

Delegated Expertise, Authority, and Communication

Inga Deimen Dezső Szalay^{*}
University of Bonn and University of Bonn

and CEPR

July 2016

Abstract

A decision-maker needs to make a choice and is forced to rely on delegated expertise. The expert's and the decision-maker's interests are imperfectly correlated and the expert decides how much to learn about each of their interests. We compare two institutions of decision-making, delegation versus communication. The model features endogenous, information driven conflicts at the communication stage. To limit the losses arising from strategic communication, the expert acquires information relevant to both the decision-maker and himself; in contrast, information acquisition under delegation is completely selfish. Hence, as a complement to delegated expertise, the decision-maker unambiguously prefers to communicate.

JEL: D82

Keywords: endogenous information, delegated expertise, strategic information transmission, delegation

*We thank Andreas Blume, Peter Eső, and Ines Moreno-de-Barreda for useful discussions and seminar participants at ESSET 2015 (Gerzensee), Bonn, Columbia, Crest, and Zürich for comments. Financial support by the DFG in the form of a SFB-TR15 grant is gratefully acknowledged. All remaining errors are our own.

Send communications to ideimen@uni-bonn.de or to szalay@uni-bonn.de or by ordinary mail to Dezső Szalay, Institute for Microeconomics, University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany.

1 Introduction

Good decision-making requires good information. Except perhaps for routine decisions, such information is not readily available but must be actively acquired. Pressed for time, decision-makers often have to delegate this job to others. We take this situation of delegated expertise as our starting point and wonder what mechanism of decision-making should ideally complement it? Should the decision-maker delegate decision-making to the expert too, or should she keep authority over decision-making and have the expert report back to her? We show that, as a complement to delegated expertise, communication unambiguously dominates delegated decision-making.

We envision an environment where the decision-maker must rely on an expert who has interests that are positively but not perfectly correlated with her own. Obviously, a completely like-minded expert would be preferred, but unfortunately such an ideal expert is generally out of reach. We assume that the expert has discretion over the breadth as well as the depth of the issues he investigates. More precisely, he chooses whether to investigate matters of interest to the decision-maker or to himself and he chooses how much information to acquire on each dimension. While the choices along these margins are observable, they are not contractible. The nature and precision of the expert's information is thus subject to moral hazard. Moreover, the nature and precision of information shape the conflicts between the expert and the advisee in communication and decision-making. If the decision-maker knows that the expert has primarily looked into matters of direct interest only to himself, then she should be cautious in following the expert's advice. She knows that the expert tends to be overly enthusiastic and hence she discounts the expert's advice to undo his cockiness. If, on the other hand, the expert looked into issues of relatively greater interest to the decision-maker than to himself, then it would be the expert who finds the decision-maker too cocky. Hence, the expert would become reluctant to share his information. Thus, the information acquired by the expert shapes the conflicts between expert and advisee.

How does the mechanism of decision-making interact with such endogenous, information driven conflicts? If the expert has authority over decision-making, then nothing keeps him from only investigating matters that are of interest to himself exclusively. Clearly, the decision-maker can rely on a perfect use of information by the expert, but the information is not as useful to her as she would have wished. In contrast, communication serves as a

safeguard mechanism against selfish information acquisition. The advice of an expert who is known to be too cocky - biased towards exaggeration in the jargon of the literature - is discounted. As a result, information can only be transmitted in a coarse fashion resulting in losses from strategic communication. So, the expert now faces a trade-off when acquiring relatively more information on matters he is interested in. On the one hand, the expert has by definition a preference for such information, because it is intrinsically more useful to him. On the other hand, the losses due to strategic communication are the larger the more selfish his information acquisition, because such information results in a more pronounced bias in communication. In our model, the two forces exactly offset each other. Intrinsically more useful information loses its appeal to the expert altogether, because its added value is exactly lost in strategic communication. Hence, acquiring information that is equally useful to the decision-maker and the expert is an equilibrium in our game. As a result, all biases in communication between the expert and the decision-maker are eliminated.

The insights from this story for organizational design are straightforward: the decision-maker is better off when retaining authority over decision-making. To avoid the consequences from ineffective communication, the expert internalizes the advisee's interests when acquiring information. In contrast, nothing stops the expert from selfish information acquisition under delegation. So, communication outperforms delegation because it changes the information the expert has. We prove these results in an environment where the expert's gains and losses from more selfish information acquisition exactly offset each other. However, all that is required is that the increased losses through strategic communication are the weakly dominating force, so the result is clearly robust beyond the specific environment.

The present paper extends a line of inquiry initiated by Deimen and Szalay (2015), DS henceforth. DS study decision-making in an organization with three parties, a sender, a receiver, and a designer, interested in aggregate surplus. E.g., such a situation arises in an organization with two divisions and a common headquarters. The roles in the organization are exogenously fixed; one division is endowed with the technology to acquire information, the other division has authority over decision-making. Headquarters steers the decision-making process by controlling the relative precision of information about issues of interest to one or the other division. Given an optimal information structure, it is irrelevant which of the divisions has formal authority over decision-making; the optimal policy from division

one's, the sender's, perspective is implemented regardless of who has formal authority.

There are several connections to and differences from our previous investigation. In contrast to DS, the allocation of authority does matter here through its impact on the expert's incentives to acquire information. Similar to DS, the model has an equilibrium outcome in which the expert acquires balanced information that eliminates conflicts in communication. DS derives this outcome as the optimum of a designer's (i.e. a headquarters) problem; the present paper derives the same outcome as an optimum from an expert's perspective. While DS sidestep a detailed analysis of communication under conflicting interests, the present analysis needs to understand precisely how the equilibrium value of decision-making is affected by the information that is acquired to understand the expert's incentives to acquire this or that piece of information. This exercise requires a closed form representation of expected utilities for all information structures that the expert considers as possibly optimal. This is infeasible in the more general statistical model of DS, so we assume a special case of theirs. Thus, our companion paper is more general in terms of assumptions, but the present analysis is far more complex.

Sobel (2013) identifies the acquisition of information as one of the important open questions for the theory of strategic communication. The problem is complex to analyze by the very nature of information: since Blackwell (1951) we think of better information in terms of more dispersed distributions of posteriors. However, such dispersed distributions make it impossible to characterize communication equilibria in closed form, a seemingly necessary step to characterize the equilibrium value of decision-making in closed form. Perhaps the most surprising of our results is that this logic is flawed. We develop a stylized model that captures the essential features of improvements in information *and* allows for a closed form representation of the value of decision-making *despite* the fact that the equilibrium itself cannot be characterized in closed form. It is this closed form representation of values that allows us to quantify the losses arising from strategic communication as a function of the sender's bias - thus ultimately, the quality of the sender's information. The essential assumption that generates these insights is that the joint distribution of states and signals is a multivariate Laplace, thus an appealing but nevertheless quite specific statistical environment.¹ Once

¹See Lehmann (1988) for a proof that Laplace location experiments are Blackwell comparable. The Laplace is a member of the elliptical class studied in DS; hence we analyze a special case of DS (but in

the hard work is done, the comparison of institutions boils down to a back-of-the-envelope computation that reveals an effect that we have not yet seen in the literature nice and simple: communication is helpful to direct an expert's search for information.

Delegation versus communication is the key question in Dessein (2002). Dessein (2002) compares the performance of these institutions in the seminal model of Crawford and Sobel (1982), where the sender's and the receiver's ideal choices differ by some constant. Our model features a very different and purely information driven bias: depending on the nature of information, the expert and the advisee differ in their responsiveness to new information. This is exactly the right assumption to make to study conflicts that arise purely on informational grounds. On top, we show in our companion paper, that the situation arises precisely when the sender and the receiver agree ex ante on a transfer price based on the action taken. The transfer price can be set based on prior information only and eliminates all biases that are known already ex ante, so that the only conflicts that remain are those that arise due and from the information that is acquired. Dessein (2002) shows that a constant bias has very different effects from those studied here: delegation always outperforms communication whenever meaningful communication is possible. An interesting question is how the classical bias interacts with information driven biases. We leave this for future work.

Alonso et al. (2008) compare centralized and decentralized forms of decision-making. It is a question of the magnitude of biases whether communication (centralized decision-making) or delegation (decentralized decision-making) dominates. For large biases, delegating is worse than deciding based on prior information alone. On the other hand, communication is always valuable. Hence, communication clearly dominates. As the sender's bias is decreased, delegation performs better and eventually outperforms communication. In contrast, communication is always the preferred mechanism in our paper, as long as interests are positively correlated but no matter how biased the sender is. The reason is that biases and information are unrelated in Alonso et. al (2008) whereas the sender's bias is a function of his information in our work. For the Pareto efficient information structure, the trade-off between communication and delegation is unambiguously resolved in favor of communication. However, as should be stressed, if a different information structure is selected for some reasons, perhaps

much more detail).

errors in the process of information acquisition, then the same trade-offs as in Alonso et al. (2008) arise here as well.

The term “delegated expertise” was coined by Demski and Sappington (1987), where an expert is defined as a person who can acquire information while others cannot. In contrast to the present paper, among many other differences, communication is prohibitively costly in their work. The organization of delegated expertise is studied in Lewis and Sappington (1997) and Gromb and Martimort (2007). Lewis and Sappington (1997) study information acquisition by an agent in a procurement context and show that information acquisition and production should ideally be delegated to different agents. Even though obtained under very different assumptions - e.g. perfect commitment and allowing for monetary payments - our result has a similar flavor: it is important that the expert, who acquires information, is not entitled to take productive decisions. Gromb and Martimort (2007) study the organization of delegated expertise allowing for collusion.

Commitment to decision-rules is an essential building block of the Demski and Sappington (1987) approach. Following this tradition, Crémer and Khalil (1992), Crémer et. al (1998), and Szalay (2009) study information acquisition in a procurement context allowing for monetary payments. Szalay (2005) studies a communication model with commitment to decision-rules with and without allowing for money payments. The main difference to this literature is the absence of commitment, a natural assumption that the literature following Crawford and Sobel (1982) maintains.

Our analysis is closely related to Argenziano et al. (forthcoming) and inspired by the questions raised there, in particular, whether delegated decision-making or communication is preferable when information needs to be acquired. However, the details of the models are very different - such as the information acquisition technology and the role of ex ante known conflicts. Argenziano et al. (forthcoming) show that communication can outperform delegation, because the expert acquires more precise information than the decision-maker would have acquired. In contrast, our model features different sources of information and we show that communication is helpful to convince the expert to look at all of them equally. Ivanov (2010) studies informational control in the communication model, that is, the receiver chooses what information the sender should observe. Informational control can outperform delegation to a completely informed sender. The main differences to the current model are

that there are two sources of information here and the expert has discretion over information acquisition.

Only quite few papers have looked at information or information acquisition in the communication model. Moscarini (2007) studies the effects of better information on equilibrium communication by a central banker. Esó and Szalay (2015) study the role of the richness of language for incentives for information acquisition. Di Pei (2015) assumes the sender can acquire coarse information. More precisely, the sender can partition the state space at a cost; the sender only acquires information that he plans to communicate. The implications for equilibrium communication are quite different from those in Crawford and Sobel (1982). Frug (2016) studies sequential information acquisition in a communication model. Ottaviani and Sorensen (2006) study communication when the sender wishes to appear well informed. Kamenica and Gentzkow (2011) study what information a sender would like a receiver to have. In contrast to Kamenica and Gentzkow (2011), the expert in the present model cannot commit on what information to pass on to the receiver once he (the expert) has observed the information.

The literature on communication is vast. We focus here on contributions with a close connection to information acquisition and leave out many important contributions. We refer to Sobel (2013) for an in depth survey of the literature. We end this review with some more technical remarks. Communication involves a limiting case where the number of induced actions goes to infinity as in Alonso et. al (2008) and Gordon (2010). The latter paper studies a very general model to obtain this characterization. We confirm the result of Gordon (2010) on the existence and non-existence of equilibria inducing a countable infinity of receiver responses. However, even though these features are the same, we cannot rely on the same methods, because our model has an unbounded state space in contrast to Gordon (2010). Moreover, our main contributions are to relate the incidence of infinite versus finite equilibria to the underlying information and to provide a closed form characterization of the value of information. None of these results appear in Gordon (2010). Alonso et. al (2008) characterize communication equilibria in closed form assuming a uniform type distribution. While we would love to follow this approach, we cannot, because uniform distributions do not allow us to capture the quality of information in a satisfactory way. The idea to characterize values of communication without necessarily characterizing equilibria explicitly appears of

course also in the analysis of optimal mediation rules as in Goltsman et al. (2009) and Alonso and Rantakari (2013). In contrast to that, our approach involves one shot communication and hence relies on very different methods.²

The remainder of the paper is organized as follows. In Section two, the model is introduced; the structure of communication equilibria is discussed in Section three; the value of communication is derived in Section four; in Section five, we study information acquisition. Section six contains our main results, the comparison of institutions. The final section concludes. All technical proofs are gathered in an appendix. Results of a mostly technical nature are stated as lemmata; technical results that are essential to understand the main trade-off are stated as propositions; economically deep results are stated as theorems.

2 The model

2.1 The decision problem

A decision-maker needs to reach a decision $y \in \mathbb{R}$. The ideal decision from her point of view depends on a state of the world, $\omega \in \mathbb{R}$. More precisely, the payoff of the decision-maker is

$$u^r(y, \omega) = -(y - \omega)^2,$$

where superscript r refers to receiver. The trouble is that the decision-maker does not know ω . However, before taking the action, she can consult an expert, henceforth referred to as the sender. The sender's preferences over actions are given by the function

$$u^s(y, \eta) = -(y - \eta)^2,$$

where η is the realization of a random variable that is correlated with ω . In the terminology of the literature, the difference $\eta - \omega$ corresponds to a state dependent bias. However, in contrast to the literature, conditional on ω , the bias is still random here.

²See also Blume et al. (2007) for a derivation of the noise mechanism identified as an optimal mediation rule in Goltsman et al. (2009).

Alonso and Rantakari (2013) discuss a limiting case of a truncated Laplace model. We go beyond the limiting case and characterize the equilibria of the model in general. Our characterization of communication values does not rely on optimal mediation but on one-shot communication.

The sender does not know the states ω and η either. However, he can observe noisy signals about the realized states according to

$$s_\omega = \omega + \varepsilon_\omega$$

and

$$s_\eta = \eta + \varepsilon_\eta.$$

To obtain a flexible and still tractable environment, we assume that the random vector $(\omega, \eta, \varepsilon_\omega, \varepsilon_\eta)$ follows a joint Laplace distribution³ where each of the marginals has a mean of zero⁴, $(\varepsilon_\omega, \varepsilon_\eta)$ is uncorrelated with (ω, η) and ε_ω is uncorrelated with ε_η . The nontrivial second moments of the distribution are denoted $Var(\omega) = \sigma_\omega^2$, $Var(\eta) = \sigma_\eta^2$, $Cov(\omega, \eta) = \sigma_{\omega\eta}$, $Var(\varepsilon_\omega) = \sigma_{\varepsilon_\omega}^2$, and $Var(\varepsilon_\eta) = \sigma_{\varepsilon_\eta}^2$. The variances $\sigma_{\varepsilon_\omega}^2$ and $\sigma_{\varepsilon_\eta}^2$ capture the amount of noise in the sender's signals.

The Laplace is a member of the elliptical class⁵ studied in DS. The reason to assume this particular stochastic environment is that the inference problems remain tractable. We reproduce the basic properties of the statistical environment and the analysis of DS for convenience of the reader below.

2.2 Timing

The strategic interaction unfolds as follows. Firstly, the decision-maker commits to an institution of decision-making. Either, she delegates both information acquisition and decision-making to the sender or she delegates information acquisition only to the sender and retains the right to choose y herself. In both cases, the decision-maker is forced to delegate information acquisition to the sender, due to, say, lack of time to acquire information herself. Secondly, the sender chooses what information to acquire. Formally, the sender chooses the

³For reasons that become obvious shortly, we defer a description of the density to Section 2.5 below.

⁴As mentioned in the introduction, such a situation arises, e.g., if a transfer price based on the action corrects for differences in prior expectations.

⁵If the distribution admits a density, then the density of a random vector \mathbf{t} that follows an elliptical distribution is of the form $f(\mathbf{t}) = k_n |\Sigma|^{-\frac{1}{2}} \phi((\mathbf{t} - \mu)' \Sigma^{-1} (\mathbf{t} - \mu))$, where k_n is a constant scale factor, ϕ a generator function, μ is the mean vector, and Σ is up to a scale factor equal to the variance matrix. E.g., for the Normal distribution, the generator function is $\phi = e^{-\frac{u}{2}}$.

amount of noise in the signals, i.e. the variances $\sigma_{\varepsilon_\omega}^2$ and $\sigma_{\varepsilon_\eta}^2$. We call the joint distribution of signals and states an *information structure*. The choice of the information structure is observed by the receiver. However, the realizations of the signals are privately observed by the sender. Finally, actions are chosen according to the selected institution of decision-making. Under delegation, the sender picks his preferred action policy. Under communication, the sender communicates with the receiver - formally, he sends a message to the receiver - and the receiver selects her preferred action, given the information that she has received. The receiver is unable to commit to an action policy before she receives the information.

The sender's choice of information structure is observable but not contractible. The sender therefore chooses the information structure with a view to using the information to his advantage in the selected institution of decision-making. The analysis of the resulting trade-offs are the subject of the present paper. All information structures are equally costly in our analysis. This allows us to focus on the purely strategic reasons to select different information structures.

2.3 Ideal policies and sufficient statistics

If both the sender and the receiver observed $\mathbf{s} = (s_\omega, s_\eta)$ directly, then their ideal choice functions would be given by

$$y^r(\mathbf{s}) = \mathbb{E}[\omega | s_\omega, s_\eta] = \alpha^r s_\omega + \beta^r s_\eta$$

and

$$y^s(\mathbf{s}) = \mathbb{E}[\eta | s_\omega, s_\eta] = \alpha^s s_\omega + \beta^s s_\eta$$

where α^j, β^j for $j = r, s$ are constants, independent of the realized signals. The first equality is due to the fact that the quadratic loss function is maximized at the conditional mean. The second equality follows because the Laplace distribution has a linear conditional mean, as all members of the elliptical class have. We refer the reader to DS for the exact expressions for the weights α^j, β^j for $j = r, s$.

Of course, the receiver does not have direct access to the sender's information but only to the sender's recommendation. Let

$$\rho \equiv \frac{\sigma_{\omega\eta}}{\sigma_\omega \sigma_\eta}$$

denote the coefficient of correlation between ω and η . The model is interesting only if $\rho > 0$, because no meaningful communication is possible if $\rho \leq 0$. Therefore, we assume that $\rho \in (0, 1)$. For this case, the first-best optimal policy functions of sender and receiver feature $y^r(\mathbf{s}) \neq y^s(\mathbf{s})$ for all $\mathbf{s} \neq (0, 0)$; thus, the model features conflicts of interest almost surely.

We focus on Bayesian equilibria in the communication game. After observing signal realizations s_ω, s_η , the sender sends a message $m \in \mathbb{M}$ to the receiver. The message space is sufficiently rich; we do not impose any restrictions on \mathbb{M} . A pure sender strategy maps the sender's information into messages $M : \mathbb{R}^2 \rightarrow \mathbb{M}, (s_\omega, s_\eta) \mapsto m$. A mixed sender-strategy is a probability distribution over pure strategies. A pure receiver strategy maps messages into actions, $X : \mathbb{M} \rightarrow \mathbb{R}, m \mapsto y$. As is well known, the receiver never mixes, due to the concavity of her payoff function. The receiver updates her belief about the sender's type after observing the sender's message and acts optimally against this belief. Define

$$\theta \equiv \mathbb{E}[\eta | s_\omega, s_\eta].$$

DS show that all sender types whose signals aggregate to the same value of θ have the same preferences. More precisely, such sender types share the same ideal policy and their preferences over any pair of choices depend only on the distance of the induced action to their ideal policy. This makes it difficult to elicit more than θ from the sender. In fact, any equilibrium is essentially equivalent to one with communication about θ only. As is standard, we can characterize any equilibrium of this kind as a partial pooling equilibrium, where sets of sender types pool on inducing the same receiver response. So, without loss of generality, we can eliminate the underlying signals (s_ω, s_η) from the picture and analyze a reduced form model where everything is as if the sender directly observed an aggregated signal θ . For the sender, θ is a sufficient statistic for the underlying signals; for the receiver, the underlying signals can be dropped, because the sender is never kind enough to reveal them.

Note that $\mathbb{E}[\eta | s_\omega, s_\eta]$ is a linear function of the signals, which are themselves linear in the underlying random variables. It follows from Kotz et al. (2001) that the joint distribution of the random vector (ω, η, θ) is a multivariate Laplace and that the marginal distribution of θ is a Laplace distribution too.⁶

⁶Note that the distribution of $\omega | \sigma$ is in general not Laplace. Fortunately, we are only interested in the first moment of the conditional distribution and its distribution, not in the posterior distribution itself.

2.4 Information and conflicts

Some caution is required when condensing the model into its reduced form. It must be true that the joint distribution of (ω, η, θ) can be generated from the underlying joint distribution by the sender's Bayesian updating.. In particular, the sender observes the underlying signals, s_ω and s_η , and forms a posterior expectation on η conditional on the signal realizations. Since θ is a function of the signals, the joint distribution is endogenous. The first moments are all zero; in particular, we necessarily have $\mathbb{E}[\theta] = 0$. The second moments are as follows. Depending on the underlying noise, the signal θ covaries more with one or the other underlying state variable. It is straightforward to show - by brute force algebra using the exact values of α^s and β^s of the sender's posterior expectation - that

$$Var(\theta) = \sigma_\eta^2 \frac{\frac{\sigma_{\varepsilon_\omega}^2}{\sigma_\omega^2} + \frac{\sigma_{\varepsilon_\eta}^2}{\sigma_\eta^2} \rho^2 + 1 - \rho^2}{\left(1 + \frac{\sigma_{\varepsilon_\omega}^2}{\sigma_\omega^2}\right) \left(1 + \frac{\sigma_{\varepsilon_\eta}^2}{\sigma_\eta^2}\right) - \rho^2}, \quad (1)$$

$$Cov(\omega, \theta) = \sigma_{\omega\eta} \frac{\frac{\sigma_{\varepsilon_\eta}^2}{\sigma_\eta^2} + \frac{\sigma_{\varepsilon_\omega}^2}{\sigma_\omega^2} + 1 - \rho^2}{\left(1 + \frac{\sigma_{\varepsilon_\omega}^2}{\sigma_\omega^2}\right) \left(1 + \frac{\sigma_{\varepsilon_\eta}^2}{\sigma_\eta^2}\right) - \rho^2}, \quad (2)$$

and

$$Cov(\eta, \theta) = Var(\theta). \quad (3)$$

Equation (3) has a natural interpretation that we explain with the help of the sender's ideal policy, (4), below. Equations (1) and (2) depend crucially on the ratios $\frac{\sigma_{\varepsilon_\omega}^2}{\sigma_\omega^2}$ and $\frac{\sigma_{\varepsilon_\eta}^2}{\sigma_\eta^2}$. E.g., $Cov(\omega, \theta)$ reaches its maximum value, $\sigma_{\omega\eta}$, if at least one of the signals is perfectly precise. DS characterize the entire set of feasible joint distributions and show in particular that, for the natural case of a symmetric prior with $Var(\omega) = Var(\eta) \equiv \sigma^2$, a joint distribution is feasible if and only if $Cov(\omega, \theta) \leq \sigma_{\omega\eta}$ and for any $Cov(\omega, \theta) = C$, $Var(\theta) \in \left[\rho C, \frac{1}{\rho} C\right]$. From now on we stick to this case.⁷ For future reference, we denote the set of joint distributions that can be generated through Bayesian updating Γ and depict the set in Figure 1.

The relative usefulness of the signal - in the sense of its relative covariation with the underlying state ω or η - determines the conflicts that the sender and the receiver face when

⁷This is not essential; we could allow for asymmetric cases and impose the weaker condition $\min\{\sigma_\omega^2, \sigma_\eta^2\} \geq \sigma_{\omega\eta}$.

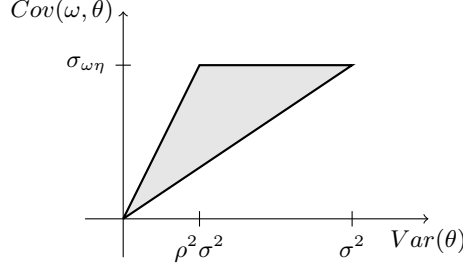


Figure 1: The set Γ of feasible second moments of the joint distribution of (ω, η, θ) .

they communicate with each other. To see this, observe that the sender's ideal policy as a function of θ is simply

$$y^s(\theta) = \frac{Cov(\eta, \theta)}{Var(\theta)} \cdot \theta = \theta, \quad (4)$$

where $\mathbb{E}[\eta|\theta] = \frac{Cov(\eta, \theta)}{Var(\theta)} \cdot \theta$ follows from the linearity of conditional means and $\frac{Cov(\eta, \theta)}{Var(\theta)} = 1$ from the fact that θ is a sufficient statistic from the sender's point of view for the underlying signals. Clearly, by construction of θ , the sender does not revise his posterior if shown his conditional expected mean again. In contrast, if the sender were kind enough to communicate θ truthfully, then the receiver's ideal policy would be

$$y^r(\theta) = \frac{Cov(\omega, \theta)}{Var(\theta)} \cdot \theta. \quad (5)$$

Compared to the sender's ideal policy, $y^s(\theta) = \theta$, the receiver is relatively more (less) conservative with respect to using the aggregated signal θ than the sender is if $Cov(\omega, \theta) < (>) Var(\theta) = Cov(\eta, \theta)$. For future reference, define

$$c \equiv \frac{Cov(\omega, \theta)}{Var(\theta)}.$$

The bias $(1 - c) \cdot \theta$ captures the sender's incentives to misrepresent the information he has. If $c < 1$, then the receiver is relatively more conservative in her use of the aggregated signal θ and the sender has incentives to exaggerate; if $c > 1$, then the receiver is relatively more enthusiastic in her use of signal θ and the sender has incentives to downplay.

It is useful to illustrate the situation for some extreme cases. Suppose the sender observes η without noise. Clearly, the signal contains all the information he is interested in and θ is identically equal to η so that $Var(\theta) = \sigma^2$. Moreover, $Cov(\omega, \theta) = \sigma_{\omega\eta}$, as is easy to see from (2). So, this signal structure corresponds to the top right corner of the set Γ in Figure 1 with $c = \frac{\sigma_{\omega\eta}}{\sigma^2} = \rho < 1$ corresponding to the slope of a ray from the origin through the top right vertex. Since the information is primarily useful to the sender, the receiver discounts the signal θ and reacts more conservatively to the signal θ than the sender would want her to respond. The situation is depicted in the left panel of figure 2 below. Similarly, suppose the sender observes ω without noise and the signal s_η contains an infinite amount of noise. It is easy to see that $\theta = \mathbb{E}[\eta|\omega] = \rho \cdot \omega$ in this case, so that $Var(\theta) = \rho^2 \cdot \sigma^2$. This information structure corresponds to the top left corner in figure 1 with $c = \frac{\sigma_{\omega\eta}}{\rho^2 \sigma^2} = \frac{1}{\rho} > 1$. In this case, the information is more useful to the receiver and hence the receiver overreacts to changes in θ from the sender's perspective, as shown in the right panel of figure 2.

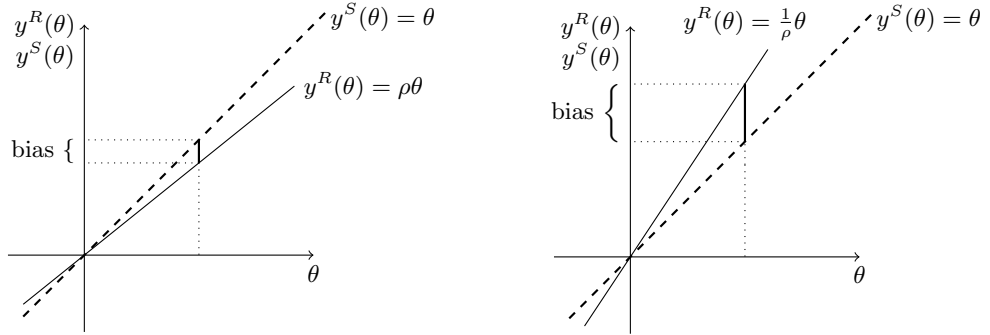


Figure 2: Conflicts as a function of the underlying information.

More generally, any feasible pair of moments $V \equiv Var(\theta)$ and $C \equiv Cov(\omega, \theta)$ in the set Γ in Figure 1, gives rise to a regression coefficient for the receiver of $c = \frac{C}{V}$, the slope of a ray from the origin through V, C . For any $c \in \left[\rho, \frac{1}{\rho}\right]$, there is a continuum of pairs V, C that generate c . By construction, the sender's ideal choice function is the identity function. Thus, any bias smaller or equal in absolute terms to the ones depicted in Figure 2 can arise from some choice of the underlying amounts of noise in the sender's signals.

2.5 Information: basic mechanics and trade-offs

The preceding discussion describes the microfoundations of a stylized yet rigorously justified model that captures the basic properties of Blackwell better information and features conflicts as a function of the information. In particular, the marginal distribution of θ is Laplace with density

$$f(\theta) = \frac{1}{2}\lambda \exp(-\lambda |\theta|) \quad (6)$$

with mean zero and variance $V = \frac{2}{\lambda^2}$, i.e., the scale parameter λ is pinned down by V . Figure 3 illustrates the distribution for different values of λ .

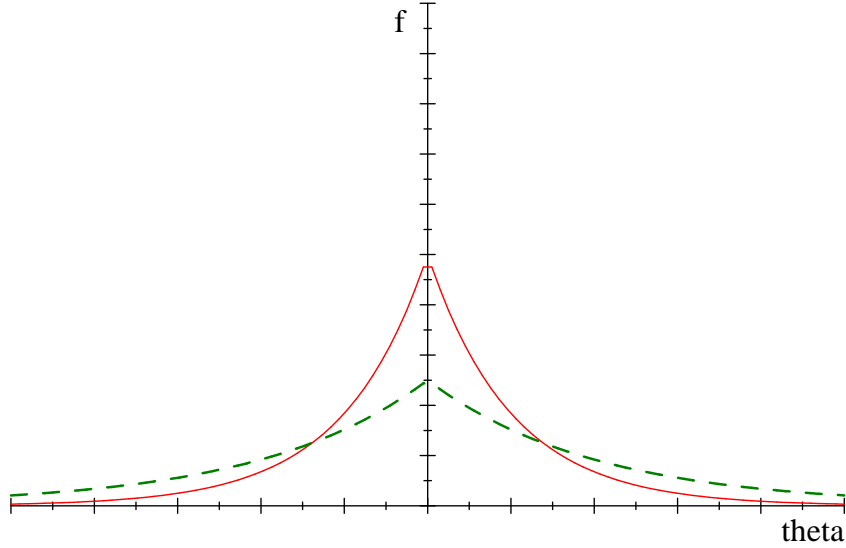


Figure 3: The density of the Laplace for $\lambda = 1$ (solid, red line) and $\lambda = \frac{1}{2}$ (dashed, green line).

Since the sender's and the receiver's interests are not perfectly correlated, information creates or resolves conflicts between sender and receiver depending on the ratio $c = \frac{C}{V}$.

The sender chooses the moments V and C with a view to their influence on the intrinsic value of information and their impact on the conflicts they generate when communicating about θ . The intrinsic value of the aggregated signal θ corresponds to the resulting expected utilities if this signal were publicly observable. However, since it isn't, some of the intrinsic

value is typically lost in conversation. The sender thus chooses the information structure so as to maximize its net value to him. Information structures that are both intrinsically more valuable and give rise to less conflicts are unambiguously preferred. However, a trade-off arises if an information structure is intrinsically more useful than another one but gives rise to more pronounced conflicts. Our model is designed to investigate precisely this trade-off in detail.

While the analysis of delegated decision-making is trivial, the analysis of communication is not. We begin with a characterization of communication equilibria and derive a closed form expression for the value of communication. This analytical result is what enables us to compare institutions. The necessity to have such a closed form representation for values is also what forces us to assume a more specific informational environment than DS.

3 Communication equilibria

We now investigate equilibria in the communication game when the sender observes signals of a given quality. At most θ is communicated in equilibrium. Indeed, an equilibrium in which θ is transmitted truthfully can be supported for all information structures featuring $c = 1$. The reason is simply that setting $c = 1$ eliminates all conflicts with respect to using the *aggregated* signal θ , even though the model still features conflicts with respect to the underlying signals almost surely. We explain this result, derived in DS, in our Theorem 2 below. The cases $c < 1$ and $c > 1$ are fundamentally different. For $c \neq 1$, any equilibrium is essentially equivalent to a partition equilibrium where sender types within intervals induce the same receiver response. We adopt the view taken in the literature, that the sender and the receiver manage to coordinate on the equilibrium that gives them the highest expected utility. Partitional equilibria are completely characterized by a sequence of marginal types, a_i , who are indifferent between pooling with types slightly below and with types slightly above them. We focus on symmetric equilibria in the main text and prove in the Appendix that this is without loss of generality for our results.⁸

⁸More precisely, for the case $c \leq 1$, we show that the highest feasible payoff is attained in a symmetric equilibrium. In addition, symmetric equilibria are the only ones that exist in this case. For the case $c > 1$, we prove the essential result also allowing for asymmetric equilibria.

Symmetric equilibria come in two classes. Class I has zero as a threshold, $a_0^n = 0$, and in addition $n \geq 0$ thresholds a_1^n, \dots, a_n^n above the prior mean. By symmetry, types $-a_n^n, \dots, -a_1^n$ are the threshold types below the prior mean. Such an equilibrium induces $2(n+1)$ actions; superscript n captures the dependence of the equilibrium threshold types on the number of induced actions. Class II has zero as an action taken by the receiver instead of a threshold. Such an equilibrium induces $2n+1$ actions. For $n \geq 1$, let

$$\mu_i^n \equiv \mathbb{E}[\theta | \theta \in [a_{i-1}^n, a_i^n]] \text{ for } i = 1, \dots, n \quad (7)$$

and $\mu_{n+1}^n \equiv \mathbb{E}[\theta | \theta \geq a_n^n]$. By convention, we take all intervals as closed from below and open from above.⁹ Clearly, given quadratic loss functions, the receiver's best reply if sender types in the interval $[a_{i-1}^n, a_i^n)$ pool is to choose

$$y(a_{i-1}^n, a_i^n) = c \cdot \mu_i^n \text{ for } i = 1, \dots, n$$

and $y(a_n^n, \infty) = c \cdot \mu_{n+1}^n$ if sender types with $\theta \geq a_n^n$ pool. This follows from the law of iterated expectations and the linearity of the conditional expectation function. Hence, a class I equilibrium that induces $2(n+1)$ actions by the receiver is completely characterized by the indifference conditions of the marginal types a_1^n, \dots, a_n^n :

$$a_i^n - c \cdot \mu_i^n = c \cdot \mu_{i+1}^n - a_i^n, \text{ for } i = 1, \dots, n. \quad (8)$$

By symmetry, this system of equations also characterizes the marginal types below the prior mean. A class II equilibrium is characterized by the same set of indifference conditions for $i = 2, \dots, n$. In that case, we let $\mu_1^n \equiv \mathbb{E}[\theta | \theta \in [a_{-1}^n, a_1^n)]$. In what follows, we are primarily concerned with class I equilibria, so no confusion will arise.

For the Laplace distribution, it is easy to show that, for $i = 1, \dots, n$ and $a_i^n > a_{i-1}^n \geq 0$

$$\mu_i^n = \frac{1}{\lambda} + a_i^n - g(a_i^n - a_{i-1}^n), \quad (9)$$

where $g(q) \equiv \frac{q}{1 - \exp(-\lambda q)}$. Moreover,

$$\mu_{n+1}^n = \frac{1}{\lambda} + a_n^n. \quad (10)$$

⁹With the obvious exception of the interval $(-\infty, -a_n^n)$.

We can now write the equilibrium conditions more explicitly. Substituting (9) and (10) into (8) and rearranging in a way suitable to the analysis, a class I equilibrium is a set of marginal types that satisfy the conditions

$$cg(a_i^n - a_{i-1}^n) = 2\frac{c}{\lambda} + c(a_{i+1}^n - a_i^n) - cg(a_{i+1}^n - a_i^n) + 2(c-1)a_i^n \quad (11)$$

for $i = 1, \dots, n-1$ and

$$cg(a_n^n - a_{n-1}^n) = 2\frac{c}{\lambda} + 2(c-1)a_n^n, \quad (12)$$

where $a_0^n = 0$. A class II equilibrium satisfies

$$a_1^n = \frac{c}{\lambda} + c(a_2^n - a_1^n) - cg(a_2^n - a_1^n) - (1-c)a_1^n, \quad (13)$$

and in addition (11) for $i = 2, \dots, n-1$ and (12).

It proves convenient to first understand properties that equilibria necessarily have if they exist.

Lemma 1 *Suppose class I and II equilibria inducing $2(n+1)$ and $2n+1$ receiver actions exist. Class I equilibria feature $a_{i+1}^n - a_i^n > a_i^n - a_{i-1}^n$ for all $i = 1, \dots, n-1$; class II equilibria always share this feature for $i = 2, \dots, n-1$.*

The result is depicted graphically in Figure 4.

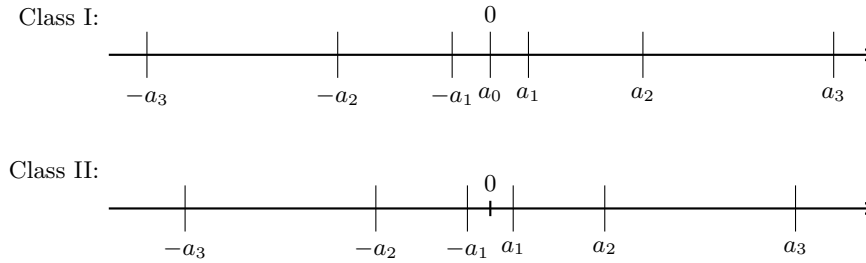


Figure 4: Intervals get longer the farther away from the agreement point they are.

The intuition is that announcements where $|\theta|$ is relatively small are relatively more credible. The reason is straightforward. At $\theta = 0$, the sender's and the receiver's ideal policies coincide. The smaller is $|\theta|$, the smaller is the sender's bias in absolute terms, $(1 - c)|\theta|$. The farther away from the agreement point, $\theta = 0$, the larger the bias, and hence the coarser the information that the sender transmits in equilibrium. Hence, intervals get longer the farther out in the type space they are.

The ideas to prove the lemma are as follows. The case $c = 1$ is straightforward. By symmetry, it suffices to characterize equilibrium thresholds on \mathbb{R}^+ . The density is decreasing for $\theta > 0$. Therefore, the truncated means are located closer to the lower bound of each interval. Consider two adjacent intervals with some dividing point between them. If the two intervals had the same length, then the distance from the dividing point to the truncated mean below would exceed the distance from the truncated mean above the dividing point to the dividing point, clearly violating the indifference condition. To restore the indifference, the first interval has to be shortened, the second one lengthened. For $c < 1$, this effect is reinforced. For $c > 1$, the proof is constructive, working backwards through the indifference conditions from the last one of type a_n^n backwards to the origin.

We now address existence and further characterization of equilibria. To streamline the exposition, we focus on class I equilibria in what follows in the main text. The extensions to class II equilibria and asymmetric equilibria, where necessary, can be found in the Appendix. The main take away message of this analysis is Equation (17) in Proposition 3, which, when combined with Propositions 1 and 2, allows us to prove our theorems. The reader uninterested in the technical details can jump directly to Section 5 at a first go.

3.1 Conservatism, enthusiasm, and limits to communication

Equilibria for $n \leq 1$ are straightforward; we leave this to the reader. For $n \geq 2$, existence of equilibria is a non-trivial question, because equilibria cannot be characterized in closed form, due to the non-linearity of equation (9) in the thresholds a_{i-1} and a_i . We construct class I equilibria for $n \geq 2$, defined as solutions of conditions (11) and (12), as follows. Take $a_1 = x$ as an arbitrary initial step length and use condition (11) as an algorithm that determines $a_2(x)$; repeat this procedure successively to determine “the next” threshold type as a function of the preceding thresholds. Provided such solutions exist up to and including

$a_n(x)$, we obtain a sequence of thresholds, $a_2(x), \dots, a_n(x)$ that satisfy condition (11) for a given initial condition $a_1 = x$. This is called a *forward solution*. We delineate exact conditions on x such that uniquely defined forward solutions exist up to and including $a_n(x)$. Next, note that condition (12) depends on a_n and a_{n-1} as well; we call this condition a *closure condition*. An equilibrium sequence of thresholds is a fixed point with respect to x : the sequence of thresholds satisfies the closure condition for the values of $a_n(x)$ and $a_{n-1}(x)$ that are generated by the forward solution.

The argument is conceptually straightforward but quite involved. We split the discussion into two cases. We begin with the case where the receiver is conservative and show that for any n , unique class I and II equilibria exist. In contrast to that, for the case of an overly enthusiastic receiver, there is necessarily a bound on the number of induced receiver actions.

Lemma 2 *Let $c \leq 1$. For all n , there exists a class I equilibrium that induces $2(n+1)$ actions and a class II equilibrium that induces $2n+1$ actions. Moreover, for each n there is only one equilibrium in each class.*

We illustrate the procedure for the case $n = 2$. The forward solution for $a_2(x)$ is the value of a_2 that solves

$$cg(x) = 2\frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) + 2(c-1)x.$$

A forward solution exists iff x is low enough. For any such x , the solution $a_2(x)$ is unique and satisfies $\lim_{x \rightarrow 0} a_2(x) = 0$ and $\frac{da_2}{dx} > 1$. Substituting the solution into condition (12), we obtain

$$cg(a_2(x) - x) = 2\frac{c}{\lambda} + 2(c-1)a_2(x).$$

Since $a_2(x) - x$ is increasing in x , the left side of the equation is increasing in x . As $(c-1) \leq 0$, the right side is nonincreasing in x . Hence, there is a unique fixed point, i.e. value of \tilde{x} , that satisfies the equality. The thresholds $a_1^2 = \tilde{x}$ and $a_2^2(\tilde{x})$ completely characterize the unique class I equilibrium inducing 6 actions. The proof generalizes these insights to the case of arbitrary n .

There is no upper bound on the number of induced actions. If the state is $\theta = 0$, then the sender's and the receiver's ideal actions coincide. This explains why communication equilibria exist that are arbitrarily fine around the agreement point. The second statement

in the lemma is that for each given n and for each given class of equilibrium, the induced actions and the equilibrium partitioning of the state space is uniquely determined. The reason for uniqueness is that the exponential distribution is loglinear and hence weakly logconcave (see Szalay (2012)).

It is instructive to investigate how equilibria within a given class differ when they have more thresholds. So, consider a class I equilibrium with n thresholds in the positive orthant, $\{a_1^n, \dots, a_n^n\}$, and compare it to the equilibrium with $n+1$ thresholds in the positive orthant, $\{a_1^{n+1}, \dots, a_{n+1}^{n+1}\}$.

Lemma 3 *Let $c \leq 1$. Equilibrium thresholds are nested. Formally, $a_1^{n+1} < a_1^n < a_2^{n+1} < \dots < a_n^{n+1} < a_n^n < a_{n+1}^{n+1}$ for all n .*

The result is intuitive; we depict it graphically in Figure 5.

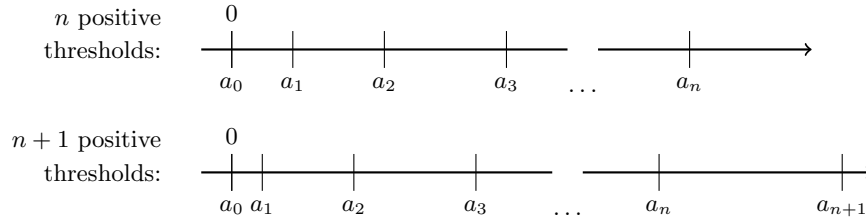


Figure 5: The effect of increasing the number of induced receiver responses on equilibrium thresholds.

The stated order of equilibrium thresholds follows from monotonicity in several ways. Thresholds within an equilibrium are monotonic in the initial length a_1 . Moreover, a threshold is monotonic in the level of the previous threshold and the length of the previous interval. Any violation of the stated order implies by monotonicity a violation of the equilibrium fixed-point condition.

Equilibria inducing the highest number of distinct receiver actions are natural to study, because the utility of sender and receiver in these equilibria is as high as possible. Since

this property makes infinite equilibria focal, we need to understand the properties of such equilibria.

Proposition 1 *Suppose that $c \leq 1$ and consider the limit as $n \rightarrow \infty$ of the finite class I and class II equilibria. The limits correspond to an infinite equilibrium of the communication game. Moreover, in any equilibrium, the equilibrium threshold a_1^n satisfies $\lim_{n \rightarrow \infty} a_1^n = 0$.*

Figure 6 illustrates the essential result.

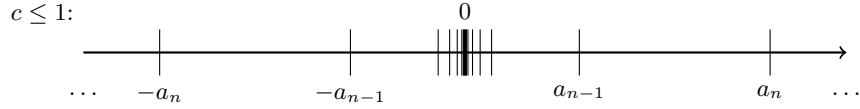


Figure 6: Intervals around the agreement point $\theta = 0$ get arbitrarily short as $n \rightarrow \infty$.

In the limit as more and more distinct receiver actions are induced, the length of the interval(s) that are closest to the agreement point, $\theta = 0$, must go to zero. The reason is that the length of intervals is the higher, the farther these intervals are away from the agreement point; formally, $a_{i+1}^n - a_i^n > a_i^n - a_{i-1}^n$. If the length of the interval that is closest to the agreement point would converge to some positive length as the number of induced choices goes out of bounds, then the last type who is indifferent between inducing two actions, a_n^n , would be arbitrarily far away from the agreement point. However, this makes it impossible to make this type indifferent between two actions. Hence, to make an infinite equilibrium possible in the first place, the length of the first interval must shrink to zero. Note that the argument refers to any infinite equilibrium, not only the limit equilibrium that corresponds to the limit of the finite equilibrium when the number of induced actions goes to infinity.

For the case $c > 1$, there is an upper bound on the number of induced actions:

Proposition 2 *Suppose that $c > 1$. Then, a_1^n is bounded away from zero in any equilibrium. Moreover, the number of induced actions is finite.*

Key to understanding the result is the forward equation for a given initial interval length x . Depending on the level of c , either one of two cases arises. For relatively large bias ($c > \frac{4}{3}$), the forward equation has no solution for x small, implying directly that the length of the first interval is bounded away from zero in any equilibrium that exists. For the case of a smaller bias, the forward equation does have a solution, but the solution ceases to satisfy the increasing interval property if the initial interval length gets small. However, since any equilibrium needs to have this property, no equilibrium with a short initial interval length can exist. In both cases, the number of induced actions needs to be finite. For our purposes, it suffices to take away that the equilibrium threshold a_1^n is necessarily bounded away from zero in any equilibrium.

4 The value of communication

In a class I equilibrium inducing $2(n+1)$ distinct receiver responses, the receiver's conditional expectation of θ , conditional on the sender's message, is a random variable μ that is supported on $\{\mu_{-(n+1)}^n, \mu_{-n}^n, \dots, \mu_{-1}^n, \mu_1^n, \dots, \mu_n^n, \mu_{n+1}^n\}$ and the receiver's optimal policy is to choose the action $y(\mu) = c \cdot \mu$. The marginal distribution of the random variable μ is derived from the marginal distribution of θ . The receiver's expected utility in such an equilibrium is

$$\mathbb{E}u^r(c\mu, \omega) = -\mathbb{E}(c\mu - \omega)^2 = c^2\mathbb{E}(\mu)^2 - \sigma^2. \quad (14)$$

The essential steps to prove this equality are the law of iterated expectations, $\mathbb{E}[\mu\omega] = \mathbb{E}_\theta\mathbb{E}_\omega[\mu\omega|\theta]$, and the linearity of conditional means, $\mathbb{E}_\theta\mathbb{E}_\omega[\mu\omega|\theta] = \mathbb{E}_\theta[\mu c\theta]$. Likewise, for the sender we have

$$\mathbb{E}u^s(c\mu, \eta) = -\mathbb{E}(c\mu - \eta)^2 = c(2-c)\mathbb{E}(\mu)^2 - \sigma^2 \quad (15)$$

where we use $\mathbb{E}_\theta\mathbb{E}_\eta[\mu\eta|\theta] = \mathbb{E}(\mu)^2$, because θ is by construction the conditional expectation of the sender.

To relate these expressions to the familiar “residual uncertainty” concepts, let $\Theta_i \equiv [a_{i-1}, a_i)$ for $i \in 1, \dots, n+1$ denote a generic partition element in \mathbb{R}^+ and let $\Theta_{n+1} \equiv [a_n, \infty)$ by convention. The probability mass over these partition elements conditional on $\theta \geq 0$ is

denoted by p_i^n for $i = 1, \dots, n+1$.¹⁰ Then, we can write $\mathbb{E}(\mu)^2 = \sum_{i=1}^{n+1} p_i^n (\mu_i^n)^2$ by symmetry of the distribution. Moreover, by a standard variance decomposition, we have

$$\sum_{i=1}^{n+1} p_i^n (\mu_i^n)^2 = \text{Var}(\theta) - \sum_{i=1}^{n+1} p_i^n [\text{Var}(\theta | \theta \in \Theta_i)]. \quad (16)$$

Using this decomposition, we can understand the receiver's and the sender's expected utilities as an intrinsic value of information net of a loss due to strategic communication. The intrinsic value of information corresponds to a situation where the sender is kind enough to communicate θ truthfully no matter what, or equivalently, θ is publicly observable. In that case, the receiver would identify μ with θ , so that $\mathbb{E}(\mu)^2 = \text{Var}(\theta)$. However, for $c \neq 1$, this is not an equilibrium and the sender behaves strategically. The resulting losses due to strategic communication are precisely proportional to $\sum_{i=1}^{n+1} p_i^n [\text{Var}(\theta | \theta \in \Theta_i)]$.

The equilibrium variability of choices has a convenient representation:

Proposition 3 *The equilibrium variability of the receiver's posterior mean in a class I equilibrium is given by*

$$\sum_{i=1}^{n+1} p_i^n (\mu_i^n)^2 = \frac{1}{2-c} \text{Var}(\theta) - \frac{c}{2-c} (\mu_1^n)^2. \quad (17)$$

The proposition provides an extremely powerful result, stating that the equilibrium variability of the receiver's posterior depends only the posterior mean taken over the interval $[0, a_1^n]$. Eső and Szalay (2015) prove the result for $c = 1$, this paper generalizes the result to the case $c \neq 1$. Clearly, it is the central result that enables us to address our question at all.

To understand the result, consider the case $n = 1$ with four induced choices. The equilibrium probability distribution on \mathbb{R}^+ can be computed from the indifference condition of the marginal type a_1^1 , $a_1^1 - c\mu_1^1 = c\mu_2^1 - a_1^1$. Substituting for $c\mu_1^1 = \frac{c}{\lambda} + ca_1^1 - c\frac{a_1^1}{1-\exp(-\lambda a_1^1)}$ from (9) and $c\mu_2^1 = \frac{c}{\lambda} + ca_1^1$ from (10) and solving for p_1^1 , we find that $p_1^1 = \frac{ca_1^1}{\frac{c}{\lambda} + ca_1^1 - c\mu_1^1}$. Using the equilibrium probability distribution, we can solve for the centered second moment of $c \cdot \mu$,

¹⁰Formally, we define $p_i^n \equiv \int_{a_{i-1}^n}^{a_i^n} \lambda \exp(-\lambda \theta) d\theta$ for $i = 1, \dots, n$ and $p_{n+1}^n \equiv \int_{a_n^n}^{\infty} \lambda \exp(-\lambda \theta) d\theta$.

the receiver's choices. We can write

$$\sum_{i=1}^2 p_i^1 \left(c\mu_i^1 - \frac{c}{\lambda} \right)^2 = ca_1^1 \left(\frac{c}{\lambda} - c\mu_1^1 \right) = \frac{c}{2-c} \left(c\mu_1^1 + \frac{c}{\lambda} \right) \left(\frac{c}{\lambda} - c\mu_1^1 \right),$$

where the first equality uses the equilibrium probability distribution and the second one the fact that $c\mu_2^1 - \frac{c}{\lambda} = ca_1^1$, which together with the indifference condition of the marginal type implies that $ca_1^1 = \frac{c}{2-c} \left(c\mu_1^1 + \frac{c}{\lambda} \right)$. Since $\frac{1}{\lambda}$ is the mean of the exponential distribution, we have $\sum_{i=1}^2 p_i^1 \mu_i^1 = \frac{1}{\lambda}$ by consistency of the conditional distributions with the marginal distribution of θ . Decentering again and cancelling c^2 , we have shown that

$$\sum_{i=1}^2 p_i^1 (\mu_i^1)^2 = \frac{1}{2-c} \frac{2}{\lambda^2} - \frac{c}{2-c} (\mu_1^1)^2 = \frac{1}{2-c} V - \frac{c}{2-c} (\mu_1^1)^2, \quad (18)$$

where the second equality results from noting that $\frac{2}{\lambda^2} = V$.

Equation (18) is precisely the expression in the proposition for the special case of $n = 1$. Surprisingly, the right-hand side of equation (18) depends on the equilibrium only through μ_1^1 , the mean over the partition element that is closest to the agreement point $\theta = 0$. The reason is the close connection between truncated means and the probability mass over any interval implied by the exponential distribution. The indifference condition allows to eliminate μ_2^1 from the picture.

If equation (18) is surprising, Proposition 3 is stunning, at least to us. Regardless of how many distinct receiver actions are induced in a class I equilibrium, the equilibrium variability of the receiver's posterior mean always depends on the communication equilibrium exclusively through μ_1^n , the mean over the partition element that is closest to the agreement point, $\theta = 0$. The formal proof of the proposition is based on an induction argument.

The proposition is extremely powerful. It allows us to quantify the usefulness of information structures in strategic communication.

5 Information acquisition

When choosing the information structure, the sender takes its effects on equilibrium communication into account. Formally, the sender's problem is to

$$\begin{aligned} \max_{C,V} c(2-c) \left(\frac{1}{2-c} V - \frac{c}{2-c} (\mu_1^n)^2 \right) - \sigma^2 \\ \text{s.t. } c = \frac{C}{V} \text{ and } C, V \in \Gamma. \end{aligned} \quad (19)$$

where the sender's objective follows from substituting for $\mathbb{E}(\mu)^2$ from (17) into (15). In face of Propositions 1 and 2, the solution to problem (19) is obvious:

Theorem 1 *The set of optimal information structures from the sender's perspective is given by $C = \sigma_{\omega\eta}$ and $V \geq C$.*

The proof of the theorem is a straightforward combination of the preceding propositions. For any C, V such that $c = \frac{C}{V} \leq 1$, there exists a class I equilibrium with arbitrarily many induced actions. In the limit where the number of induced actions goes out of bounds, we have $\lim_{n \rightarrow \infty} \mu_1^n = 0$ and the equilibrium variability of the receiver's conditional mean converges to $\frac{1}{2-c}V$. Hence, the sender's expected utility is equal to

$$\mathbb{E}u^s(c\mu, \eta) = c(2-c)V \cdot \frac{1}{2-c}.$$

The term $c(2-c)V$ corresponds to the intrinsic value of the signal θ to the sender if it is publicly observed and the receiver follows the policy $y^r(\theta) = c \cdot \theta$. Since θ isn't publicly observed, some of its value is lost in strategic communication and only the fraction $\frac{1}{2-c}$ of the intrinsic value materializes. Within the set of joint distributions that satisfy $\frac{C}{V} \leq 1$, the intrinsic value of the signal θ is strictly increasing in V and C . Hence, the highest intrinsic value is obtained if $C = \sigma_{\omega\eta}$ and $V = \frac{1}{\rho}\sigma_{\omega\eta}$. Note that this distribution minimizes c . On the other hand, the fraction of the intrinsic value that survives in communication, $\frac{1}{2-c}$, is an increasing function of c . Interestingly, the strategic effect completely undoes any impact that V has on the sender's expected utility. In the limit equilibrium where n goes out of bounds, the sender's expected utility thus becomes equal to C , and depends on V only in that V

cannot exceed C . Hence, within the set of distributions that satisfy $\frac{C}{V} \leq 1$ any distribution with $C = \sigma_{\omega\eta}$, the highest feasible covariance, is optimal for the sender, generating an expected utility of $\sigma_{\omega\eta} - \sigma^2$ to the sender.

Consider now all other information structures featuring $\frac{C}{V} > 1$. By Proposition 2, there is no equilibrium in which the interval around the agreement point $\theta = 0$ gets small, so $\mu_1^n > 0$. Therefore, for all $\frac{C}{V} > 1$,

$$\mathbb{E}u^s(c\mu, \eta) = c(2 - c) \left(\frac{1}{2 - c}V - \frac{c}{2 - c}(\mu_1^n)^2 \right) - \sigma^2 < \sigma_{\omega\eta} - \sigma^2.$$

Clearly, the expression for the equilibrium variability of the receiver's conditional means can be extended in straightforward fashion to cover all equilibria, not just class I equilibria. In any such equilibrium, the sender's expected utility is less than $\sigma_{\omega\eta} - \sigma^2$, because there necessarily is a loss due to biased communication. Hence, we have shown that the sender reaches the highest payoff by choosing any information structure with highest feasible covariance and $\text{Var}(\theta) \geq \sigma_{\omega\eta}$.

Consider now the receiver's payoff as a function of the information structure that the sender chooses. For $C = \text{Cov}(\omega, \eta)$ and any $V \geq C$, the receiver's payoff in the limit class I equilibrium where $n \rightarrow \infty$ is

$$\mathbb{E}u^r(c\mu, \omega) = c^2 \frac{1}{2 - c}V - \sigma^2 = \frac{C^2}{2V - C} - \sigma^2,$$

a decreasing function of V . Clearly, the receiver suffers if the sender chooses an information structure with a higher V ; at the same time, the sender derives no benefit from such behaviour. The following theorem is now obvious:

Theorem 2 *The set of sender optimal information structures contains the uniquely Pareto efficient element $\text{Var}(\theta)^* = \text{Cov}(\omega, \theta)^* = \sigma_{\omega\eta}$. The sender achieves the highest feasible payoff by choosing this information structure and communicating θ truthfully to the receiver, who rubberstamps the sender's proposal.*

For convenience, the result is depicted graphically in figure 7.

If the sender chooses the receiver optimal information structure among those that are privately optimal for himself, then we have $c = \frac{C}{V} = 1$, and the bias with respect to communicating θ is eliminated. Hence, it is an equilibrium for the sender to follow the message strategy

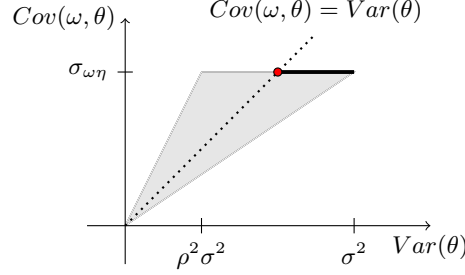


Figure 7: The solid line represents the sender-optimal information structures. The red dot represents the receiver-optimal information structure within the set of sender-optimal information structures.

$m(\theta) = \theta$ for all θ , and for the receiver to follow the action strategy $y(m) = \frac{Cov(\omega, \theta)}{Var(\theta)} \cdot m = m$ for all m , because the receiver correctly identifies m with θ in her belief. In other words, the situation corresponds to what DS have termed a “smooth communication equilibrium”, where smoothness refers to the differentiability of the sender’s and the receiver’s strategy.

6 Delegation versus communication

We are now ready to address the question we have set out to answer: if a receiver is forced to delegate information acquisition to a sender, should she also delegate decision-making to this sender? The answer is unambiguously negative.

If the receiver retains the right to make choices, then the sender is happy to look into issues also of relevance to the receiver. The receiver’s expected payoff in the smooth communication equilibrium is

$$\mathbb{E}u^r(\theta, \omega) = \sigma_{\omega\eta} - \sigma^2. \quad (20)$$

If the sender has the right to choose the action directly, then he will follow the action policy $y^s(\theta) = \theta$ for all θ , resulting in expected utility for the sender of

$$\mathbb{E}u^s(\theta, \eta) = -\mathbb{E}(\theta - \eta)^2 = V - \sigma^2,$$

where we used the fact that $Cov(\theta, \eta) = V$ by construction of θ . Clearly, the optimal information structure from the sender's perspective, if he is authorized to choose the action y , is $\hat{C} = \sigma_{\omega\eta}$ and $\hat{V} = \frac{1}{\rho}\sigma_{\omega\eta}$, because this information structure maximizes V within the set Γ . The receiver's expected utility is

$$\mathbb{E}u^r(\theta, \omega) = -\hat{V} + 2\hat{C} - \sigma^2 = \left(2 - \frac{1}{\rho}\right)\sigma_{\omega\eta} - \sigma^2. \quad (21)$$

We can now state our main result:

Theorem 3 *Suppose the sender selects a privately optimal information structure; in case there are several optimal ones, he picks the receiver's preferred information structure among them. Then, the receiver strictly prefers delegating information acquisition only and communicating with the sender over delegating both information acquisition and decision-making to the sender.*

The formal proof of the theorem consists again simply of pulling insights together. In particular, direct comparison of (20) and (21) reveals that communication is the preferred mode of decision-making, because $2 - \frac{1}{\rho} < 1$ for any $\rho \in (0, 1)$. Note also that the receiver always benefits from communicating with the sender, while the gain from delegation is only positive for $\rho > \frac{1}{2}$, that is, if interests are relatively well aligned.

The result stands in stark contrast to what is known for the case of exogenously given information structures and biases. Key to understanding the difference between the results is the selection of the pareto efficient information structure. Clearly, selecting the most efficient equilibria is exactly in the tradition of the communication literature following Crawford and Sobel (1982).¹¹ To play devil's advocate - and to reconcile results - suppose that for whatever reason we selected the worst information structure from the receiver's perspective out of the set of sender optimal information structures. This corresponds to the one that is uniquely optimal under delegation, $C = \sigma_{\omega\eta}$ and $V = \frac{1}{\rho}\sigma_{\omega\eta}$. Obviously, the comparison between delegation and communication is now exactly as if the information structure were exogenously given, simply because the selection criterion picks the same information structure under both institutions.

¹¹See also Chen et al. (2008) for a more recent result in this tradition.

Theorem 4 *Suppose the sender picks a privately optimal information structure; in case there are several optimal ones, he selects the least preferred one from the receiver's perspective. Then, communication is strictly preferred to delegation for $\rho \in (0, \frac{2}{3})$ and delegation is strictly preferred for $\rho \in (\frac{2}{3}, 1)$.*

Substituting $C = \sigma_{\omega\eta}$ and $V = \frac{1}{\rho}\sigma_{\omega\eta}$, receiver's expected utility under communication is

$$\mathbb{E}u^r(c\mu, \omega) = \frac{\sigma_{\omega\eta}^2}{\frac{2}{\rho}\sigma_{\omega\eta} - \sigma_{\omega\eta}} - \sigma^2 = \frac{\sigma_{\omega\eta}}{\frac{2}{\rho} - 1} - \sigma^2.$$

It is easy to see why communication is the preferred mode for badly aligned interests. The receiver benefits from communication for all $\rho > 0$, while delegation is beneficial only if $\rho > \frac{1}{2}$. Similarly, the gain from delegation is rather small for values of ρ larger than but close to $\frac{1}{2}$. Hence, communication is still the preferred mode for $\rho < \frac{2}{3}$. As interests get well aligned, in particular for $\rho > \frac{2}{3}$, delegation becomes the preferred institution. While communication entails a loss of information due to strategic communication, delegation does not. On the other hand, delegation results in a choice of action that is not ideal from the receiver's point of view. We take matters to the extreme by selecting the receiver's least preferred information structure. The qualitative findings remain unchanged if we select a less extreme information structure.¹²

In sum, the current model provides strong support for communication as an institution. The reason is that the receiver retains some indirect means of control over the kind of information that is acquired. Delegation provides no such safeguard mechanism and therefore results in maximally selfish information acquisition by the sender.

7 Conclusions and extensions

We have studied a model of delegated expertise, where the expert's interests are positively but imperfectly correlated with a decision-maker's interests. Conflicts between the two parties depend on the nature of information that the expert acquires. When given decision-rights

¹²In particular, for $\rho < \frac{2}{3}$, communication dominates delegation, regardless of which information structure the sender picks. For $\rho \geq \frac{2}{3}$, the comparison depends on which information structure is selected.

as well, the expert neglects the decision-maker's interests when choosing what information to acquire. In contrast, when forced to communicate with the decision-maker, the expert anticipates that selfish information acquisition results in coarse communication. To avoid the losses from such coarse communication, the expert acquires balanced information that is equally useful to both the decision-maker and himself. The resulting outcome unambiguously dominates the delegation outcome, irrespective of the degree of conflicting interests as long as interests are positively correlated.

The model lends itself to further investigation. Interesting avenues for future research include costs of information acquisition, noisy communication, and mediated communication, to name but a few. The current model assumes away any conflicts that are known ex ante or that they have been successfully eliminated. This is exactly the right assumption to make to understand how information creates and resolves conflicts. However, it is clearly also interesting to learn about the impact of ex ante known conflicts and how they interact with endogenous conflicts as they arise here. We are pursuing these questions in ongoing work.

Appendix

Lemma A1 The function $g(q) = \frac{q}{1-\exp(-\lambda q)}$ satisfies $\lim_{q \rightarrow 0} g(q) = \frac{1}{\lambda}$ and has limits $\lim_{q \rightarrow \infty} g(q) = \infty$, and $\lim_{q \rightarrow \infty} (q - g(q)) = 0$. Moreover, the function is increasing and convex, with a slope satisfying $\lim_{q \rightarrow 0} g'(q) = \frac{1}{2}$ and attaining the limit $\lim_{q \rightarrow \infty} g'(q) = 1$.

Proof of Lemma A1. By l'Hôpital's rule $\lim_{q \rightarrow 0} g(q) = \frac{1}{\lambda}$. The limit $\lim_{q \rightarrow \infty} 1 - \exp(-\lambda q) = 1$ implies that $\lim_{q \rightarrow \infty} g(q) = \infty$. Using $q - g(q) = -\frac{q \exp(-\lambda q)}{1 - \exp(-\lambda q)}$ and $\lim_{q \rightarrow \infty} q \exp(-\lambda q) = 0$, we have $\lim_{q \rightarrow \infty} (q - g(q)) = 0$.

The slope of the function is

$$g'(q) = \frac{(1 - (1 + \lambda q) e^{-q\lambda})}{(1 - e^{-q\lambda})^2} \geq 0.$$

The inequality is strict for $q > 0$ since $\lim_{q \rightarrow 0} (1 + \lambda q) e^{-q\lambda} = 1$ and $\frac{\partial}{\partial q} (1 - (1 + \lambda q) e^{-q\lambda}) = \lambda^2 q e^{-q\lambda} > 0$ for $q > 0$. Applying l'Hôpital's rule twice, one finds that $\lim_{q \rightarrow 0} g'(q) = \frac{1}{2}$, and since $\lim_{q \rightarrow \infty} \lambda q e^{-q\lambda} = 0$, we have $\lim_{q \rightarrow \infty} g'(q) = 1$.

Differentiating $g(q)$ twice, we obtain

$$g''(q) = \lambda \frac{e^{-q\lambda}}{(1 - e^{-q\lambda})^3} (2e^{-q\lambda} + q\lambda + q\lambda e^{-q\lambda} - 2).$$

The sign of the second derivative is equal to the sign of the expression in brackets. At $q = 0$, the expression is zero. The change of the expression is given by

$$\frac{\partial}{\partial q} (2e^{-q\lambda} + q\lambda + q\lambda e^{-q\lambda} - 2) = \lambda (1 - (1 + \lambda q) e^{-q\lambda}) \geq 0,$$

by the same argument as given above. Hence, $g(q)$ is convex. ■

Proof of Lemma 1. Consider first class I equilibria for given $n \geq 2$. For $n < 2$, the question is meaningless. Define

$$z_i^n \equiv a_i^n - a_{i-1}^n \text{ for } i = 1, \dots, n,$$

where we skip the superscript n whenever this causes no confusion - in particular, when n is constant.

We first consider the case $c \leq 1$. The typical indifference condition for type a_i is

$$a_i = \frac{c}{2} (\mathbb{E}[\theta | \theta \in [a_i - z_i, a_i]] + \mathbb{E}[\theta | \theta \in [a_i, a_i + z_{i+1}]]).$$

Suppose that $z_i = z_{i+1} = \Delta$. Because the density is decreasing on \mathbb{R}^+ , $\mathbb{E}[\theta | \theta \in [a_i - \Delta, a_i]] < a_i - \frac{\Delta}{2}$ and $\mathbb{E}[\theta | \theta \in [a_i, a_i + \Delta]] < a_i + \frac{\Delta}{2}$. Hence,

$$\frac{1}{2} (\mathbb{E}[\theta | \theta \in [a_i - \Delta, a_i]] + \mathbb{E}[\theta | \theta \in [a_i, a_i + \Delta]]) < \frac{1}{2} 2a_i = a_i.$$

For given a_{i-1} and given a_{i+1} , the expression $a_i - \frac{c}{2} (\mathbb{E}[\theta | \theta \in [a_{i-1}, a_i]] + \mathbb{E}[\theta | \theta \in [a_i, a_{i+1}]])$ is increasing in a_i , because the loglinear density is weakly logconcave (see Szalay (2012)). Hence, a_i needs to decrease relative to the situation where $z_i = z_{i+1} = \Delta$. This implies that $z_{i+1} > z_i$.

Consider now the case $c \in (1, 2)$. For $c \geq 2$, no equilibrium of the considered kind exists. For $n = 2$, the indifference condition of type a_2 and a_1 are, in that order,

$$cg(z_2) = 2\frac{c}{\lambda} + 2(c-1)(z_1 + z_2),$$

and

$$cg(z_1) = 2\frac{c}{\lambda} + c(z_2 - g(z_2)) + 2(c-1)z_1.$$

Substituting the former condition into the latter and simplifying, we have

$$z_2 = \frac{c}{2-c}g(z_1).$$

Since $g(z) > z$ and $\frac{c}{2-c} > 1$, we have $z_2 > z_1$.

For $n \geq 3$, the indifference conditions of types a_n , a_{n-1} , and a_{n-2} , respectively, can be written as

$$cg(z_n) = 2\frac{c}{\lambda} + 2(c-1)\sum_{j=1}^n z_j,$$

$$cg(z_{n-1}) = 2\frac{c}{\lambda} + c(z_n - g(z_n)) + 2(c-1)\sum_{j=1}^{n-1} z_j,$$

and

$$cg(z_{n-2}) = 2\frac{c}{\lambda} + c(z_{n-1} - g(z_{n-1})) + 2(c-1)\sum_{j=1}^{n-2} z_j.$$

Adding $-2\frac{c}{\lambda} - 2(c-1)\sum_{j=1}^n z_j + cg(z_n) = 0$ to the indifference condition of type a_{n-1} , we get

$$cg(z_{n-1}) = (2-c)z_n,$$

and hence

$$z_n = \frac{c}{2-c}g(z_{n-1}).$$

Since $\frac{c}{2-c} > 1$ for $c > 1$ and $g(z) > z$, this implies that $z_n > z_{n-1}$. By Lemma A1, we therefore have $g(z_n) - z_n < g(z_{n-1}) - z_{n-1}$. Hence, we also have

$$\begin{aligned} cg(z_{n-1}) &= 2\frac{c}{\lambda} + c(z_n - g(z_n)) + 2(c-1)\sum_{j=1}^{n-2} z_j + 2(c-1)z_{n-1} \\ &> 2\frac{c}{\lambda} + c(z_{n-1} - g(z_{n-1})) + 2(c-1)\sum_{j=1}^{n-2} z_j = cg(z_{n-2}), \end{aligned}$$

where the first equality is the indifference condition of type a_{n-1} and the second equality the one for type a_{n-2} . Hence, we can conclude that $z_{n-2} < z_{n-1}$.

Likewise, suppose as an inductive hypothesis that $z_i < z_{i+1}$. Consider the indifference conditions of types a_i and a_{i-1} , respectively,

$$cg(z_i) = 2\frac{c}{\lambda} + c(z_{i+1} - g(z_{i+1})) + 2(c-1) \sum_{j=1}^{i-1} z_j + 2(c-1)z_i$$

and

$$cg(z_{i-1}) = 2\frac{c}{\lambda} + c(z_i - g(z_i)) + 2(c-1) \sum_{j=1}^{i-1} z_j.$$

By Lemma A1, the value of the right-hand side of the former equation exceeds the value of the right-hand side of the latter equation, and hence we have shown that $z_{i-1} < z_i$.

Class II equilibria have the same indifference conditions for the marginal types a_i for $i = 2, \dots, n-1$. Hence, the same argument applies.

Note that we do not invoke symmetry of the equilibrium in any way. Therefore, except for notation, the same argument applies also to asymmetric equilibria. ■

Proof of Lemma 2. We prove the result for the case of class I equilibria first. The argument is structured as follows. In a first step, we investigate the forward solution, addressing first properties of solutions and then existence. In a second step, we address existence and uniqueness of a fixed point. In a third step, we give an inductive argument for the existence of equilibria for all n . Finally, the extension to the case of class II equilibria is presented.

1. The forward solution

i. Properties

For $c \leq 1$, it is easy to see from the proof of Lemma 1 that the forward solution satisfies $a_2(x) - x > x$ and $a_{i+1}(x) - a_i(x) > a_i(x) - a_{i-1}(x)$ for $i = 3, \dots, n-1$.

We first show that the forward solution $a_2(x)$ satisfies $\lim_{x \rightarrow 0} (a_2(x) - x) = 0$ and $\frac{da_2}{dx} > 1$, implying that $a_2(x) - x$ is increasing in x . Then we show that the forward solutions $a_i(x) - a_{i-1}(x)$ for $i = 3, \dots, n$ all satisfy $\lim_{x \rightarrow 0} (a_i(x) - a_{i-1}(x)) = 0$ and $\frac{da_{i+1}(x)}{dx} > \frac{da_i(x)}{dx}$, implying that $a_i(x) - a_{i-1}(x)$ is increasing in x .

Consider the equation determining the forward solution for $a_2(x)$, that is condition (11) for $i = 1$, $a_0 = 0$, and $a_1 = x$; formally, $a_2(x)$ is the value of a_2 that solves

$$cg(x) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) + 2(c-1)x.$$

In the limit as $x \rightarrow 0$, we obtain $\lim_{x \rightarrow 0} a_2(x) = 0$ from the fact that $\lim_{q \rightarrow 0} g(q) = \frac{1}{\lambda}$. Totally differentiating, we obtain

$$(cg'(x) + c(1 - g'(a_2(x) - x)) - 2(c - 1))dx - c(1 - g'(a_2(x) - x))da_2 = 0,$$

so that

$$\frac{da_2}{dx} = \frac{(cg'(x) + c(1 - g'(a_2(x) - x)) - 2(c - 1))}{c(1 - g'(a_2(x) - x))} > 0.$$

Moreover, $\frac{da_2}{dx} > 1$ by the fact that $cg'(x) - 2(c - 1) > 0$ for $c \leq 1$. Hence, we have that $\lim_{x \rightarrow 0} (a_2(x) - x) = 0$ and $\frac{d}{dx}(a_2(x) - x) > 0$.

For $i = 2$, consider the forward equation for $a_3(x)$. Formally, $a_3(x)$ is the value of a_3 that solves

$$cg(a_2(x) - x) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_3 - a_2(x)) - cg(a_3 - a_2(x)) + 2(c - 1)a_2(x).$$

Since $\lim_{x \rightarrow 0} a_2(x) = 0$ and $\lim_{x \rightarrow 0} (a_2(x) - x) = 0$, we also have $\lim_{x \rightarrow 0} a_3(x) = 0$ and $\lim_{x \rightarrow 0} (a_3(x) - a_2(x)) = 0$. Totally differentiating, we obtain

$$\frac{da_3(x)}{da_2(x)} = \frac{cg'(a_2(x) - x) \left(\frac{da_2(x)}{dx} - 1 \right) + (c(1 - g'(a_3(x) - a_2(x))) - 2(c - 1)) \frac{da_2(x)}{dx}}{c(1 - g'(a_3(x) - a_2(x))) \frac{da_2(x)}{dx}}.$$

Since $\frac{da_2(x)}{dx} > 1$, we have $\frac{da_3(x)}{da_2(x)} > 0$, and moreover $\frac{da_3(x)}{da_2(x)} > 1$. Finally,

$$\frac{da_3(x)}{dx} = \frac{da_3(x)}{da_2(x)} \frac{da_2(x)}{dx} > \frac{da_2(x)}{dx}.$$

Hence, we have that $\lim_{x \rightarrow 0} (a_3(x) - a_2(x)) = 0$ and $\frac{d}{dx}(a_3(x) - a_2(x)) > 0$.

Suppose as an inductive hypothesis that the forward solutions up to and including $a_i(x)$ have the properties that $\lim_{x \rightarrow 0} (a_i(x) - a_{i-1}(x)) = 0$, $\lim_{x \rightarrow 0} a_i(x) = 0$, and $\frac{da_i(x)}{da_{i-1}(x)} > 1$, so that $a_i(x) - a_{i-1}(x)$ increasing in x . Consider the equation for a_{i+1} with solution $a_{i+1}(x)$,

$$cg(a_i(x) - a_{i-1}(x)) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_{i+1} - a_i(x)) - cg(a_{i+1} - a_i(x)) + 2(c - 1)a_i(x).$$

The inductive assumptions for $a_i(x)$ and $a_{i-1}(x)$ imply that $\lim_{x \rightarrow 0} (a_{i+1}(x) - a_i(x)) = 0$,

so that $\lim_{x \rightarrow 0} a_{i+1}(x) = 0$. Totally differentiating, we obtain

$$\begin{aligned} & \frac{da_{i+1}(x)}{da_i(x)} \\ &= \frac{cg'(a_i(x) - a_{i-1}(x)) \left(\frac{da_i(x)}{da_{i-1}(x)} - 1 \right) + (c(1 - g'(a_{i+1}(x) - a_i(x))) - 2(c-1)) \frac{da_i(x)}{da_{i-1}(x)}}{c(1 - g'(a_{i+1}(x) - a_i(x))) \frac{da_i(x)}{da_{i-1}(x)}}. \end{aligned}$$

The assumption $\frac{da_i(x)}{da_{i-1}(x)} > 1$ implies that $\frac{da_{i+1}(x)}{da_i(x)} > 1$. We can conclude that, $a_{i+1}(x) - a_i(x)$ is increasing in x for all $i = 1, \dots, n$.

ii. Existence

We now address existence of forward solutions and show that for each $i = 2, \dots, n$, there is x_i^* such that a unique, finite forward solution for $a_i(x)$ exists for all $x \in [0, x_i^*]$. In the limit as $x \rightarrow x_i^*$, $\lim_{x \rightarrow x_i^*} a_i(x) = \infty$. Furthermore, we show that $x_{i+1}^* < x_i^*$.

The forward solution $a_2(x)$ solves

$$cg(x) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) + 2(c-1)x.$$

The left-hand side satisfies $\lim_{x \rightarrow 0} cg(x) - \frac{c}{\lambda} = 0$ and increases in x . The right-hand side satisfies

$$\lim_{a_2 \rightarrow x} \left\{ \frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) + 2(c-1)x \right\} = 2(c-1)x \leq 0,$$

where the inequality is strict for $c < 1$ and $x > 0$. Moreover, the right-hand side is increasing and concave in a_2 with limiting value

$$\lim_{a_2 \rightarrow \infty} \left\{ \frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) + 2(c-1)x \right\} = \frac{c}{\lambda} + 2(c-1)x.$$

Hence, there exists a finite forward solution $a_2(x)$ if and only if

$$cg(x) - \frac{c}{\lambda} < \frac{c}{\lambda} + 2(c-1)x.$$

Since $cg(x) - \frac{c}{\lambda}$ is nonnegative and increasing in x and $\frac{c}{\lambda} + 2(c-1)x$ is positive for $x = 0$ and nonincreasing in x , there exists a unique value x_2^* such that

$$cg(x_2^*) - \frac{c}{\lambda} = \frac{c}{\lambda} + 2(c-1)x_2^*. \quad (22)$$

Hence, a finite forward solution $a_2(x)$ exists for all $x \in [0, x_2^*)$. In the limit as $x \rightarrow x_2^*$, we have $\lim_{x \rightarrow x_2^*} a_2(x) = \infty$.

Consider now the forward solution for $a_i(x)$ for $i = 3, \dots, n$. The forward solution a_i solves

$$cg(a_{i-1}(x) - a_{i-2}(x)) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_i - a_{i-1}(x)) - cg(a_i - a_{i-1}(x)) + 2(c-1)a_{i-1}(x).$$

The left-hand side satisfies $\lim_{x \rightarrow 0} cg(a_{i-1}(x) - a_{i-2}(x)) - \frac{c}{\lambda} = 0$ and is increasing in x . The right-hand side satisfies

$$\lim_{a_i \rightarrow a_{i-1}(x)} \frac{c}{\lambda} + c(a_i - a_{i-1}(x)) - cg(a_i - a_{i-1}(x)) + 2(c-1)a_{i-1}(x) = 2(c-1)a_{i-1}(x) \leq 0,$$

with strict inequality for $x > 0$ and $c < 1$. Moreover, the right-hand side is increasing and concave in a_{i-1} with limiting value

$$\lim_{a_i \rightarrow \infty} \frac{c}{\lambda} + c(a_i - a_{i-1}(x)) - cg(a_i - a_{i-1}(x)) + 2(c-1)a_{i-1}(x) = \frac{c}{\lambda} + 2(c-1)a_{i-1}(x).$$

Therefore, a unique solution for a_i exists if and only if

$$cg(a_{i-1}(x) - a_{i-2}(x)) - \frac{c}{\lambda} < \frac{c}{\lambda} + 2(c-1)a_{i-1}(x).$$

Given the derived properties of the forward solution, we have that $cg(a_{i-1}(x) - a_{i-2}(x)) - \frac{c}{\lambda}$ is nonnegative and increasing in x and $\frac{c}{\lambda} + 2(c-1)a_{i-1}(x)$ is positive for $x = 0$ and nonincreasing in x . Therefore, there exists a unique value x_i^* such that

$$cg(a_{i-1}(x_i^*) - a_{i-2}(x_i^*)) - \frac{c}{\lambda} = \frac{c}{\lambda} + 2(c-1)a_{i-1}(x_i^*). \quad (23)$$

Hence a finite forward solution $a_i(x)$ exists for all $x \in [0, x_i^*)$. In the limit as $x \rightarrow x_i^*$, we have $\lim_{x \rightarrow x_i^*} a_i(x) = \infty$.

Define

$$A_i(x) \equiv cg(a_{i-1}(x) - a_{i-2}(x)) - \frac{c}{\lambda} - \left(\frac{c}{\lambda} + 2(c-1)a_{i-1}(x) \right),$$

and similarly

$$A_{i+1}(x) \equiv cg(a_i(x) - a_{i-1}(x)) - \frac{c}{\lambda} - \left(\frac{c}{\lambda} + 2(c-1)a_i(x) \right).$$

Since $a_i(x) - a_{i-1}(x) > a_{i-1}(x) - a_{i-2}(x)$ and $a_i(x) > a_{i-1}(x)$ for all x , we have $A_{i+1}(x) > A_i(x)$. Moreover, both $A_{i+1}(x)$ and $A_i(x)$ are increasing in x . Letting x_i^* and x_{i+1}^* denote the values of x such that $A_i(x_i^*) = 0$ and $A_{i+1}(x_{i+1}^*) = 0$, we have $x_{i+1}^* < x_i^*$.

2. The fixed point argument

Take the forward solution for $a_i(x)$ for $i = 2, \dots, n$ and consider the difference between the left and the right side of (12), which we define as

$$\Delta_n(x) \equiv cg(a_n(x) - a_{n-1}(x)) - \frac{c}{\lambda} - \frac{c}{\lambda} - 2(c-1)a_n(x).$$

Differentiating $\Delta_n(x)$ with respect to x we get

$$\begin{aligned} \frac{d\Delta_n(x)}{dx} &= cg'(a_n(x) - a_{n-1}(x)) \left(\frac{da_n(x)}{dx} - \frac{da_{n-1}(x)}{dx} \right) - 2(c-1) \frac{da_n(x)}{dx} \\ &= cg'(a_n(x) - a_{n-1}(x)) \left(\frac{da_n(x)}{da_{n-1}(x)} - 1 \right) \frac{da_{n-1}(x)}{dx} - 2(c-1) \frac{da_n(x)}{dx}. \end{aligned}$$

Since $\frac{da_n(x)}{da_{n-1}(x)} > 1$, $\Delta_n(x)$ is strictly monotonic in x . This implies that there is at most one value of x that solves $\Delta_n(x) = 0$. Let \tilde{x}_n denote the value of x that satisfies $\Delta_n(\tilde{x}_n) = 0$ for given n , if it exists. To show that a fixed point exists, we need to show that \tilde{x}_n is such that the forward solution for $a_n(\tilde{x}_n)$ exists. To see this is true, note simply that $\Delta_n(\tilde{x}_n) = 0$ for $\tilde{x}_n = x_{n+1}^*$. That is, \tilde{x}_n is the value of x , such that forward solutions for $a_i(x)$ for $i = 2, \dots, n+1$ exist and are finite for all $x \in [0, \tilde{x}_n]$. Since $x_{n+1}^* < x_n^*$, the forward solutions for $i = 2, \dots, n$ exist and are finite at $x = \tilde{x}_n$. Hence, this completes the proof that there exists exactly one fixed point.

3. Induction

From the main text we have existence of an equilibrium for $n = 2$. We now show that existence of an equilibrium inducing $2(n+1)$ receiver actions implies the existence of an equilibrium inducing $2((n+1)+1)$ receiver actions. The equilibrium inducing $2(n+1)$ receiver actions is characterized by the thresholds $a_1^n = \tilde{x}_n$, $a_2^n(\tilde{x}_n), \dots, a_n^n(\tilde{x}_n)$ and the condition $\Delta_n(x) = cg(a_n(x) - a_{n-1}(x)) - \frac{c}{\lambda} - \frac{c}{\lambda} - 2(c-1)a_n(x) = 0$. Note that the forward solution for $a_{n+1}^n(x)$ satisfies $\lim_{x \rightarrow \tilde{x}_n} a_{n+1}^n(x) = \infty$. This implies that for all $x \in [0, \tilde{x}_n]$ the forward solution with $n+1$ steps exists. Now consider the function

$$\Delta_{n+1}(x) \equiv cg(a_{n+1}(x) - a_n(x)) - \frac{c}{\lambda} - \frac{c}{\lambda} - 2(c-1)a_{n+1}(x).$$

Clearly, $\lim_{x \rightarrow \tilde{x}_n} \Delta_{n+1}(x) > 0$, by the fact that $\lim_{x \rightarrow \tilde{x}_n} a_{n+1}^n(x) = \infty$. As $\Delta_{n+1}(x)$ is increasing in x , and $\lim_{x \rightarrow 0} \Delta_{n+1}(x) = -\frac{c}{\lambda} < 0$, there exists a unique value \tilde{x}_{n+1} such that $\Delta_{n+1}(\tilde{x}_{n+1}) = 0$. The thresholds $a_1^{n+1} = \tilde{x}_{n+1}, a_2^{n+1}(\tilde{x}_{n+1}), \dots, a_{n+1}^{n+1}(\tilde{x}_{n+1})$ characterize the unique class I equilibrium inducing $2(n+2)$ receiver actions.

Hence, for all n there exists a unique equilibrium inducing exactly $2(n+1)$ receiver actions.

Class II Equilibria:

A class II equilibrium is characterized by

$$a_1 = \frac{c}{\lambda} + c(a_2 - a_1) - cg(a_2 - a_1) - (1 - c)a_1$$

in addition to condition (11) for $i = 2, \dots, n-1$ and condition (12).

To construct a forward solution, take an arbitrary initial value x for the first threshold as given and compute $a_2(x)$ as the solution to

$$x = \frac{c}{\lambda} + c(a_2(x) - x) - cg(a_2(x) - x) - (1 - c)x.$$

We have $\lim_{a_2 \rightarrow x} \left(\frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) - (1 - c)x \right) = -(1 - c)x$ and $\lim_{a_2 \rightarrow \infty} \left(\frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) - (1 - c)x \right) = \frac{c}{\lambda} - (1 - c)x$. Hence, there is a unique finite forward solution $a_2(x)$ if and only if $x < \frac{c}{\lambda} - (1 - c)x$, or equivalently $(2 - c)x < \frac{c}{\lambda}$. Since $c \leq 1$, this is equivalent to $x < \frac{c}{\lambda(2-c)}$. We have $\lim_{x \rightarrow \frac{c}{\lambda(2-c)}} a_2(x) = \infty$. Likewise, for $x = 0$ we have $a_2(x)|_{x=0} = 0$.

Differentiating totally, we find

$$0 = (c(1 - g'(a_2(x) - x))) da_2 - (c(1 - g'(a_2(x) - x)) + (2 - c)) dx,$$

and so

$$\frac{da_2}{dx} = \frac{(c(1 - g'(a_2(x) - x)) + (2 - c))}{(c(1 - g'(a_2(x) - x)))} > 1.$$

Since the forward equations for $a_i(x)$ for $i = 3, \dots, n$ as well as the fixed point condition (12) are unchanged, all the remaining arguments are unchanged. We conclude that there is a unique class II equilibrium for all n . ■

Proof of Lemma 3. Since $a_1^n = \tilde{x}_n = x_{n+1}^*$ and $a_1^{n+1} = \tilde{x}_{n+1} = x_{n+2}^*$ it follows immediately from Lemma 2 that $a_i^{n+1} < a_i^n$ for $i = 1, \dots, n$. In particular, the argument follows from the

fact that the solution of the forward equation is monotonic in the initial condition, x . Hence, it suffices to prove that $a_i^{n+1} > a_i^n$ for $i = 1, \dots, n$.

We start with two preliminary observations.

Firstly, the “next” solution of the forward equation, $a_{i+1}^k(x)$ for $i = 1, \dots, k-1$, $k = n, n+1$ is monotonic in $a_i^k(x)$, and the length of the previous interval, $a_i^k(x) - a_{i-1}^k(x)$. To see this, note that the forward equations for a_2^k , a_3^k , and a_{i+1}^k , for $i = 3, \dots, k-1$ and $k = n, n+1$, satisfy:

$$cg(x) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_2^k - x) - cg(a_2^k - x) + 2(c-1)x,$$

$$cg(a_2^k(x) - x) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_3^k - a_2^k(x)) - cg(a_3^k - a_2^k(x)) + 2(c-1)a_2^k(x),$$

and

$$cg(a_i^k(x) - a_{i-1}^k(x)) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_{i+1}^k - a_i^k(x)) - cg(a_{i+1}^k - a_i^k(x)) + 2(c-1)a_i^k(x).$$

The conclusion follows from the fact that $a_i^k(x)$ decreases the value of the right-hand side and increases the value of the left-hand side. Moreover, the left-hand side is increasing in $a_i^k(x) - a_{i-1}^k(x)$.

Secondly, it is impossible that $a_{n+1}^{n+1}(\tilde{x}_{n+1}) < a_n^n(\tilde{x}_n)$ and $a_{n+1}^{n+1}(\tilde{x}_{n+1}) - a_{n+1}^{n+1}(\tilde{x}_{n+1}) < a_n^n(\tilde{x}_n) - a_{n-1}^n(\tilde{x}_n)$. If these conditions would hold, then one of the fixed-point conditions,

$$0 = cg(a_n^n(\tilde{x}_n) - a_{n-1}^n(\tilde{x}_n)) - \frac{c}{\lambda} - \frac{c}{\lambda} - 2(c-1)a_n^n(\tilde{x}_n)$$

and

$$0 = cg(a_{n+1}^{n+1}(\tilde{x}_{n+1}) - a_n^{n+1}(\tilde{x}_{n+1})) - \frac{c}{\lambda} - \frac{c}{\lambda} - 2(c-1)a_{n+1}^{n+1}(\tilde{x}_{n+1})$$

would necessarily be violated.

We now show that $a_{j+1}^{n+1} > a_j^n$ for all $j \leq n$.

Suppose that this were not true and let the property be violated for the first time at $j = l$.

Now suppose $a_{j+1}^{n+1}(\tilde{x}_{n+1}) > a_j^n(\tilde{x}_n)$ for all $j = 1, \dots, l-1$ and $a_{l+1}^{n+1}(\tilde{x}_{n+1}) < a_l^n(\tilde{x}_n)$. Taken together, these inequalities immediately imply that $a_{l+1}^{n+1}(\tilde{x}_{n+1}) - a_l^{n+1}(\tilde{x}_{n+1}) < a_l^n(\tilde{x}_n) - a_{l-1}^n(\tilde{x}_n)$. In turn, the monotonicity property of the next forward solution implies then that $a_{l+2}^{n+1}(\tilde{x}_{n+1}) < a_{l+1}^n(\tilde{x}_n)$.

It also follows then that $a_{l+2}^{n+1}(\tilde{x}_{n+1}) - a_{l+1}^{n+1}(\tilde{x}_{n+1}) < a_{l+1}^n(\tilde{x}_n) - a_l^n(\tilde{x}_n)$. To see this, suppose instead that $a_{l+2}^{n+1}(\tilde{x}_{n+1}) - a_{l+1}^{n+1}(\tilde{x}_{n+1}) \geq a_{l+1}^n(\tilde{x}_n) - a_l^n(\tilde{x}_n)$ or equivalently that $a_{l+2}^{n+1}(\tilde{x}_{n+1}) \geq a_{l+1}^n(\tilde{x}_n) + \{a_{l+1}^{n+1}(\tilde{x}_{n+1}) - a_l^n(\tilde{x}_n)\}$. However, this is impossible since both $a_{l+2}^{n+1}(\tilde{x}_{n+1}) < a_{l+1}^n(\tilde{x}_n)$ and $a_{l+1}^{n+1}(\tilde{x}_{n+1}) < a_l^n(\tilde{x}_n)$. Hence, the claim follows.

However, if $a_{l+2}^{n+1}(\tilde{x}_{n+1}) < a_{l+1}^n(\tilde{x}_n)$ and $a_{l+2}^{n+1}(\tilde{x}_{n+1}) - a_{l+1}^{n+1}(\tilde{x}_{n+1}) < a_{l+1}^n(\tilde{x}_n) - a_l^n(\tilde{x}_n)$, then $a_{l+3}^{n+1}(\tilde{x}_{n+1}) < a_{l+2}^n(\tilde{x}_n)$ and so forth. Hence, we would have $a_{j+1}^{n+1}(\tilde{x}_{n+1}) < a_j^n(\tilde{x}_n)$ and $a_{j+1}^{n+1}(\tilde{x}_{n+1}) - a_j^{n+1}(\tilde{x}_{n+1}) < a_j^n(\tilde{x}_n) - a_{j-1}^n(\tilde{x}_n)$ for all $j \geq l$ and in particular for $j = n$, leading to a violation of one of the fixed-point conditions.

The same argument can be given for a class II equilibrium. This is omitted. ■

Proof of Proposition 1. i) The limit of $\{a_1^n, a_2^n, \dots, a_n^n\}$ as n goes to infinity is an equilibrium.

To see this, note that the sequence $\{\tilde{x}_n\}_{n=2}^\infty$ is monotone decreasing and bounded from below by zero. Hence it converges. As shown in Lemma 3, the sequence $\{a_n^n\}_{n=2}^\infty$ is monotone increasing. Moreover, the fixed point condition implies that the sequence is bounded from above, $a_n^n < \frac{c}{1-c} \frac{1}{2\lambda}$ for all n . Hence, it also converges.

Given that the largest threshold a_n^n converges, by construction of the equilibrium all thresholds below must also converge: We have constructed an equilibrium as a forward equation for given initial interval length $a_1^n = x$ together with a closure condition (12), that determines the equilibrium value of x . The same equilibrium can also be constructed as a backward equation, that starts with a given threshold $a_n^n = q$, which recursively determines $a_{n-1}^n(q)$ and so on, and that together with a closure condition determines the equilibrium point q . An equilibrium set of thresholds satisfies both the forward and the backward equation. Given that the largest threshold converges, all thresholds below must also converge since, they are linked recursively to a_n^n via the backward equation. Hence by construction, the limit is an equilibrium.

ii) From Lemma 3, we have $\tilde{x}_{n+1} < \tilde{x}_n$: the higher the number of thresholds, the shorter the first interval. We now show that in the limit as $n \rightarrow \infty$, we have $\lim_{n \rightarrow \infty} \tilde{x}_n = 0$.

Recall that the forward solution for $a_n(x)$ exists for $x \leq x_n^*$, where x_n^* satisfies

$$cg(a_{n-1}(x_n^*) - a_{n-2}(x_n^*)) - \frac{c}{\lambda} = \frac{c}{\lambda} + 2(c-1)a_{n-1}^n(x_n^*).$$

Monotonicity of the forward solutions, $a_k(x) > a_{k-1}(x)$, and increasing length of the in-

tervals, $a_k(x) - a_{k-1}(x) > a_{k-1}(x) - a_{k-2}(x)$, imply that for any $x > 0$, there is a k such that

$$cg(a_{k-1}(x) - a_{k-2}(x)) - \frac{c}{\lambda} \leq \frac{c}{\lambda} + 2(c-1)a_{k-1}(x)$$

and

$$cg(a_k(x) - a_{k-1}(x)) - \frac{c}{\lambda} > \frac{c}{\lambda} + 2(c-1)a_k(x).$$

For a fixed length x of the first interval, the forward solution will necessarily cease to have a solution at some point. Hence, in an infinite equilibrium we have $\lim_{n \rightarrow \infty} \tilde{x}_n = 0$, the length of the first interval goes to zero.

The proof for the case of a class II equilibrium is virtually the same and hence omitted.

■

Proof of Proposition 2. Let $c > 1$. Consider class I equilibria first. The forward equation for a_2 is given by

$$cg(x) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) + 2(c-1)x. \quad (24)$$

The left-hand side satisfies $\lim_{x \rightarrow 0} cg(x) - \frac{c}{\lambda} = 0$ and is increasing and convex in x , with slope between $\frac{c}{2}$ and c . The right-hand side satisfies

$$\lim_{a_2 \rightarrow x} \frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) + 2(c-1)x = 2(c-1)x \geq 0,$$

where the inequality is strict for $x > 0$. Moreover, the right-hand side is increasing and concave in a_2 with limit

$$\lim_{a_2 \rightarrow \infty} \frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) + 2(c-1)x = \frac{c}{\lambda} + 2(c-1)x.$$

Hence, there exists a forward solution $a_2(x)$ if and only if

$$2(c-1)x < cg(x) - \frac{c}{\lambda} < \frac{c}{\lambda} + 2(c-1)x.$$

There are three cases to distinguish. For $c \in (1, \frac{4}{3}]$, we have $2(c-1) \leq \frac{c}{2}$ and thus $2(c-1) \leq cg'(x)$ for all x , since $g'(x) \geq \frac{1}{2}$ for all x . Therefore, the former inequality is satisfied for all $x > 0$. The latter inequality holds for x small since $\lim_{x \rightarrow 0} cg(x) - \frac{c}{\lambda} = 0 < \frac{c}{\lambda}$. As x increases, the latter inequality eventually ceases to hold, since $c > 2(c-1)$ and thus

$cg'(x) > 2(c-1)$ for x sufficiently large, as $g'(x)$ tends to one as $x \rightarrow \infty$. Hence, there exists a solution $a_2(x)$ for x small enough.

For $c \in (\frac{4}{3}, 2)$, we have $\frac{c}{2} < 2(c-1) < c$. Since $\lim_{x \rightarrow 0} g'(x) = \frac{1}{2}$, we have $2(c-1)x \geq cg(x) - \frac{c}{\lambda}$ for x positive and small, so that the former inequality is violated for x small. Thus, no solution for $a_2(x)$ exists if x is close to zero.

Finally, for $c \geq 2$ we have $2(c-1) \geq c$ and therefore $cg'(x) \leq 2(c-1)$ for all x . Hence, no solution exists for $a_2(x)$. This implies that at most two actions can be induced in equilibrium.

Hence, it follows immediately that x is bounded away from zero for $c > \frac{4}{3}$. Consider therefore the case where $c \in (1, \frac{4}{3}]$. By Lemma 3, the solution must satisfy $a_2(x) - x > x$ for any equilibrium. We show that this condition is violated for small x . Suppose that $a_2 - x = x$. We define the difference between the rhs and the lhs of condition (24) at $a_2 - x = x$ as

$$D(x) \equiv \frac{c}{\lambda} + cx - cg(x) + 2(c-1)x + \frac{c}{\lambda} - cg(x).$$

If $D(x)$ is positive (negative), then a_2 needs to decrease (increase) to satisfy the forward equation, since the right-hand side of (24) is increasing in a_2 . We have $\lim_{x \rightarrow 0} D(x) = 0$. Moreover, the slope of $D(x)$ at $x = 0$ is $D'(x)|_{x=0} = 2(c-1) > 0$. Hence, for x small, the solution to the forward equation would satisfy $a_2(x) - x < x$, violating the increasing interval property. However, since any equilibrium needs to have this property, x is bounded away from zero.

Note that this argument extends to any equilibrium with zero as a threshold, not just symmetric equilibria.

Consider now a class II equilibrium. Given x , $a_2(x)$ is the value of a_2 that solves

$$cg(a_2 - x) - \frac{c}{\lambda} = c(a_2 - x) + (c-2)x. \quad (25)$$

Note first that no solution $a_2(x)$ exists for $c \geq 2$. To see this, note that

$$\lim_{a_2 \rightarrow x} \frac{c}{\lambda} + c(a_2(x) - x) - cg(a_2(x) - x) - (2-c)x = -(2-c)x \geq 0$$

for any $c \geq 2$ and any $x \geq 0$. Therefore, we consider $1 < c < 2$ from now on. Equation (25) has a solution for $x < \frac{c}{\lambda(2-c)}$, which satisfies $\lim_{x \rightarrow 0} a_2(x) = 0$ and moreover,

$$\frac{da_2}{dx} = \frac{c(1 - g'(a_2 - x)) + (2-c)}{c(1 - g'(a_2 - x))} > 1.$$

Rearranging (25) we can write

$$-2 \frac{(c-1)}{(c-2)} \left(cg(a_2(x) - x) - \frac{c}{\lambda} - c(a_2(x) - x) \right) = -2(c-1)x.$$

Given x and $a_2(x)$, $a_3(x)$ is the value of a_3 that solves

$$cg(a_2(x) - x) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_3 - a_2(x)) - cg(a_3 - a_2(x)) + 2(c-1)a_2(x). \quad (26)$$

Adding up both equations and rearranging, we can conclude that $a_3(x)$ is the value of a_3 that solves

$$0 = \frac{c}{\lambda} + c(a_3 - a_2(x)) - cg(a_3 - a_2(x)) + 4 \frac{c-1}{2-c} (a_2(x) - x) - \frac{c}{2-c} \left(cg(a_2(x) - x) - \frac{c}{\lambda} \right). \quad (27)$$

Note that the right-hand side of this equation is increasing in a_3 and that $a_3(x)$ is the unique value that sets the expression equal to zero. We wish to show that the expression is strictly positive for $a_3 - a_2(x) = a_2(x) - x$, to get that we would have to have $a_3(x) - a_2(x) < a_2(x) - x$, in contradiction to the increasing interval property.

Note that the right-hand side of (27) depends only on the differences $a_2(x) - x$ and $a_3 - a_2(x)$. Moreover, note that $a_2(x) - x$ goes to zero as x goes to zero. Let $z = a_2(x) - x$ and evaluate the rhs of (27) at $a_3 - a_2(x) = z$. We obtain

$$F(z) \equiv cz + 4 \frac{c-1}{2-c} z + \frac{2}{2-c} \left(\frac{c}{\lambda} - cg(z) \right).$$

$F(z)$ is concave in z . In the limit as x and hence z tends to zero, we find

$$F'(z)|_{z=0} = \frac{5c - c^2 - 4}{2 - c},$$

where we use that $g'(z)|_{z=0} = \frac{1}{2}$. For $c \in (1, 2)$, we have $5c - c^2 - 4 > 0$ and we know that $F(z) > 0$ for z small. Since, the right-hand side of (27) is increasing in a_3 , to restore equality with zero, a_3 needs to decrease, which would imply that $a_3(x) - a_2(x) < a_2(x) - x$. However, this contradicts the the increasing interval property of any equilibrium. This implies that x must be bounded away from zero.

Consider now an asymmetric interval around zero. Fix an arbitrary point $a_{-1} = -y < 0$ and an arbitrary point $a_1 = x > 0$. We have $\Pr(\theta \in (0, x]) = \frac{1}{2}(1 - e^{-\lambda x})$ and $\Pr(\theta \in (-y, 0]) =$

$\Pr(\theta \in [0, y)) = \frac{1}{2}(1 - e^{-\lambda y})$. Let $\delta(x, y) \equiv \frac{(1 - e^{-\lambda x})}{(1 - e^{-\lambda x}) + (1 - e^{-\lambda y})}$, then the conditional expectation over the interval $[-y, x]$ is

$$\omega(x, y) \equiv \delta(x, y) \left(\frac{1}{\lambda} + x - g(x) \right) - (1 - \delta(x, y)) \left(\frac{1}{\lambda} + y - g(y) \right).$$

Clearly, $\omega(x, y) \geq 0$ for $x \geq y$. The forward solution $a_2(x, y)$ is the value of a_2 that solves

$$-c\omega(x, y) = \frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) + (c - 2)x. \quad (28)$$

Note first that for $c \geq 2$ necessarily $x < y$. However, we need to have $y < x$ to get a solution for the isomorphic problem on the negative orthant. Hence for $c \geq 2$ the forward solution does not exist in both directions.

Now consider $1 < c < 2$. A solution $a_2(x, y)$ exists if and only if

$$(c - 2)x < -c\omega(x, y) < \frac{c}{\lambda} + (c - 2)x.$$

Note that this is always satisfied for $x = y$, and hence by continuity also for x close to y . The condition determining a_3 is unchanged,

$$cg(a_2(x, y) - x) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_3 - a_2(x, y)) - cg(a_3 - a_2(x, y)) + 2(c - 1)a_2(x, y). \quad (29)$$

Rearranging (28), we can write

$$\frac{2(c - 1)}{(c - 2)}c\omega(x, y) - \frac{2(c - 1)}{(c - 2)} \left(cg(a_2 - x) - \frac{c}{\lambda} - c(a_2 - x) \right) = -2(c - 1)x.$$

Adding up with (29),

$$\begin{aligned} & \frac{2(c - 1)}{(c - 2)}c\omega(x, y) \\ &= \frac{c}{\lambda} + c(a_3(x, y) - a_2(x, y)) - cg(a_3(x, y) - a_2(x, y)) + 4\frac{c - 1}{2 - c}(a_2(x, y) - x) \\ & \quad - \frac{c}{2 - c} \left(cg(a_2(x, y) - x) - \frac{c}{\lambda} \right). \end{aligned}$$

For $x > y$, the left-hand side is strictly negative. On the other hand, the right-hand side is strictly positive at $a_3(x, y) - a_2(x, y) = a_2(x, y) - x = z$ for z small. Hence, the argument

extends to this case. Note that by symmetry of the distribution, the case $x < y$ causes the isomorphic problem on the negative orthant. Hence, the size of the interval around zero must be bounded away from zero.

Finally, note that if the interval closest to the origin has length bounded away from zero, then all intervals must be of strictly positive length. Hence, only a finite number of intervals can exist and a finite number of receiver actions is induced. ■

Proof of Proposition 3. Define $z_i^n \equiv a_i^n - a_{i-1}^n$ for $i = 1, \dots, n$. We wish to derive a closed form representation for $\sum_{i=1}^{n+1} p_i^n (c\mu_i^n - \frac{c}{\lambda})^2$.

For $k = 1, \dots, n+1$, we define

$$X_k^n(a_{k-1}^n) \equiv \frac{\sum_{i=k}^{n+1} p_i^n (c\mu_i^n - \frac{c}{\lambda})^2}{\sum_{i=k}^{n+1} p_i^n} \quad \text{and} \quad \hat{p}_k^n(a_{k-1}^n) \equiv \frac{p_k^n}{\sum_{i=k}^{n+1} p_i^n}.$$

$X_k^n(a_{k-1}^n)$ is the expected squared deviation of the receiver's choice from its mean value, $\frac{c}{\lambda}$, conditional on truncation above a_{k-1}^n . Moreover, $\hat{p}_k^n(a_{k-1}^n) = \Pr(\theta \in \Theta_k^n | \theta \geq a_{k-1}^n)$.

We first derive an exact expression for $X_n^n(a_{n-1}^n)$. Then, we impose as an inductive hypothesis that $X_k^n(a_{k-1}^n)$ has the same functional form as $X_n^n(a_{n-1}^n)$ has and show that this implies that $X_{k-1}^n(a_{k-2}^n)$ does have the same functional form that $X_k^n(a_{k-1}^n)$ has.

Clearly, we have $\hat{p}_{n+1}^n(a_n^n) = 1$ and since $\mu_{n+1}^n = \frac{1}{\lambda} + a_n^n$, we get

$$X_{n+1}^n(a_n^n) = \left(c\mu_{n+1}^n - \frac{c}{\lambda}\right)^2 = \left(c \sum_{j=1}^n z_j^n\right)^2.$$

For $n = 1$, the expression takes value $c^2(z_1^1)^2$.

We now use $X_{n+1}^n(a_n^n)$ to compute $\sum_{i=n}^{n+1} \frac{p_i^n}{p_n^n + p_{n+1}^n} (c\mu_i^n - \frac{c}{\lambda})^2$, the expected squared deviation of the receiver's choice from its mean conditional on $\theta \geq a_{n-1}^n$. We can write

$$X_n^n(a_{n-1}^n) = \hat{p}_n(a_{n-1}^n) \left(c\mu_n^n - \frac{c}{\lambda}\right)^2 + (1 - \hat{p}_n(a_{n-1}^n)) X_{n+1}^n(a_n^n).$$

For the exponential distribution, $\sum_{i=n}^{n+1} \frac{p_i^n}{p_n^n + p_{n+1}^n} = 1 - \exp(-\lambda z_n^n)$. The indifference condition of type a_n^n pins down $\hat{p}_n(a_{n-1}^n)$. To see this, note that the indifference condition of type a_n^n ,

$-c\mu_n^n = c\mu_{n+1}^n - 2\sum_{j=1}^n z_j^n$, written explicitly takes the form:

$$c \frac{z_n^n}{1 - \exp(-\lambda z_n^n)} - \frac{c}{\lambda} - c \sum_{j=1}^n z_j^n = \frac{c}{\lambda} + c \sum_{j=1}^n z_j^n - 2 \sum_{j=1}^n z_j^n.$$

Rearranging and simplifying, we have

$$\hat{p}_n^n(a_{n-1}^n) = 1 - \exp(-\lambda z_n^n) = \frac{cz_n^n}{2\frac{c}{\lambda} - 2(1-c)\sum_{j=1}^n z_j^n}.$$

Hence, we can write

$$\begin{aligned} X_n^n(a_{n-1}^n) &= \frac{cz_n^n}{2\frac{c}{\lambda} - 2(1-c)\sum_{j=1}^n z_j^n} \left((2-c)\sum_{j=1}^n z_j^n - \frac{2c}{\lambda} \right)^2 + \left(1 - \frac{cz_n^n}{2\frac{c}{\lambda} - 2(1-c)\sum_{j=1}^n z_j^n} \right) \left(c \sum_{j=1}^n z_j^n \right)^2. \end{aligned}$$

Tedious but straightforward algebra shows that

$$X_n^n(a_{n-1}^n) = cz_n^n \left(\frac{2c}{\lambda} - ((2-c)z_n^n + 2(1-c)a_{n-1}^n) \right) + c^2(a_{n-1}^n)^2.$$

Using the indifference condition of type a_n^n , $(2-c)\sum_{j=1}^n z_j^n = \frac{c}{\lambda} + c\mu_n^n$, we can also write

$$X_n^n(a_{n-1}^n) = \frac{c}{(2-c)} \left(\frac{c}{\lambda} + c\mu_n^n - (2-c)a_{n-1}^n \right) \left(\frac{c}{\lambda} - c\mu_n^n + ca_{n-1}^n \right) + c^2(a_{n-1}^n)^2.$$

Notice that for $n = 1$, which implies that $a_{n-1}^n = 0$, we obtain the exact value of the expected squared deviation of the receiver's choice from their mean.

Suppose now as an inductive hypothesis that

$$X_k^n(a_{k-1}^n) = \frac{c}{(2-c)} \left(\frac{c}{\lambda} + c\mu_k^n - (2-c)a_{k-1}^n \right) \left(\frac{c}{\lambda} - c\mu_k^n + ca_{k-1}^n \right) + c^2(a_{k-1}^n)^2.$$

We can write

$$\begin{aligned} X_{k-1}^n(a_{k-2}^n) &= \hat{p}_{k-1}^n(a_{k-2}^n) \left(c\mu_{k-1}^n - \frac{c}{\lambda} \right)^2 + (1 - \hat{p}_{k-1}^n(a_{k-2}^n)) X_k^n(a_{k-1}^n) \\ &= (1 - \exp(-\lambda z_{k-1}^n)) \left(c\mu_{k-1}^n - \frac{c}{\lambda} \right)^2 + \exp(-\lambda z_{k-1}^n) X_k^n(a_{k-1}^n). \end{aligned}$$

Note that $\frac{\exp(-\lambda z_{k-1}^n)}{\sum_{i=k}^{n+1} p_i^n} = \frac{\exp(-\lambda(a_{k-1}^n - a_{k-2}^n))}{\exp(-\lambda a_{k-1}^n)} = \exp(\lambda a_{k-2}^n) = \frac{1}{\sum_{i=k-1}^{n+1} p_i^n}$, so the factor $\exp(-\lambda z_{k-1}^n)$ adjusts the term $X_k^n(a_{k-1}^n)$ to reflect the change in the truncation point from a_{k-1}^n to a_{k-2}^n .

Writing the indifference condition of type a_{k-1}^n , $-c\mu_{k-1}^n = c\mu_k^n - 2\sum_{j=1}^{k-1} z_j^n$, in explicit form, we have

$$c \frac{z_{k-1}^n}{1 - \exp(-\lambda z_{k-1}^n)} - \frac{c}{\lambda} - c \sum_{j=1}^{k-1} z_j^n = c\mu_k^n - 2 \sum_{j=1}^{k-1} z_j^n.$$

This delivers the probability distribution as

$$\hat{p}_{k-1}^n(a_{k-2}^n) = 1 - \exp(-\lambda z_{k-1}^n) = \frac{cz_{k-1}^n}{\frac{c}{\lambda} + c\mu_k^n - (2-c) \sum_{j=1}^{k-1} z_j^n}.$$

Substituting the inductive hypothesis and the probability distribution, we have

$$\begin{aligned} X_{k-1}^n(a_{k-2}^n) &= \frac{cz_{k-1}^n}{\frac{c}{\lambda} + c\mu_k^n - (2-c) \sum_{j=1}^{k-1} z_j^n} \left(c\mu_{k-1}^n - \frac{c}{\lambda} \right)^2 \\ &\quad + \frac{\frac{c}{\lambda} + c\mu_k^n - (2-c) \sum_{j=1}^{k-1} z_j^n - cz_{k-1}^n}{\frac{c}{\lambda} + c\mu_k^n - (2-c) \sum_{j=1}^{k-1} z_j^n} \\ &\quad \left(\frac{c}{(2-c)} \left(\frac{c}{\lambda} + c\mu_k^n - (2-c) a_{k-1}^n \right) \left(\frac{c}{\lambda} - c\mu_k^n + ca_{k-1}^n \right) + c^2 (a_{k-1}^n)^2 \right). \end{aligned}$$

Substituting for $c \sum_{j=1}^{k-1} z_j^n = ca_{k-1}^n = ca_{k-2}^n + cz_{k-1}^n$, and using the indifference condition of type a_{k-1}^n , $-c\mu_{k-1}^n = c\mu_k^n - 2(a_{k-2}^n + z_{k-1}^n)$, we find

$$\begin{aligned} X_{k-1}^n(a_{k-2}^n) &= \frac{cz_{k-1}^n}{\frac{c}{\lambda} - c\mu_{k-1}^n + ca_{k-2}^n + cz_{k-1}^n} \left(c\mu_{k-1}^n - \frac{c}{\lambda} \right)^2 \\ &\quad + \frac{\frac{c}{\lambda} - c\mu_{k-1}^n + ca_{k-2}^n}{\frac{c}{\lambda} - c\mu_{k-1}^n + ca_{k-2}^n + cz_{k-1}^n} \\ &\quad \left(\frac{c}{(2-c)} \left(\frac{c}{\lambda} - c\mu_{k-1}^n + ca_{k-2}^n + cz_{k-1}^n \right) \left(\frac{c}{\lambda} - (2-c) (a_{k-2}^n + z_{k-1}^n) + c\mu_{k-1}^n \right) \right. \\ &\quad \left. + c^2 (a_{k-2}^n + z_{k-1}^n)^2 \right). \end{aligned}$$

Tedious but straightforward algebra shows that

$$X_{k-1}^n(a_{k-2}) = \frac{c}{(2-c)} \left(\frac{c}{\lambda} + c\mu_{k-1}^n - (2-c)a_{k-2}^n \right) \left(\frac{c}{\lambda} - c\mu_{k-1}^n + ca_{k-2}^n \right) + c^2 (a_{k-2}^n)^2,$$

completing the inductive argument.

To obtain the statement in the proposition, note that in a class I equilibrium $a_0^n = 0$ and so we can write

$$X_1^n(0) = \frac{c}{(2-c)} \left(\frac{c}{\lambda} + c\mu_1^n \right) \left(\frac{c}{\lambda} - c\mu_1^n \right).$$

Since $X_1^n(0) = \sum_{i=1}^{n+1} p_i^n (c\mu_i^n - \frac{c}{\lambda})^2$, we have shown that

$$\sum_{i=1}^{n+1} p_i^n \left(c\mu_i^n - \frac{c}{\lambda} \right)^2 = \frac{c}{2-c} \left(\left(\frac{c}{\lambda} \right)^2 - (c\mu_1^n)^2 \right).$$

Since $\sum_{i=1}^{n+1} p_i^n (c\mu_i^n - \frac{c}{\lambda})^2 = \sum_{i=1}^{n+1} p_i^n (c\mu_i^n)^2 - (\frac{c}{\lambda})^2$ - by the fact that $\sum_{i=1}^{n+1} p_i^n \mu_i^n = \frac{1}{\lambda}$ - we can write

$$c^2 \sum_{i=1}^{n+1} p_i^n (\mu_i^n)^2 = \frac{2}{2-c} \left(\frac{c}{\lambda} \right)^2 - \frac{c}{2-c} (c\mu_1^n)^2.$$

Hence we have shown that

$$\sum_{i=1}^{n+1} p_i^n (\mu_i^n)^2 = \frac{1}{2-c} \frac{2}{\lambda^2} - \frac{c}{2-c} (\mu_1^n)^2.$$

Noting that $Var(\theta) = \frac{2}{\lambda^2}$, completes the proof. ■

References

- [1] Alonso, R., Dessein, W., & Matouschek, N. (2008). "When does coordination require centralization?". *American economic review*, 98(1), 145-179.
- [2] Alonso, R., & Rantakari, H. (2013). "The art of brevity". SSRN discussion paper 2306824.
- [3] Argenziano, R., Severinov, S., and Squintani, F. "Strategic Information Acquisition and Transmission" forthcoming *AEJ:microeconomics*.
- [4] Blackwell, D., (1951). "Comparison of experiments" In: Neyman, J. (Ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley, Univ. of California Press, 93–102.
- [5] Blume, A., Board, O.J., and Kawamura, K. (2007) "Noisy talk" *Theoretical Economics* 2(4), 395-440.
- [6] Chen, Y., Kartik, N., and Sobel J. (2008) "Selecting Cheap Talk Equilibria", *Econometrica*, 76(1), 117–136.
- [7] Crawford, V. and Sobel, J. (1982) "Strategic Information Transmission" *Econometrica*, 50(6), 1431–1451.
- [8] Crémer, J. and Khalil, F. (1992) "Gathering Information before Signing a Contract" *American Economic Review*, 82, 566-578.
- [9] Crémer, J., Khalil, F., and Rochet, J.-C. (1998) "Contracts and Productive Information Gathering", *Games and Economic Behavior*, 25, 174-193.
- [10] Deimen, I. and Szalay, D. (2014) "Smooth, strategic communication" CEPR DP 10190.
- [11] Deimen, I. and Szalay, D. (2015) "Information, Authority, and Smooth Communication in Organizations" CEPR DP 10969.

- [12] Demski, J. S., & Sappington, D. E. (1987). Delegated expertise. *Journal of Accounting Research*, 25(1), 68-89.
- [13] Dessein, W. (2002) "Authority and Communication in Organizations" *Review of Economic Studies* 69, 811-838.
- [14] Di Pei, H. (2015) "Communication with Endogenous Information Acquisition" *Journal of Economic Theory*, 160, 132-149.
- [15] Eső, P, and Szalay, D. (2015) "Incomplete Language as an Incentive Device" unpublished manuscript.
- [16] Frug, A. (2016) "Strategic Gradual Learning and Information Transmission", unpublished manuscript.
- [17] Goltsman, M., Horner, J., Pavlov, G., and Squintani, F. (2009) "Mediation, arbitration and negotiation". *Journal of Economic Theory* 144(4), 1397 – 1420.
- [18] Gordon, S. (2010) "On Infinite Cheap Talk Equilibria" unpublished manuscript.
- [19] Gromb, D., & Martimort, D. (2007) "Collusion and the organization of delegated expertise". *Journal of Economic Theory*, 137(1), 271-299.
- [20] Ivanov, M. (2010) "Informational Control and Organizational Design." *Journal of Economic Theory*, 145, 721-751.
- [21] Kamenica, E and Gentzkow, M. (2011). "Bayesian persuasion" *American Economic Review*, 101(6), 2590-2615.
- [22] Kotz, S., Kozubowski, T., and K. Podgorski (2001) "The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance", Springer, New York.
- [23] Lewis, T. R., & Sappington, D. E. (1997). "Information management in incentive problems". *Journal of Political Economy*, 105(4), 796-821.

- [24] Moscarini, G. (2007) "Competence implies credibility", *American Economic Review* 97(1), 37-63.
- [25] Ottaviani, M. and Sorensen, P. (2006), "Professional Advice" *Journal of Economic Theory* 126(1), 120–142.
- [26] Sobel, J. (2013) "Giving and Receiving Advice", *Advances in Economics and Econometrics*, Tenth World Congress.
- [27] Szalay, D. (2005). "The economics of clear advice and extreme options". *Review of Economic Studies*, 72(4), 1173-1198.
- [28] Szalay, D. (2009) "Contracts with Endogenous Information" *Games and Economic Behavior* 65, 586-625.
- [29] Szalay, D. (2012), "Strategic Information Transmission and Stochastic Orders" working paper University of Bonn.