\gamma_{m_i^\star}$ by F^γ, and $F^\gamma_{m_i'}$ by G^γ.

To prove that μ^γ is more informative than ν^γ, we show that

$$\int_0^y G^\gamma(x)\mathrm{d}x \leq \int_0^y F^\gamma(x)\mathrm{d}x \qquad \text{for all } y \in [0,1].$$

We first derive the following condition:

Let μ and ν with $\mu \succ \nu$ be distributions over posteriors with cdfs F and G, respectively. Let γ be the distribution of preference parameters. Then, the lower (upper) expertise bound $\underline{b}_\mu$ of μ is weakly smaller (greater) or equal than the lower (upper) expertise bound $\underline{b}_\nu$ of ν in the optimal equilibria with partisans.

This holds by the following argument. Suppose that $\underline{b}_\nu < \underline{b}_\mu$, and use Lemma 1 to see that

$$\begin{aligned}
\gamma(0)\left(p_0 - \underline{b}_\nu\right) &= \gamma(\lambda_0) \cdot \int_0^{\underline{b}_\nu} G(x)\mathrm{d}x \\
&\leq \gamma(\lambda_0) \cdot \int_0^{\underline{b}_\nu} F(x)\mathrm{d}x \\
&\leq \gamma(\lambda_0) \cdot \int_0^{\underline{b}_\mu} F(x)\mathrm{d}x = \gamma(0)\left(p_0 - \underline{b}_\mu\right).
\end{aligned}$$

Hence, $\underline{b}_\mu \leq \underline{b}_\nu$, which is a contradiction. The proof for the upper expertise bound is analogous.

Using this insight, it holds that $\underline{b}_\mu \leq \underline{b}_\nu$ and $\bar{b}_\mu \geq \bar{b}_\nu$. This allows us to check the inequality separately on the three intervals $[0, \underline{b}_\nu]$, $\left[\underline{b}_\nu, \bar{b}_\nu\right]$, and $\left[\bar{b}_\nu, 1\right]$.

For all $y \in [0, \underline{b}_\nu]$, it holds that

$$\int_0^y G^\gamma(x)\mathrm{d}x = 0 \leq \int_0^y F^\gamma(x)\mathrm{d}x.$$

For all $y \in \left[\underline{b}_\nu, \bar{b}_\nu\right]$ it holds that

$$\begin{aligned}
\int_0^y G^\gamma(x)\mathrm{d}x &= \int_{\underline{b}_\nu}^y \gamma(0) + \gamma(\lambda_0)G(x)\mathrm{d}x \\
&= \gamma(0)(y - \underline{b}_\nu) + \gamma(\lambda_0)\int_0^y G(x)\mathrm{d}x - \gamma(\lambda_0)\int_0^{\underline{b}_\nu} G(x)\mathrm{d}x \\
&= \gamma(0)(y - p_0) + \gamma(\lambda_0)\int_0^y G(x)\mathrm{d}x
\end{aligned}$$

$$
\begin{aligned}
&\leq \gamma(0)(y-p_0)+\gamma(\lambda_0)\int_0^y F(x)\mathrm{d}x \\
&= \int_0^y F^\gamma(x)\mathrm{d}x.
\end{aligned}
$$

The first equality follows from the definition of virtual posteriors and the equilibrium strategies. For the third equality, we apply Lemma 1. The inequality follows from the assumption that $\mu \succ \nu$.

As $G^\gamma(x) = 1$ for $x \geq \bar{b}_\nu$, it follows that for all $y \in \left[\bar{b}_\nu, 1\right]$, it holds that

$$
\int_y^1 G^\gamma(x)\mathrm{d}x \geq \int_y^1 F^\gamma(x)\mathrm{d}x.
$$

The expected value of both distributions is consistent with the common prior, i.e.,$\int_0^1 F^\gamma(x)\mathrm{d}x = \int_0^1 G^\gamma(x)\mathrm{d}x = 1 - p_0$. Thus, we conclude that

$$
\int_0^y G^\gamma(x)\mathrm{d}x \leq \int_0^y F^\gamma(x)\mathrm{d}x,
$$

for all $y \in [\bar{b}_\nu, 1]$, which concludes the proof. □

Proof of Lemma 3. Let $\mu_{m_i^\star}$ be the distribution over virtual posteriors under the truthful strategy $m_i^\star$. Any distribution $\mu_{m_i'}$ that is induced by another strategy m_i' can be constructed from $\mu_{m_i^\star}$ by an application of garblings. We do not restrict strategies to use only a finite set of messages. Therefore, we apply a result from Blackwell [1953] that generalizes Theorem 12.2.2 in Blackwell and Girshick [1979] (see Proposition A in Appendix B) to the case with continuous signals. Thereby, we conclude that $\mu_{m_i^\star}$ is more informative than $\mu_{m_i'}$. □

Proof of Lemma 4. We will prove this by contradiction. Let us assume that an equilibrium exists that violates the statement of the lemma. Recall that in any equilibrium, the receiver strategy is to follow a weighted majority rule. Fix a state of the world and senders' types (p, γ) s.t. if partisans were to imitate and devalue expertise, wlog action 1 would be taken, but instead, action 0 is taken. We will show that partisan types do not play a best response, thereby implying a contradiction.

In this state, if all senders with preference type 1 were to change their message to that associated with the highest virtual posterior, action 1 would be taken. If any partisan is pivotal, he is not best responding, and we directly have a contradiction.

Let us instead assume that multiple 1 partisans need to change their message, to change the action of the receiver and fix any ordering of them. Then there will be a first critical 1 partisan that can change the receiver's action provided that all partisans before him have already selected a message corresponding to the highest virtual posterior.

Now let us replace all partisan senders that were before the critical one with their respective aligned types that send the message associated with the highest virtual posterior. They need to exist in any non-babbling equilibria. This event happens with positive probability. In this situation the critical partisan is pivotal and hence his strategy was not a best response to begin with.

□

This concludes the proof of optimality. □

Proof of Theorem 2. The distribution over posteriors of a sender whose partisan types have been exchanged with incompetent types is given by:

$$\mu^{\gamma}_{incompetent} = \gamma(\lambda_0)\mu(x) + (\gamma(0) + \gamma(1))\,\delta_{p_0}$$

with average informativeness:

$$\pi(\mu_{incompetent}) = \gamma(\lambda_0)\pi(\mu).$$

The distribution over virtual posteriors derived from the equilibrium in Theorem 1 is given by

$$\mu^{\gamma}_{partisan}(x) = \begin{cases} \gamma(\lambda_0)F(x) + \gamma(0) & \text{if } x = \underline{b} \\ \gamma(\lambda_0)\mu(x) & \text{if } x \in \left(\underline{b}, \overline{b}\right) \\ \gamma(\lambda_0)\left(1 - F(x)\right) + \gamma(1) & \text{if } x = \overline{b} \\ 0 & \text{else.} \end{cases}$$

We first check whether $\mu_{partisan}$ is less informative than $\mu_{incompetent}$. For this, let us denote with $F^{\gamma}_{partisan}$, $F^{\gamma}_{incompetent}$, and F the cdfs of $\mu^{\gamma}_{partisan}$, $\mu^{\gamma}_{incompetent}$, and μ, respectively. We then must show that

$$\int_0^y F_{incompetent}(x)\mathrm{d}x \geq \int_0^y F_{partisan}(x)\mathrm{d}x \;\forall y \in [0,1].$$

When $y \in [0, \underline{b}]$, this is true, because $F_{partisan}$ is constant and equal to 0 on this interval. The case for the interval $y \in [\overline{b}, 1]$ follows by a symmetric argument, because the integrals become equal to the prior at $y = 1$. Let us hence focus on $y \in [\underline{b}, \overline{b}]$. We then get

$$\int_0^y F_{incompetent}(x)\mathrm{d}x = \int_0^y \gamma(\lambda_0)F(x) + (1 - \gamma(\lambda_0))\mathbb{1}\{x \geq p_0\})\mathrm{d}x$$

$$= \int_0^{\underline{b}} \gamma(\lambda_0)F(x)\mathrm{d}x + \int_{\underline{b}}^{y} \gamma(\lambda_0)F(x) + (1-\gamma(\lambda_0))\mathbb{1}\{x \geq p_0\})\mathrm{d}x$$

$$\geq \gamma(0)(p_0 - \underline{b}) + \int_{\underline{b}}^{y} \gamma(\lambda_0)F(x) + \gamma(0)\mathbb{1}\{x \geq p_0\})\mathrm{d}x$$

$$\geq \int_{\underline{b}}^{y} \gamma(\lambda_0)F(x) + \gamma(0)\mathrm{d}x$$

$$= \int_0^{y} F_{partisan}(x)\mathrm{d}x.$$

The first inequality makes use of Lemma 1. In the last equality, we use the fact that $F_{partisan}$ is equal to 0 on $[0, \underline{b}]$. Lastly, we verify the average informativeness of $\mu_{partisan}$ to be

$$\begin{aligned}\pi(\mu^{\gamma}_{partisan}) =& \int_0^1 |x - p_0|\mathrm{d}\mu_{partisan}\\ =&\gamma(\lambda_0)\int_0^1 |x-p_0|\mathrm{d}\mu + \gamma(0)(p_0 - \underline{b}) - \int_0^{\underline{b}} (\underline{b} - x)\mathrm{d}\mu + \gamma(1)(\overline{b} - p_0) - \int_{\overline{b}}^{1} (x - \overline{b})\mathrm{d}\mu\\ =&\gamma(\lambda_0)\pi(\mu).\end{aligned}$$

Here the last equation is a consequence of Lemma 1.

□

Proof of Theorem 3. We prove the proposition in two steps. We start to show that by monotonicity and continuity of $\underline{b}$ and $\overline{b}$, there exists $c_0, c_1 \in (0,1)$, such that the receiver can only form two expected posteriors in the optimal equilibrium. Then, we prove that there exist c_0 and c_1 such that $c_0 + c_1 < 1$. For both parts, we use Lemma 1, which characterizes the lower expertise bound by the equation

$$\gamma(0)\,(p_0 - \underline{b}) = \gamma(\lambda_0) \cdot \int_0^{\underline{b}} F(x)\mathrm{d}x.$$

The lower expertise bound can take any value in $\underline{b} \in [0, \max\{x : F(x) = 0\}]$ if $\gamma(0) = 0$. Further, it is p_0 if $\gamma(0) = 1$. Rewriting the above equation yields

$$\frac{\gamma(0)}{\gamma(\lambda_0)} = \frac{\int_0^{\underline{b}} F(x)\mathrm{d}x}{p_0 - \underline{b}} \tag{2}$$

which exhibits that $\underline{b}$ is monotonically increasing in $\gamma(0)$, monotonically decreasing in

$\gamma(\lambda_0)$, and continuous in $\gamma(0), \gamma(\lambda_0) \in (0,1)$.

As μ is never-ignorant, there exists a highest type strictly smaller than the prior, $p_L := \max\{x | x < p_0 \wedge x \in \mathcal{P}\}$. The proposition is fulfilled if the lower expertise bound equals this type $\underline{b} = p_L$. The continuity and monotonicity of $\underline{b}$ imply that the right-hand side of Equation (2) is positive and finite, and hence, $\gamma(0) < 1$ if $\underline{b} = p_L$. The proof for the upper part with type $p_H := \min\{x | x > p_0 \wedge x \in \mathcal{P}\}$ is analogous, so that constants $c_0, c_1 \in (0,1)$ are implicitly given by

$$c_0\,(p_0 - p_L) = \gamma(\lambda_0) \cdot \int_0^{p_L} F(x)\mathrm{d}x \quad \text{and} \quad c_1\,(p_H - p_0) = \gamma(\lambda_0) \cdot \int_{p_H}^{1} 1 - F(x)\mathrm{d}x. \tag{3}$$

To see that $c_0 + c_1 < 1$, divide Equations (3) by $(p_0 - p_L)$ and $(p_H - p_0)$, respectively. Adding both equations yields

$$c_0 + c_1 = \gamma(\lambda_0) \cdot \frac{\int_0^{p_L} F(x)\mathrm{d}x}{p_0 - p_L} + \gamma(\lambda_0) \cdot \frac{\int_{p_H}^{1} 1 - F(x)\mathrm{d}x}{p_H - p_0}.$$

As $\frac{\int_0^{p_L} F(x)\mathrm{d}x}{p_0 - p_L}, \frac{\int_{p_H}^{1} 1 - F(x)\mathrm{d}x}{p_H - p_0} > 0$, it follows that $\gamma(\lambda_0) > 0$. This implies that $c_0 + c_1 = \gamma(0) + \gamma(1) = 1 - \gamma(\lambda_0) < 1$, which completes the proof. □

Proof of Proposition 3. By Lemma 1, we know that as long as $\gamma(0)$ and $\gamma(1)$ remain above some fixed $\epsilon > 0$, we have that as $\gamma(\lambda_0)$ converges to 0, expertise bounds converge to the prior.

Hence, there exists c_0 and c_1 for which both μ^γ and ν^γ have, at most, one mass point above and below the prior, and for large enough partisanships, the masses of these points become arbitrarily similar, as almost no aligned senders are left. However, the relative difference in the average informativeness of both virtual posterior measures stays fixed. Therefore, there exist c_0 and c_1 for which the mass points, which are not at the prior, of μ^γ are necessarily further away from the prior than those of ν^γ. However, this implies that μ^γ is more informative than ν^γ, which implies our result by Proposition 2.

□

B Appendix: Basic Properties of the Informativeness Order

In this part of the appendix, we collect certain tools from the literature that we use throughout and that are related to definition 1. This allows us to compare the receiver's utility for different distributions over posteriors of senders. The methods and results in this Subsection are borrowed from Chapter 12 in Blackwell and Girshick [1979], building on majorization theory, first developed by Hardy et al. [1929]. To apply their machinery to our problem, we adjust our setting, and translate our notation into theirs.

The following results rely on the assumption that the action space of the receiver is a closed bounded convex subset of $\mathbb{R}$. To fulfill this assumption, we extend the action space of the receiver from $\{0, 1\}$ to $\Delta\{0, 1\}$, so that her action space is the interval $[0, 1]$. An action $a \in \Delta\{0, 1\}$ corresponds to the probability that the receiver takes action 1. Note that we can use this extended action space throughout the whole article without changing any result. In all statements on the receiver's best response, one of the two extreme actions $\{0, 1\} \subset \Delta\{0, 1\}$ is optimal. We use the action space $\{0, 1\}$ in the main text of the article to simplify the exposition.

To present the next results, it is also helpful to introduce some of the notation of Blackwell and Girshick [1979].[13] For a distribution over posteriors μ, we define a $2 \times N$ matrix P, where $N = |\mathcal{P}|$ is the number of possible posteriors. The rows represent the two states of the world 0 and 1. Each column represents one possible posterior. The value P_{ij} is the probability of observing the posterior represented by column j in state i. Note that matrix P is Markov, which means that $P_{ij} > 0$ for all i and j and that $\sum_{j=1}^{N} P_{ij} = 1$ for all i. With the notation, we are equipped to remind the reader of Theorem 12.2.2 in Blackwell and Girshick [1979].

Proposition A (Blackwell and Girshick [1979])**.** *Let P and Q be two $2 \times N_1$ and $2 \times N_2$ Markov matrices of distributions over posteriors μ and ν. μ is more informative than ν if and only if there is an $N_1 \times N_2$ Markov matrix M with $PM = Q$.*

Matrix M is said to *garble* information by transforming matrix P to Q. This means that distribution ν can be constructed from distribution μ. This interpretation justifies the statement that μ is more informative than ν.

The next result generalizes the previous proposition by enabling the comparison of sets of distributions. Each sender sends a conditionally independent posterior. Consider two sets of senders with different distributions over posteriors. Then, Theorem 12.3.2 in

[13] We also enjoyed reading the notes of Borgers [2009] on Chapter 12 of Blackwell and Girshick [1979] and borrow of his notation.

Blackwell and Girshick [1979] allows us to compare the information of both groups in the following sense.

Proposition B (Blackwell and Girshick [1979])**.** *Let* $(\mu_i)_{i=1}^n$ *and* $(\nu_i)_{i=1}^n$ *be two sets of distributions over posteriors. Suppose that* μ_i *is more informative than* ν_i *for every* i*. Then, the combination of distributions over posteriors* $(\mu_i)_{i=1}^n$ *is more informative than* $(\nu_i)_{i=1}^n$*.*

The proposition allows us to compare the information that is transmitted to the receiver from different distributions. Theorem 12.2.2 (4) in Blackwell and Girshick [1979] allows us to use this result for a statement on the utility of the receiver.

Proposition C (Blackwell and Girshick [1979])**.** *Let* μ *and* ν *be two distributions over posteriors such that* μ *is more informative than* ν*. Then for every continuous convex function* $\phi : [0,1] \to \mathbb{R}$*, we have*

$$\mathbb{E}_{\mu}\left[\phi(x)\right] \geq \mathbb{E}_{\mu'}\left[\phi(x)\right].$$

Note that the utility function $u^{\star}(q)$ is convex in q. Thus, if there are two distributions over posteriors with $\mu \succ \nu$, the proposition implies that the expected utility for the receiver with distribution μ is at least as high, as with distribution ν.

C Appendix: Discussion of Assumptions

More Actions

One can easily imagine situations in which a receiver not only has a binary choice but also might prefer to take any of a range of intermediate actions whenever his belief is sufficiently far away from either 0 or 1. In these cases, our analysis mostly generalizes as long as aligned senders still have identical utility functions as the receiver and the partisans, they still have monotone preferences with regard to the receiver's belief regardless of the true state of the world. To describe such more general models, one would likely abstract the exposition to a reduced form in which each agent's utility is simply a function of the receiver's belief. We have forgone this extra generality to present the model with a concrete and maximally simple decision problem. As a byproduct, the interpretation of our numerous illustrations is also simplified.

Also note that our focus on a binary decision problem is conservative in the role specialization and the loss of it plays for the receiver. Any additional action that does not dominate any previous actions can only make the receiver's utility more convex, hence increasing the gains from specialization.

Continuous Support

In this article, we focused on situations in which every signal only has a finite set of possible realizations. We do not see this as more or less natural than to assume a distribution over posteriors that is absolutely continuous but rather find that working with general probability measures would likely add little insight, yet complicate our exposition.

The only significant change that we anticipate, if one would pursue to rewrite this model with absolutely continuous posterior functions is a necessity to redefine *never-ignorant* as we have used it in Theorem 3, to mean that a neighborhood around the prior is not included in the support of posterior distribution. In essence, we require a lower bound of the information contained in every possible posterior, which in the discrete case, is given whenever the prior is not itself a possible posterior.

Heterogeneous Senders

We decided to restrict our analysis to symmetric senders for ease of exposition. However, in some of the applications, it stands to reason that the receiver can discriminate among the senders based on prior knowledge. A regulator might understand that one of his advising experts has previously worked on the approval of similar regulations and might hence believe his distribution over posteriors to be more informative than average.

As we have seen in our discussion of the symmetric senders, the receiver's learning from messages happens sender by sender, i.e., the message of one sender does not change how the message of another is processed. Consequently, all that changes in our model, when we allow for asymmetric senders, is that the weighting function that the receiver uses needs to be individualized based on each sender's posterior and preference distribution.

Learning about the informativeness and preferences of senders can, of course, only improve the receiver's situation as she can always choose to ignore that information regardless. Hence, our symmetric case can also be interpreted as a worst-case benchmark for the effect of partisanship on information transmission.

Departing from that benchmark, our previous analysis suggests that the largest gains from knowledge about the individual sender's informativeness and preferences are generated by the possibility of finding a specialist with a low probability of partisanship. As we have seen, specialization can have great benefits as long as the expert is also well trusted.

References

Ricardo Alonso, Wouter Dessein, and Niko Matouschek. When does coordination require centralization? *American Economic Review*, 98(1):145–79, 2008.

David Austen-Smith. Information transmission in debate. *American Journal of Political Science*, pages 124–152, 1990.

Yaron Azrieli. The price of 'one person, one vote'. *Social Choice and Welfare*, 50(2): 353–385, 2018.

Marco Battaglini. Public protests and policy making. *The Quarterly Journal of Economics*, 132(1):485–549, 2017.

Michael Baye, Dan Kovenock, and Casper De Vries. Rigging the lobbying process: an application of the all-pay auction. *The American Economic Review*, 83(1):289–294, 1993.

Roland Benabou and Guy Laroque. Using privileged information to manipulate markets: Insiders, gurus, and credibility. *The Quarterly Journal of Economics*, 107(3):921–958, 1992.

David Blackwell. Comparison of experiments. Technical report, Howard University, Washington United States, 1950.

David Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, pages 265–272, 1953.

David Blackwell and Meyer Girshick. *Theory of Games and Statistical Decisions*. Courier Corporation, 1979.

Tilman Borgers. Notes on Blackwell's comparison of experiments, 2009. URL `http://www-personal.umich.edu/~tborgers/Blackwell-Girshick.pdf`.

James Buchanan, Robert Tollison, and Gordon Tullock. *Toward a theory of the rent-seeking society*. Number 4. Texas A-&-M University Press, 1980.

Marquis de Condorcet. Essay on the application of analysis to the probability of majority decisions. *Imprimerie Royale Paris*, 1785.

Vincent Crawford and Joel Sobel. Strategic information transmission. *Econometrica*, pages 1431–1451, 1982.

Ronald A Dye. Disclosure of nonproprietary information. *Journal of accounting research*, pages 123–145, 1985.

Timothy Feddersen and Wolfgang Pesendorfer. The swing voter's curse. *The American Economic Review*, pages 408–424, 1996.

Timothy Feddersen and Wolfgang Pesendorfer. Voting behavior and information aggregation in elections with private information. *Econometrica*, pages 1029–1058, 1997.

Alexander Frankel and Emir Kamenica. Quantifying information and uncertainty. *American Economic Review*, 109(10):3650–80, 2019.

Dino Gerardi, Richard McLean, and Andrew Postlewaite. Aggregation of expert opinions. *Games and Economic Behavior*, 65(2):339–371, 2009.

Thomas Gilligan and Keith Krehbiel. Asymmetric information and legislative rules with a heterogeneous committee. *American Journal of Political Science*, pages 459–490, 1989.

Jacob Glazer, Helios Herrera, and Motty Perry. Fake reviews. Working paper, 2019.

Godfrey Hardy, John. Littlewood, and George Polya. Some simple inequalities satisfied by convex functions. *Messenger of Mathematic*, 58(58):225–243, 1929.

Patrick Hummel, John Morgan, and Phillip Stocken. A model of flops. *The RAND Journal of Economics*, 44(4):585–609, 2013.

Jordi Blanes i Vidal. Credibility and strategic communication: Theory and evidence from securities analysts. Technical report, 2006.

Woon-Oh Jung and Young K Kwon. Disclosure when the market is unsure of information endowment of managers. *Journal of Accounting research*, pages 146–153, 1988.

Kyungmin Kim and Jonathan Pogach. Honesty vs. advocacy. *Journal of Economic Behavior & Organization*, 105:51–74, 2014.

Vijay Krishna and John Morgan. A model of expertise. *The Quarterly Journal of Economics*, 116(2):747–775, 2001.

Hao Li, Sherwin Rosen, and Wing Suen. Conflicts and common interests in committees. *American Economic Review*, 91(5):1478–1497, 2001.

Ming Li and Kristóf Madarász. When mandatory disclosure hurts: Expert advice and conflicting interests. *Journal of Economic Theory*, 139(1):47–74, 2008.

Andrew McLennan. Consequences of the Condorcet jury theorem for beneficial information aggregation by rational agents. *American Political Science Review*, 92(2):413–418, 1998.

Joseph McMurray. Ideology as opinion: A spatial model of common-value elections. *American Economic Journal: Microeconomics*, 9(4):108–40, 2017.

John Morgan and Phillip C Stocken. An analysis of stock recommendations. *RAND Journal of conomics*, pages 183–203, 2003.

Shmuel Nitzan and Jacob Paroush. Optimal decision rules in uncertain dichotomous choice situations. *International Economic Review*, pages 289–297, 1982.

Lloyd Shapley and Bernard Grofman. Optimizing group judgmental accuracy in the presence of interdependencies. *Public Choice*, 43(3):329–343, 1984.

Joel Sobel. A theory of credibility. *The Review of Economic Studies*, 52(4):557–573, 1985.

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