Delegated Expertise, Authority, and Communication*

Inga Deimen
University of Arizona

and

Dezső Szalay
University of Bonn and CEPR

January 2018

Abstract

A decision-maker needs to reach a decision and relies on an expert to acquire information. Ideal actions of expert and decision-maker are partially aligned and the expert chooses what to learn about each. The decision-maker can either get advice from the expert or delegate decision-making to him. Under delegation, the expert learns his privately optimal action and chooses it. Under communication, advice based on such information is discounted, resulting in losses from strategic communication. We characterize the communication problems that make the expert acquire information of equal use to expert and decision-maker. In these problems, communication outperforms delegation.

JEL: D82

Keywords: endogenous information, delegated expertise, strategic information transmission, delegation, effectiveness of biased communication

*We thank Jeff Ely and four anonymous referees for excellent comments. Many thanks to Andreas Blume, Jacques Crémer, Wouter Dessein, Peter Eső, Johannes Hörner, Ian Jewitt, Steve Matthews, Ines Moreno-de-Barreda, Herakles Polemarchakis, Soenje Reiche, Andrew Rhodes, Hamid Sabourian, and seminar participants at Arizona, Arizona State, Bonn, Cambridge, Columbia, Crest, ESSET (Gerzensee), Southampton, St. Gallen, Toulouse, UPenn, Warwick, and Zürich for insightful discussions. 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@email.arizona.edu 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[1] 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? This paper makes a case for communication as a complement to delegated expertise, providing conditions under which communication unambiguously dominates delegated decision-making.

We envision a decision-problem that involves a change of policy away from some known status quo, e.g., adapting a design to new market conditions, adjusting a portfolio in response to new information, choosing a new project, adapting a business plan in response to changes in the environment of the firm, and so on. Naturally, the status quo is the optimal action based on the information currently available, but additional information will likely lead to a revision of plans. The decision-maker can consult an expert for advice or help. The contractual options for the decision-maker are incomplete (Grossman and Hart (1986) and Hart and Moore (1990)): as in Aghion and Tirole (1997), the decision-maker can only choose the allocation of authority. That is, she can either simply ask for advice or entrust the expert with decision-making altogether. In contrast to their approach, the decision-maker has no time or means to become informed herself. However, she can communicate with the expert and infer her preferred action indirectly, at least to some extent.[2]

We construct a novel model with linear Bayesian updating rules in which the decision-maker's inferred optimal action is a compromise between the expert's preferred action and the status quo. The expert observes noisy signals about the optimal actions from his and the decision-maker's perspective; he is free to choose the precision of each of the signals. We think of a transparent environment where the expert's information acquisition is overt,

[1] Demski and Sappington (1987) have coined this term. An expert is an agent endowed with a technology to acquire information.

[2] In Aghion and Tirole (1997), an uninformed individual would stick to the status quo, because there exist disastrous projects. We drop this assumption.

as is the case, e.g., for inhouse consulting. While the precision levels of the signals are observable, the actual realizations are privately observed by the expert. We abstract from real costs of information acquisition and focus on the strategic costs of different information acquisition strategies instead. The extent to which the decision-maker follows the expert's advice depends crucially on what type of information the expert acquires. Moreover, when allocating authority the decision-maker takes into account that the expert's information acquisition will depend on her choice of institution.

Our main findings are as follows. If the decision-maker transfers formal authority to the expert, then the expert acquires perfect information about his preferred action and takes it. The decision-maker benefits from this policy, but only to some extent. The advantage is that the expert's action policy is highly sensitive to his information, the disadvantage is that the policy is optimal from the expert's perspective, not the decision-maker's. Imagine now that the decision-maker keeps formal authority so that the expert has to report back to her. If the expert followed the same information acquisition strategy, then the decision-maker would discount the expert's advice, resulting in losses from strategic communication. The expert can afford to incur these losses if they are limited. However, if the losses from strategic communication are severe, then the expert has an incentive to change his information acquisition strategy in a way that reduces these losses. The optimal information acquisition strategy has the following features in this case. Firstly, the expert acquires perfect information about the preferred action from the decision-maker's perspective. Secondly, he acquires only a noisy signal about the optimal action from his perspective. That is, he remains partially ignorant about his own preferred choice, to signal credibly to the decision-maker that his advice is useful to her. The expert benefits only to some extent from the resulting action policy. However, this is still better than perfectly knowing the ideal action but not getting the decision-maker to follow the advice.

Put differently, an expert who wishes to have an impact on the decision-maker's choices needs to acquire information of primary concern to the decision-maker and needs to reassure the decision-maker of his unbiasedness. Our analysis reveals that advice based on privately optimal information is the less effective the more likely is extreme disagreement between expert and decision-maker. If extreme disagreement is likely, communication based on privately optimal information of the expert has very little impact on the decision-maker's choice and hence does not work well. The shadow of such ineffective communication makes the expert avoid these kind of situations and gives him incentives to acquire information that impacts decisions.

We draw upon and contribute to several literatures. Our first contribution is to introduce a rich model of information and strategic information transmission à la Crawford and Sobel (1982) into a problem of adapting to new information. The defining feature of the problem is that there is no conflict of interest with respect to the status quo; conflicts arise only ex post depending on the information that is acquired. In contrast to known models, conflicts are endogenous here and intertwined with information. Our linear model allows for closed form expressions of the value of information in a rich set of environments that we characterize. The technique to compute these closed forms is new to the literature.[3] Likewise is the statistical model, that allows for linear updating in a more tractable way than the multivariate normal case allows.[4] The model is rich enough to allow us to quantify the effectiveness of strategic communication, a measure of the amount of information transmitted through strategic communication. Making this kind of comparative statics analysis feasible is perhaps the major contribution of this paper.

Adapting to news is a natural application, but our approach is not confined to such problems; different scenarios give rise to the same abstract incentive problems. For example, think of a situation where incentive contracts have been used to align incentives with respect to everything that is known already. Information arriving after this contracting stage still creates conflicts, e.g., when a project- or division-manager's pay depends on the division's profit to a greater extent than overall profits do. Likewise, in the financial industry, even if regulation makes every effort to eliminate known conflicts of interests, requesting perfect foresight is probably asking too much. Similarly, differences in the lengths of horizons may create wedges of the sort envisioned here. For example, a consultant will care relatively more

[3]See Goltsman et al. (2009) and Alonso and Rantakari (2013) for contributions that characterize values of communication via mediation.

[4]The idea is to construct the multivariate distribution from marginals with linear inference rules based on truncations to the tails – linear tail conditional expectations. In addition, the classical linear conditional expectations rules apply. Our leading case is the joint Laplace distribution (Kotz et al. (2001)). However, we describe the entire class with these features. See Section 8 for details.

about the short-term impact of his advice than the advised firm does.[5]

Our paper adds to the comparison of institutions. Dessein (2002) investigates the optimal allocation of authority in a decision-problem à la Crawford and Sobel (1982) in which the informed party is uniformly biased in one direction. Delegation entails a loss of control, communication a loss from strategic information transmission. Delegation outperforms communication if the bias is small. We look at the optimal allocation of authority in problems of acquiring and adapting to new information, or more generally in problems where a priori known biases have been eliminated.[6]

The comparison of institutions has implications for the organization of hierarchies. The literature has studied the interplay between adaptation and coordination problems (see Alonso et al. (2008) and Rantakari (2008)). In particular, Alonso et al. (2008) show that decentralized decision-making is better than centralized decision-making for small conflicts of interests and the reverse is true for larger conflicts. Our present approach abstracts from the coordination motive and shows that information acquisition may tilt the dice in favor of communication. Our results have parallels in richer hierarchies. In companion work, (Deimen and Szalay (2018)), we allow for information provision in an organization with division of labor and show that the optimal information provision by headquarters aligns incentives. Thus, the informational policy of the organization may serve as a substitute for the allocation of authority.

We contribute to the literature on information acquisition in communication problems.[7] Most closely related in terms of conclusions is Argenziano et al. (2016), which allows for endogenous information acquisition in the Crawford and Sobel (1982) model and shows that communication creates better incentives for information acquisition than delegation. Appropriate off-path beliefs of the receiver can make life particularly unpleasant for a sender

[5] See Antic and Persico (2017), among other results, for an analysis of conflicts due to differences in horizons.

[6] Communication in adaptation problems with similar (linear) reduced forms have been studied, e.g., by Melumad and Shibano (1991) and Stein (1989). The most general analysis is due to Gordon (2010). Our contribution is the rich informational model. For tractability, we assume quadratic loss functions, which is more structure than Dessein (2002) imposes.

[7] See Austen-Smith (1994) for an early contribution. See also Pei (2015) for a recent contribution. The effects of better information are studied in Moscarini (2007) and Ottaviani and Sørensen (2006). Blume et al. (2007) study noise in communication.

who acquires too little information relative to what the receiver expects, even if the receiver expects more from the sender than what she would do herself.[8] The most important difference to our problem is that we address a very different kind of decision-problem, with conflicts that arise only ex post but are absent ex ante. A further difference is the statistical model we employ: we face a two-dimensional state space[9] and develop a general theory of location experiments with linear posteriors. While more information acquisition is induced in Argenziano et al. (2016), it is information that correlates better with the decision-maker's preferred choice that is acquired here.

Clearly, our analysis has its limitations. In the incomplete contracting approach, actions are not contractible. If they were, then much more complicated institutions, in particular optimally constrained delegation, would become feasible.[10] In the context of information acquisition, a problem of this sort is analyzed in Szalay (2005). The optimal way to deal with a problem of moral hazard in information acquisition is to prohibit actions that are optimal given prior information. Allowing for costs and contractible actions is an extension worth pursuing. We stick to the case of overt and costless information acquisition here, as in Kamenica and Gentzkow (2011). This is a reasonable description of an inhouse consultant who is to combine his knowledge with the data owned by the firm. It is very easy to monitor which files the consultant requests and which not. The crucial assumption is that the expert can somehow generate this kind of transparency. With covert information acquisition the expert would have no incentive to adjust his information acquisition and communication would perform badly. Thus, if he can, the expert wants to choose the overt mode.[11]

The paper is structured as follows. In Section 2, we introduce the model. In Section 3, we prove an essential reduction that serves to simplify the analysis dramatically. We

[8]See Eső and Szalay (2017) for an argument showing that the nuancedness of language reduces the incentive to become informed for related reasons.

[9]For multi-dimensional models, see Battaglini (2002), Chakraborty and Harbaugh (2007), and Levy and Razin (2007). Our aggregation model, with a two-dimensional state and a one-dimensional action space has not yet been studied. Our approach differs also from the one taken in random bias models, such as Li and Madarász (2008), Dimitrakas and Sarafidis (2005), and Morgan and Stocken (2003), which typically look at independent biases.

[10]For an analysis of such institutions, see Holmstrom (1982), Alonso and Matouschek (2008), and Amador and Bagwell (2013).

[11]We thank Steve Matthews for pointing this out to us.

characterize equilibria at the communication stage for arbitrary information in Section 4 and derive the value of information arising from communication in Section 5. Proceeding backwards along the timeline, we study the expert's incentives to acquire information in Section 6 and then draw the implications for the choice of institutions in Section 7. To this point, our analysis is confined to our leading case, the joint Laplace distribution. In Section 8, we generalize our findings to a rich class of informational models with linear updating rules, all ordered by a single parameter that captures the effectiveness of communication. Section 9 looks at a variation in timing. Section 10 discusses a number of extensions and concludes our investigation. All longer proofs are gathered in appendices, proofs of theorems are discussed in the text.

2 The Model

2.1 The decision-problem

A decision-maker, henceforth the receiver, 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 receiver is

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

Unfortunately, the receiver 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

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

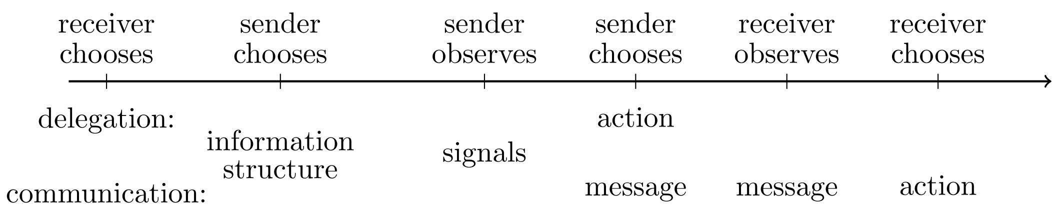
where η is the realization of a random variable that is positively correlated with ω. We assume that ω as well as η have a mean of zero, so that the sender and the receiver agree that the status quo action, $y = 0$, is optimal absent additional information. Moreover, we assume identical variances $Var(\omega) = Var(\eta) = \sigma^2$ and denote the covariance by $Cov(\omega, \eta) = \sigma_{\omega\eta}$. The coefficient of correlation is denoted by $\rho \equiv \frac{\sigma_{\omega\eta}}{\sigma^2} \in (0, 1)$.[12]

[12] Our analysis easily extends to not too asymmetric prior variances with $\min\{Var(\omega), Var(\eta)\} \geq \sigma_{\omega\eta}$. The model is interesting only if $\rho > 0$, because no meaningful communication is possible for $\rho \leq 0$.

The sender does not know the states ω and η either. However, he can observe information about the realized states. In particular, the sender's information consists of the noisy signals $s_\omega = \omega + \varepsilon_\omega$ and $s_\eta = \eta + \varepsilon_\eta$, where ε_ω and ε_η are uncorrelated noise terms. Let $\boldsymbol{\tau} \equiv (\omega, \eta, \varepsilon_\omega, \varepsilon_\eta)$ and let $\boldsymbol{\Sigma}$ denote the covariance matrix of $\boldsymbol{\tau}$. To make our model tractable, we assume that $\boldsymbol{\tau}$ follows a joint Laplace distribution, i.e., we assume that the characteristic function of the distribution is $\Phi(\mathbf{t}) = \frac{1}{1+\frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}$ (see Kotz et al. (2001)). We explain in detail why this assumption is useful in Section 3 and generalize our analysis beyond the Laplace case in Section 8.

2.2 Timing

Timeline:



The strategic interaction unfolds as indicated in the timeline.[13] Firstly, the receiver commits to an institution of decision-making, $d \in \mathcal{D} = \{delegation, communication\}$. If she chooses *delegation*, then she delegates both information acquisition and decision-making to the sender. If she chooses *communication*, then she retains the right to choose y herself and only delegates information acquisition to the sender. Note that the receiver is always forced to delegate information acquisition to the sender, because she has no time to acquire information herself. Secondly, the sender chooses what information to acquire. Formally, the sender chooses the variances of the noise terms in the signals, $Var(\varepsilon_\omega) = \sigma^2_{\varepsilon_\omega}$ and $Var(\varepsilon_\eta) = \sigma^2_{\varepsilon_\eta}$. 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

[13] See Section 9 for a discussion of a variation with an alternative timing.

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 Strategies, beliefs, and equilibria

A sender strategy consists of two parts. Firstly, for a given institution of decision-making, $d \in \mathcal{D}$, the sender chooses a feasible information structure; formally, he chooses the variances $(\sigma^2_{\varepsilon_\omega}, \sigma^2_{\varepsilon_\eta})$ in the covariance matrix $\mathbf{\Sigma}$ of the joint Laplace distribution of $\boldsymbol{\tau}$. Secondly, given $d =$ *communication*, given the information structure Σ, and given a signal realization $(s_\omega, s_\eta) \in \mathbb{R}^2$, the sender chooses what message $m \in \mathbb{M}$ to send. Formally, a pure sender strategy is a pair of functions $\mathcal{D} \to \mathbb{R}^2_+$, $d \mapsto (\sigma^2_{\varepsilon_\omega}, \sigma^2_{\varepsilon_\eta})$ and $M : \mathbb{R}^2 \times \mathbb{R}^2_+ \to \mathbb{M}$, $(s_\omega, s_\eta, \sigma^2_{\varepsilon_\omega}, \sigma^2_{\varepsilon_\eta}) \mapsto m$. A mixed sender strategy is a probability distribution over the pure strategies. The message space is sufficiently rich; we do not impose any restrictions on $\mathbb{M}$. Given $d =$ *delegation*, the latter part of the sender's strategy is replaced by an optimal action policy for each given information structure and signal realization, $Y : \mathbb{R}^2 \times \mathbb{R}^2_+ \to \mathbb{R}$, $(s_\omega, s_\eta, \sigma^2_{\varepsilon_\omega}, \sigma^2_{\varepsilon_\eta}) \mapsto y$.

A receiver strategy consists of the choice of institution, $d \in \mathcal{D}$, and, for $d =$ *communication*, a mapping from information structures and messages into actions, $Y : \mathbb{M} \times \mathbb{R}^2_+ \to \mathbb{R}$, $(m, \sigma^2_{\varepsilon_\omega}, \sigma^2_{\varepsilon_\eta}) \mapsto y$. As is well known, the receiver never mixes, due to the strict concavity of her payoff function.

There is commitment to $\mathcal{D}$ and Σ, but no commitment in the communication game. A Bayesian equilibrium of our game corresponds to the standard notion. For each observed information structure and each message, the receiver forms a belief as to what type sent the message. The belief is derived from the prior and the sender's strategy. The receiver's equilibrium strategy maximizes her payoff given her belief and the sender's equilibrium strategy. Likewise, the sender's choice of information structure and his message strategy maximize his payoff given the receiver's strategy.

When analyzing the game we focus on the most informative equilibria for all possible information structures, that is, on and off equilibrium path. In particular, this implies that we do not allow for strategies where the receiver can threaten not to listen to the sender if the latter does not choose the receiver's preferred information structure. We find this assumption reasonable in situations where the receiver has to justify her actions ex post to some third party. For example, a CEO may have to explain to the members of the board of directors why she took certain actions and what information she had when she made decisions.[14]

We begin with an analysis of the sender's and the receiver's ideal choices as a function of the information.

3 Key Properties of the Informational Environment

3.1 A useful reduction

Since the quadratic loss function is maximized at the conditional mean, the sender's ideal choice after observing (s_ω, s_η) is given by $y^s(s_\omega, s_\eta) = \mathbb{E}[\eta|\, s_\omega, s_\eta]$. Moreover, all sender types with the same conditional mean have exactly the same posterior distribution. This observation makes the following lemma obvious.

Lemma 1 *Any equilibrium under communication is essentially equivalent to one where the sender's message strategy is a function of* $\mathbb{E}[\eta|s_\omega, s_\eta]$ *only and all sender types* (s_ω, s_η) *such that* $\mathbb{E}[\eta|s_\omega, s_\eta]$ *=constant induce the same action.*

[14]Argenziano et al. (2016) allow for the case where the receiver can threaten with babbling and find that driven by this threat the sender will overinvest in information acquisition. We rule out such threats here. With a slight variation of our model – decomposing information further into public and private components – we can account for any kind of off-path threats.

All sender types whose signals aggregate to the same value have the same preferences. More precisely, such sender types share the same posterior and hence the same ideal policy. Moreover, 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 the level of $\mathbb{E}[\eta|s_\omega, s_\eta]$ from the sender. In fact, any equilibrium is essentially equivalent to one with communication about this level only. Before the signals are realized, the level that the conditional expectation takes is random and we denote this random variable by

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

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.

3.2 Information in the reduced form

Given Lemma 1, we can focus on a reduced form model corresponding to the joint distribution over the receiver's and the sender's ideal actions ω, η, and the aggregated signal θ, that indicates the sender's conditionally optimal action. For the Laplace distribution, the conditional expectation is a linear function of the signals, $\mathbb{E}[\eta|s_\omega, s_\eta] = \gamma_\omega s_\omega + \gamma_\eta s_\eta$ for some constants $\gamma_\omega, \gamma_\eta$ that are provided in the proof of the following Lemma. This feature allows us to compute the joint distribution over the random vector (ω, η, θ):

Lemma 2 *i) For any given* $\left(\sigma^2_{\varepsilon_\omega}, \sigma^2_{\varepsilon_\eta}\right) \in \mathbb{R}^2_+$, *the vector of random variables* (ω, η, θ) *follows a joint Laplace distribution with first moments* $\mathbb{E}[\omega] = \mathbb{E}[\eta] = \mathbb{E}[\theta] = 0$ *and second moments*

$$Cov(\eta, \theta) \equiv \sigma_{\eta\theta} = Var(\theta) \equiv \sigma^2_\theta = \sigma^2 \frac{\frac{\sigma^2_{\varepsilon_\omega}}{\sigma^2} + \frac{\sigma^2_{\varepsilon_\eta}}{\sigma^2}\rho^2 + 1 - \rho^2}{\left(1 + \frac{\sigma^2_{\varepsilon_\omega}}{\sigma^2}\right)\left(1 + \frac{\sigma^2_{\varepsilon_\eta}}{\sigma^2}\right) - \rho^2}, \tag{1}$$

and

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

ii) Moreover, a joint distribution of (ω,η,θ) *can be generated through Bayesian updating from signals* (s_ω,s_η) *if and only if* $\sigma_{\omega\theta}\in[0,\sigma_{\omega\eta}]$ *and for any given* $\sigma_{\omega\theta}=C$, $\sigma_\theta^2\in\left[\rho C,\frac{1}{\rho}C\right]$.

Given the linearity of the conditional expectation function, the precise expressions in (1) and (2) are straightforward to compute.[15] The equality $\sigma_{\eta\theta}=\sigma_\theta^2$ in (1) follows from the fact that θ is the conditional expectation of η given the signals. This construction reduces the degrees of freedom in the covariance matrix. The lemma also describes all the feasible covariance matrices that can be generated from the underlying joint distribution by the sender's Bayesian updating. Note that this set is smaller than the set of positive semidefinite matrices. For future reference, we denote the set of feasible joint distributions Γ and depict it in Figure 1.

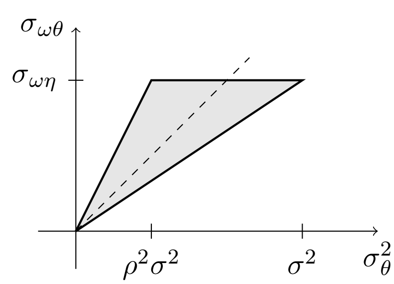


Figure 1: The set Γ of feasible second moments of the joint distribution of (ω,η,θ).

The extreme points of the feasible set are easy to understand. The origin corresponds to completely noisy information. The top right vertex corresponds to the case where the sender gets signals with $\sigma_{\varepsilon_\eta}^2=0$ (and arbitrary $\sigma_{\varepsilon_\omega}^2$) and hence knows his state, η. The top left vertex corresponds to the case where the sender gets signals with $\sigma_{\varepsilon_\omega}^2=0$ and $\sigma_{\varepsilon_\eta}^2\to\infty$. More generally, the top edge is generated from signals featuring $\sigma_{\varepsilon_\omega}^2=0$ and some $\sigma_{\varepsilon_\eta}^2\geq 0$. On this edge, $\sigma_{\omega\theta}$ reaches its maximum value, $\sigma_{\omega\eta}=\rho\sigma^2$.

[15]For example, $Var\left(\theta\right)=Var\left(\gamma_\omega s_\omega+\gamma_\eta s_\eta\right)=\gamma_\omega^2Var\left(s_\omega\right)+\gamma_\eta^2Var\left(s_\eta\right)+2\gamma_\omega\gamma_\eta Cov\left(s_\omega,s_\eta\right).$

Three further properties of the Laplace distribution contribute to making our model tractable.

Lemma 3 *The joint distribution of (ω, η, θ) features linear conditional expectations. In particular,*
i) the conditional expectation of ω given θ is the linear regression of ω on θ,

$$\mathbb{E}\left[\omega|\,\theta\right] = \frac{\sigma_{\omega\theta}}{\sigma_\theta^2} \cdot \theta \qquad \forall\theta \in \mathbb{R}, \tag{3}$$

ii)

$$\mathbb{E}\left[\omega|\,\theta \in \left[\underline{\theta}, \overline{\theta}\right]\right] = \frac{\sigma_{\omega\theta}}{\sigma_\theta^2} \cdot \mathbb{E}\left[\theta|\,\theta \in \left[\underline{\theta}, \overline{\theta}\right]\right] \qquad \forall\underline{\theta} \leq \overline{\theta}, \tag{4}$$

iii) the marginal distribution of θ is a one-dimensional Laplace distribution, and the tail conditional expectation of θ satisfies

$$\mathbb{E}\left[\theta|\,\theta \geq \bar{\theta}\right] = \mathbb{E}\left[\theta|\,\theta \geq 0\right] + \alpha \cdot \bar{\theta} \qquad \forall\bar{\theta} \geq 0 \tag{5}$$

with $\alpha = 1$.

All the results of this subsection are not limited to the Laplace distribution. In other words, our assumptions are sufficient but not necessary for our results. In fact, all we need is that the reduced form model has the features described in Lemma 2 and the linearity properties as in Lemma 3. The proofs of Lemma 2 and of Parts i) and ii) of Lemma 3 only use the fact that the Laplace distribution is an elliptically contoured distribution (see, e.g., Fang et al. (1990)), hence the results are valid for this entire class.[16] The only feature that is not shared by all elliptically contoured distributions is the linearity of the tail conditional expectation, (5) , a feature we need to compute closed form values of communication in what follows. We stick to the special case of the Laplace distribution with $\alpha = 1$ for now, to keep the exposition simple. However, we generalize our results in Section 8 to the more general case where $\alpha \neq 1$.

[16]See Mailath and Nöldeke (2008) for a model using elliptically contoured distributions in information economics.

3.3 Information and biases

Information is useful to improve decision-making. Depending on the choice of information structure, the aggregate signal θ is relatively more or less useful to the receiver than to the sender. Clearly, the relative usefulness of information shapes the ideal decision-rules if given access to the aggregate signal directly. Since the aggregate signal is privately observed by the sender only, the choice of information structure endogenously creates or resolves biases in decision-making between the sender and the receiver. These are the main points we explain in this subsection.

3.3.1 Decision rules with a public aggregate signal

If given access to information θ, the receiver's ideal policy is, by Lemma 3, $y^r(\theta) = \mathbb{E}[\omega|\,\theta] = \frac{\sigma_{\omega\theta}}{\sigma_\theta^2} \cdot \theta$, while the sender's ideal policy is, by construction, $y^s(\theta) = \theta$. Define the regression coefficient

$$c \equiv \frac{\sigma_{\omega\theta}}{\sigma_\theta^2}.$$

Information is the more useful from a player's perspective the higher the covariance with the player's ideal action. As proven in Lemma 2, σ_θ^2 is identically equal to $\sigma_{\eta\theta}$. Hence, compared to the sender's, the receiver's ideal policy is less (more) responsive to information θ, $c < (>)\, 1$, if θ is less (more) informative about ω than about η.

To illustrate, consider the two extreme cases. Firstly, suppose the sender observes η without noise. Since η is all the sender is interested in, θ is identically equal to η so that $\sigma_\theta^2 = \sigma^2$ and $\sigma_{\omega\theta} = \sigma_{\omega\eta}$. This information structure corresponds to the top right corner of the set Γ in Figure 1 with $c = \frac{\sigma_{\omega\eta}}{\sigma^2} = \rho < 1$ equal to the slope of the 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 θ than the sender would want her to respond. The situation is depicted in the left panel of Figure 2.

Secondly, suppose the sender observes ω without noise and the signal s_η contains an infinite amount of noise implying that $\theta = \mathbb{E}[\eta|\,\omega] = \rho\cdot\omega$ so that $\sigma_\theta^2 = \rho^2\cdot\sigma^2$ and $\sigma_{\omega\theta} = \rho\cdot\sigma^2$. This information structure corresponds to the top left corner in Figure 1 with $c = \frac{1}{\rho} > 1$. 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.

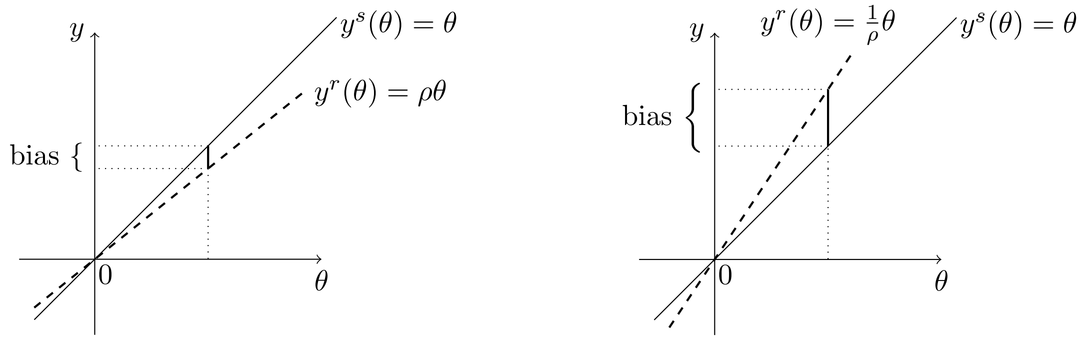


Figure 2: Ideal choices as a function of the underlying information.

Any relative usefulness in between these extreme cases can result from choosing some feasible information structure. Formally, any pair of moments $(\sigma_\theta^2, \sigma_{\omega\theta})$ in the set Γ in Figure 1, gives rise to a regression coefficient for the receiver satisfying $c \in \left[\rho, \frac{1}{\rho}\right]$. Even though all information structures in Γ on a ray with given slope are equal in relative usefulness, they differ in absolute usefulness: pairs $(\sigma_\theta^2, \sigma_{\omega\theta})$ farther to the northeast are more informative for both players.

3.3.2 Information and endogenous biases

Since θ is privately observed by the sender, the relative usefulness of information incentivizes the sender to misrepresent his information in different ways. As is customary in the literature, define the bias $b(\theta) \equiv (1-c) \cdot \theta$. If $c < 1$, then $b(\theta) > (<) \, 0$ for $\theta > (<) \, 0$ and the sender has an incentive to exaggerate positive realizations and downplay negative realizations. In contrast, if $c > 1$, then the sign of the bias is reversed and the sender has an incentive to downplay the positive realizations and exaggerate the negative ones. In both cases, the sender and receiver agree on the optimal action for $\theta = 0$. They agree on the optimal policy for all θ if and only if $c = 1$.

4 Equilibria in the Communication Game

We now investigate equilibria in the communication game when the sender observes signals with a given amount of noise. Due to Lemma 1 and the single crossing condition of utilities in θ and y, any equilibrium is essentially equivalent to an interval partition on $\mathbb{R}$ for $c \neq 1$, inducing a countable number of distinct receiver actions. For $c = 1$, there is also an equilibrium inducing an uncountably infinite number of receiver actions (for details see discussion after Theorem 2). Our results are in line with the literature (Gordon (2010)). However, since we work from different assumptions, we have to prove everything from scratch. We state only our main results here and refer to Appendix B for details. As standard, partitional equilibria are characterized by indifferent sender types $a_i^n \equiv a_i(n)$ with n relating to the number of induced receiver actions. We let a_1^n denote the first marginal type above zero.

Proposition 1 *Suppose that $c \leq 1$.*
i) For all n, there exists a unique equilibrium, which is symmetric and induces $2(n+1)$ actions (Class I) and a unique equilibrium, which is symmetric and induces $2n+1$ actions (Class II).
ii) For $n \to \infty$, the limits of the finite Class I and Class II equilibria exist and correspond to infinite equilibria of the communication game.
iii) Within any of the two classes of equilibria, the sequence of first thresholds above zero $(a_1^n)_n$ satisfies $\lim_{n\to\infty} a_1^n = 0$.

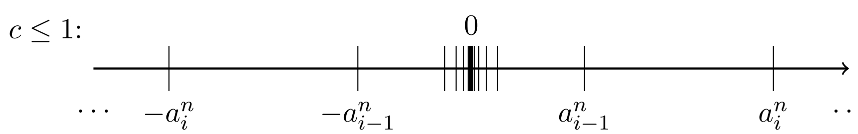


Figure 3: Intervals around the agreement point $\theta = 0$ get arbitrarily short as $n \to \infty$.

For future reference, we denote the limits of the finite equilibria as *limit equilibria*.

Proposition 2 *Suppose that $c > 1$. Then, in any equilibrium, the first threshold below or the first threshold above zero is bounded away from zero and at most a finite number of receiver actions is induced in equilibrium.*

The important insight to take away is that communication is arbitrarily precise around the agreement point, $\theta = 0$, in case $c \leq 1$, and coarse in case $c > 1$.

5 The Value of Communication

Let μ and v^2 denote the receiver's posterior mean and variance after receiving the sender's message. We define $\mu_i^n \equiv \mathbb{E}\left[\theta | \, \theta \in \left[a_{i-1}^n, a_i^n\right)\right]$ for $i = 1, \ldots, n$ and $\mu_{n+1}^n \equiv \mathbb{E}\left[\theta | \, \theta \geq a_n^n\right]$. Hence, for example 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 supported on $\left\{\mu_{-(n+1)}^n, \mu_{-n}^n, \ldots, \mu_{-1}^n, \mu_1^n, \ldots, \mu_n^n, \mu_{n+1}^n\right\}$, and likewise for the conditional variance. By Lemma 3 ii), the receiver's action is c times the appropriate mean. The distributions of the random variables μ and v^2 are derived from the marginal distribution of θ.

Consider now the expected utilities of sender and receiver. Suppose first the sender would naïvely transmit information θ honestly to the receiver and the receiver would follow the optimal policy to choose $y\,(\theta) = c \cdot \theta$. In this case, the receiver's expected utility would be $\mathbb{E}[-(c\theta - \omega)^2] = c^2\sigma_\theta^2 - \sigma^2$, while the sender's expected utility would be $\mathbb{E}[-(c\theta - \eta)^2] = c\,(2-c)\,\sigma_\theta^2 - \sigma^2$. The following Lemma shows that expected utilities under strategic communication can be computed using the same functional form with the amount of information transmitted equal to $\mathbb{E}\left[\mu^2\right] \leq \sigma_\theta^2$.

Lemma 4 *The receiver's expected equilibrium utility is*

$$\mathbb{E}u^r\,(c\mu, \omega) = c^2\mathbb{E}\left[\mu^2\right] - \sigma^2. \tag{6}$$

The sender's expected equilibrium utility is

$$\mathbb{E}u^s\,(c\mu, \eta) = c\,(2-c)\,\mathbb{E}\left[\mu^2\right] - \sigma^2. \tag{7}$$

By a standard variance decomposition, $\mathbb{E}\left[\mu^2\right] = \sigma_\theta^2 - \mathbb{E}\left[v^2\right]$, 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 the naïve sender scenario where the sender is kind enough to communicate θ truthfully, or equivalently, θ is publicly observable. In that case, the receiver would identify μ with θ, so that $\mathbb{E}[\mu^2] = \sigma_\theta^2$. 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 the expected conditional variance, $\mathbb{E}[v^2]$.

We denote $\mu_+ \equiv \mathbb{E}[\theta | \theta \geq 0]$. The following proposition shows that the equilibrium variability of choices has a convenient representation. This representation is at the heart of the analysis and enables us to solve for the value of communication.

Proposition 3 *The equilibrium variability of the receiver's posterior mean in a Class I equilibrium inducing $2(n+1)$ distinct receiver actions is given by*

$$\mathbb{E}\left[\mu^2\right] = \frac{2}{2-c}\mu_+^2 - \frac{c}{2-c}\left(\mu_1^n\right)^2. \tag{8}$$

In a Class II equilibrium inducing $2n+1$ distinct receiver actions, the equilibrium variability is

$$\mathbb{E}\left[\mu^2\right] = \left(1 - \Pr\left[\theta \in \left[-\frac{c\mu_2^n}{2}, \frac{c\mu_2^n}{2}\right)\right]\right) \cdot \left(\frac{2}{2-c}\mu_+^2 + \frac{c}{2-c}\mu_2^n\mu_+\right). \tag{9}$$

In a limit equilibrium, which exists if and only if $c \leq 1$,

$$\mathbb{E}\left[\mu^2\right] = \frac{2}{2-c}\mu_+^2. \tag{10}$$

For the Laplace distribution, $2\mu_+^2 = \sigma_\theta^2$ and moreover, for any $c \leq 1$, $\mathbb{E}[\mu^2]$ is maximal in a limit equilibrium.

The proposition provides a closed form solution for the equilibrium variability of the receiver's choices in any Class I and Class II equilibrium, and for a limit equilibrium. This is possible, because the expected variability, $\mathbb{E}[\mu^2]$, can be expressed in terms of the constant μ_+ and the receiver's conditional expectation in the first interval partition, μ_1^n for Class I (8) and μ_2^n for Class II (9). In a limit equilibrium the first partition length converges to zero and as a result the expected variability depends on the bias c only (10).

To get a feeling for how this works, consider a Class I equilibrium with $n = 1$ and four intervals in total. Threshold a_1^1 partitions $\mathbb{R}^+$ into two subintervals with conditional means $\mu_1^1 = \mathbb{E}[\theta | \theta \in [0, a_1^1)]$ and $\mu_2^1 = \mathbb{E}[\theta | \theta \geq a_1^1]$, respectively. The Laplace distribution satisfies $\mathbb{E}[\theta | \theta \geq a_1^1] = \mu_+ + a_1^1$ (Lemma 3 iii) with $\alpha = 1$). The indifference condition of the marginal sender type, a_1^1, is $a_1^1 - c \cdot \mu_1^1 = c \cdot \mu_2^1 - a_1^1$. Together, this makes for a linear relationship of the three moments

$$\mu_2^1 - \mu_+ = \frac{c}{2-c} \cdot \left(\mu_1^1 + \mu_+\right),$$

which allows us to eliminate one of them. In addition, the Laplace distribution is a one-parameter distribution, so we can write its variance as a function of the scale parameter, μ_+. In particular, we have $\sigma_\theta^2 = 2\mu_+^2$. Finally, letting $p_1^1 \equiv \Pr\left[\theta \in [0, a_1^1] | \, \theta \geq 0\right]$, and using the law of iterated expectations, $p_1^1 \mu_1^1 + (1 - p_1^1)\, \mu_2^1 = \mu_+$, to substitute for the probability distribution as a function of the truncated means, we obtain

$$\sum_{i=1}^{2} p_i^1 (\mu_i^1 - \mu_+)^2 = \left(\mu_2^1 - \mu_+\right)\left(\mu_+ - \mu_1^1\right).$$

Using these insights, we observe that the expected variability of choices is exactly (8) with $n = 1$. The argument for Class II is essentially the same.

The general proof is based on an induction argument. In a first step, the expected variability of choices is computed over the last two subintervals, conditional on $\theta \geq a_{n-1}^n$, the second to last threshold in an equilibrium interval partition. Due to the aforementioned linearity, this is a function of μ_n^n and a_{n-1}^n only. Using the indifference condition of the marginal type a_{k-1}^n, the inductive proof then shows that the expected variability conditional on $\theta \geq a_{k-2}^n$ takes the same functional form as the one conditional on $\theta \geq a_{n-1}^n$, for all $k < n$.

The derivation of expressions (8) through (10) depends on the linearity of the tail conditional expectation. Even though the proposition is stated for the Laplace case ($\alpha = 1$), in the proof we allow for more general values of α. Hence we can state closed form solutions for the value of communication in a class of general linear environments (see Section 8).

6 Information Acquisition

When choosing the information structure, the sender takes its effects on equilibrium communication into account. We assume that sender and receiver coordinate on an equilibrium that gives them the highest possible expected utility in the set of all equilibria. If we select such an equilibrium, then we say that $\mathbb{E}\left[\mu^2\right]$ is maximal in the equilibrium set. The sender's problem is

$$\max_{\sigma_{\omega\theta},\sigma_\theta^2} c\,(2-c)\,\mathbb{E}\left[\mu^2\right]-\sigma^2 \tag{11}$$
$$s.t.\ \sigma_{\omega\theta},\sigma_\theta^2\in\Gamma \text{ and}$$
$$\mathbb{E}\left[\mu^2\right] \text{ is maximal in the equilibrium set.}$$

In face of Propositions 1, 2, and 3, the solution to problem (11) is as follows:

Theorem 1 *The set of optimal information structures from the sender's perspective is given by* $\sigma_{\omega\theta}=\sigma_{\omega\eta}$ *and* $\sigma_\theta^2\in\left[\sigma_{\omega\eta},\sigma^2\right]$.

The proof of the theorem is a straightforward combination of the preceding propositions. For any $(\sigma_{\omega\theta},\sigma_\theta^2)$ such that $c=\frac{\sigma_{\omega\theta}}{\sigma_\theta^2}\leq 1$, there exists for each n a unique Class I and a unique Class II equilibrium and no other equilibrium. The variability of choices in either class converges to $\mathbb{E}\left[\mu^2\right]=\frac{1}{2-c}\sigma_\theta^2$ from below as n goes out of bounds. Thus, for any feasible information structure featuring $c\leq 1$, the highest expected utility of the sender is equal to

$$\mathbb{E}u^s\left(c\mu,\eta\right)=\max_{\substack{\sigma_{\omega\theta},\sigma_\theta^2\\ s.t.\ \sigma_{\omega\theta},\sigma_\theta^2\in\Gamma,\ c\leq 1}} c\,(2-c)\,\sigma_\theta^2\frac{1}{2-c}-\sigma^2=\sigma_{\omega\eta}-\sigma^2.$$

The term $c\,(2-c)\,\sigma_\theta^2$ corresponds to the intrinsic value of the signal θ to the sender if it is publicly observed and the receiver follows the policy $y^r\left(\theta\right)=c\cdot\theta$. Because θ isn't publicly observed, only a fraction $\frac{1}{2-c}$ of the intrinsic value materializes and the remainder is lost in strategic communication. As a result of the combination of these forces, the sender becomes indifferent between all information structures stated in Theorem 1. For an illustration see the solid line in Figure 4.

Why would the sender never want to choose any $(\sigma_{\omega\theta},\sigma_\theta^2)$ such that $c>1$, irrespective of what equilibrium would be played in the communication game following such a choice of information, if the most informative equilibrium is played on equilibrium path? Intuitively, such information structures have strategic disadvantages *and* are intrinsically suboptimal. As shown in Proposition 2, there is no equilibrium in which θ is communicated perfectly, so $\mathbb{E}\left[\mu^2\right]<\sigma_\theta^2$, implying that the sender's expected utility in any equilibrium is strictly below the value the sender obtains if information θ were public, $c\,(2-c)\,\sigma_\theta^2$. Moreover, among information structures featuring $c \geq 1$, the intrinsically most useful information structure has $\sigma_{\omega\theta} = \sigma_{\omega\eta} = \sigma_\theta^2$.

Consider now the receiver's payoff as a function of the information structure that the sender chooses. For $\sigma_{\omega\theta} = \sigma_{\omega\eta}$ and any $\sigma_\theta^2 \geq \sigma_{\omega\eta}$, the receiver's payoff in the limit equilibrium where $n \to \infty$ is

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

a decreasing function of σ_θ^2. Clearly, the receiver suffers if the sender chooses an information structure with a higher σ_θ^2; at the same time, the sender derives no benefit from such behavior. The following theorem is now obvious.

Theorem 2 *The set of sender optimal information structures contains the uniquely Pareto efficient element $\sigma_\theta^{2*} = \sigma_{\omega\theta}^* = \sigma_{\omega\eta}$. The ensuing communication continuation game following the Pareto efficient information selection has an equilibrium in which the sender communicates θ truthfully to the receiver, who follows the sender's proposal one-for-one.*

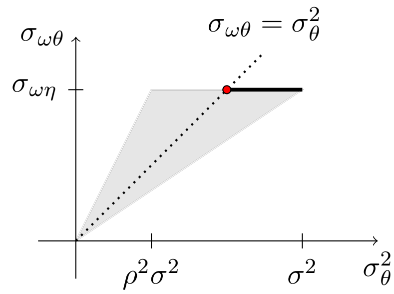


Figure 4: The solid line represents the sender-optimal information structures. The dot represents the receiver-optimal information structure within the set of sender-optimal information structures.

For convenience, we depict the theorem and the discussion preceding it graphically in Figure 4. If the sender chooses the receiver optimal information structure among those that are privately optimal for himself, then we have $c = 1$, and the bias with respect to communicating θ is eliminated. Hence, it is an equilibrium for the sender to follow the message strategy

$m(\theta) = \theta$ for all θ, and for the receiver to follow the action strategy $y(m) = \frac{\sigma^*_{\omega\theta}}{\sigma^{2*}_\theta} \cdot m = m$ for all m, because the receiver correctly identifies m with θ in her belief. We call this a *smooth communication equilibrium*, because the equilibrium involves differentiable strategies in the communication game. Note that expected utilities in the smooth communication equilibrium are the same as in the equilibrium with countably infinitely many induced actions for $c = 1$.[17] In terms of the underlying noise terms, the sender is perfectly informed about the receiver's ideal action, $\sigma^2_{\varepsilon_\omega} = 0$, but must remain partially ignorant about his preferred choice, $\sigma^2_{\varepsilon_\eta} = \frac{1-\rho^2}{\rho}\sigma^2$, to convince the receiver of his unbiasedness.

We now turn to the receiver's choice of institution of decision-making. We assume that the sender chooses the Pareto efficient information structure out of the ones that are optimal from his perspective. Since this stacks the deck in favor of communication, we give reasons beyond Pareto efficiency why our equilibrium selection is compelling after presenting our main result.

7 Delegation versus Communication

If the receiver retains the right to make choices (communication), then the sender's proposal also reflects the receiver's ideal action instead of just the sender's. The receiver's expected payoff in the smooth communication equilibrium is

$$\mathbb{E}u^r(\theta, \omega) = \sigma_{\omega\eta} - \sigma^2. \tag{12}$$

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

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

where we used the fact that $\sigma_{\eta\theta} = \sigma^2_\theta$ by construction of θ. Clearly, the optimal information structure from the sender's perspective is $\hat{\sigma}_{\omega\theta} = \sigma_{\omega\eta}$ and $\hat{\sigma}^2_\theta = \sigma^2 = \frac{1}{\rho}\sigma_{\omega\eta}$, because this information structure maximizes σ^2_θ within the set Γ. This means that the sender's action

[17] The smooth communication equilibrium is also optimal with respect to maximization of joint surplus of sender and receiver, as shown in Deimen and Szalay (2018).

reflects information about η exclusively. The receiver's expected utility under delegation is

$$\mathbb{E}u^r(\theta,\omega) = -\hat{\sigma}_\theta^2 + 2\hat{\sigma}_{\omega\theta} - \sigma^2 = \left(2 - \frac{1}{\rho}\right)\sigma_{\omega\eta} - \sigma^2. \tag{13}$$

We can now state our main result:

Theorem 3 *Suppose the sender selects privately optimal information structures for both choices of institution; in case there are several optimal ones, he picks the receiver's preferred information structure among them. Then, the receiver strictly prefers communication over delegation.*

The formal proof of the theorem consists simply of pulling insights together. In particular, direct comparison of equations (12) and (13) 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 contrast to what is known for the case of exogenously given information structures and biases.[18] 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).[19] To play devil's advocate and to reconcile results, suppose that 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, $\hat{\sigma}_{\omega\theta} = \sigma_{\omega\eta}$ and $\hat{\sigma}_\theta^2 = \frac{1}{\rho}\sigma_{\omega\eta}$. 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.

Theorem 4 *Suppose the sender selects privately optimal information structures for both choices of institution; in case there are several optimal ones, he selects the least preferred*

[18] In particular, Alonso et al. (2008) show that decentralized decision-making is better than centralized decision-making for small conflicts of interests whereas the reverse is true for larger conflicts. Dessein (2002) shows that delegation outperforms communication whenever meaningful communication is possible.

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

one from the receiver's perspective. Then, communication is strictly preferred to delegation for $\rho \in \left(0, \frac{2}{3}\right)$ *and delegation is strictly preferred for* $\rho \in \left(\frac{2}{3}, 1\right)$.

Substituting $\hat{\sigma}_{\omega\theta} = \sigma_{\omega\eta}$, $\hat{\sigma}_{\theta}^{2} = \frac{1}{\rho}\sigma_{\omega\eta}$, and $c = \rho$, the receiver's expected utility under communication is

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

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.[20]

Our model provides support for communication as an institution. However, the strength of this support depends on which information structure is selected among the optimal ones from the sender's perspective. Even though the Laplace model is simple, it does not produce a unique sender-optimal information structure, a point that we address next.

8 The Effectiveness of Biased Communication

When selecting information structures, the sender trades off intrinsic gains against strategic losses. Our result rests on the fact that the strategic losses for the sender weakly outweigh the intrinsic gains from learning too much about own interests. This dominance becomes strict when biased communication is less effective than in our leading case; the effectiveness of the communication problem is captured by the parameter α for all elliptical distributions that satisfy the linearity conditions of Lemma 3.

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

We now explain this point formally. We embed the Laplace distribution with $\alpha = 1$ into a more general family of distributions with linear tail conditional expectations. For $\alpha \in (0, 2)$, the distributions in this family have a marginal density of

$$f_x(x;\alpha) = \frac{1}{2\sigma_x}\sqrt{\frac{2\alpha^2}{(2-\alpha)}}\left(1 - \sqrt{\frac{2}{2-\alpha}}(1-\alpha)\frac{|x|}{\sigma_x}\right)^{\frac{2\alpha-1}{1-\alpha}} \cdot \mathbb{I}\left(\frac{|x|}{\sigma_x}\right), \tag{14}$$

where σ_x is the standard deviation of x and $\mathbb{I}$ is the indicator function describing the support of the distribution. For $\alpha > 1$, the support is $\mathbb{R}$; for $\alpha < 1$, the support is an interval. For $\alpha \in \left[\frac{1}{2}, 1\right)$, the interval is closed and defined by $\frac{|x|}{\sigma_x} \leq \sqrt{\frac{2-\alpha}{2}}\frac{1}{1-\alpha}$; for $\alpha < \frac{1}{2}$, this interval is open. The following Lemma states formally why the family of distributions described by (14) is of interest:

Lemma 5 *If the distribution of θ has a density given by equation* (14) *with $\alpha \in (0, 2)$, then the tail conditional expectation of θ satisfies condition* (5)*. The variance of θ is finite and related to μ_+ via $\sigma_\theta^2 = \frac{2\mu_+^2}{2-\alpha}$.*

Condition (5) can be restated as a differential equation that can be solved, resulting in the density (14). The formulation nests the classical Laplace as a special case: for $\alpha \to 1$ the density (14) converges to the density of the Laplace. Another special case worth noting is $\alpha = \frac{1}{2}$, the uniform density. So, even though (14) might be regarded as pretty special, all the known distributions that deliver closed form solutions for problems of our type have a density that satisfies (14).[21] We illustrate the density (14) for different values of α and variances equal to one in Figure 5. Note that the density has heavy tails for $\alpha > 1$.

To obtain the generalization to the multivariate distribution we proceed exactly in the same way as for the Laplace: we use (14) to compute the characteristic function of the marginal distributions. By symmetry, the characteristic function is a function of $t^2\sigma_x^2$ only, $\Phi_\alpha(t) = \phi_\alpha(t^2\sigma_x^2)$. Define the characteristic function of the multivariate distribution as $\Phi_\alpha(\mathbf{t}) = \phi_\alpha(\mathbf{t}'\mathbf{\Sigma}\mathbf{t})$, that is, take the characteristic function as invariant with respect to

[21] The density can be derived for arbitrary, positive α. However, in the general case, it can only be expressed in terms of μ_+, not the variance, since the latter substitution requires that the variance be finite. Since expected utilities are only defined for finite variance in a quadratic loss model, we restrict attention to the case $\alpha < 2$.

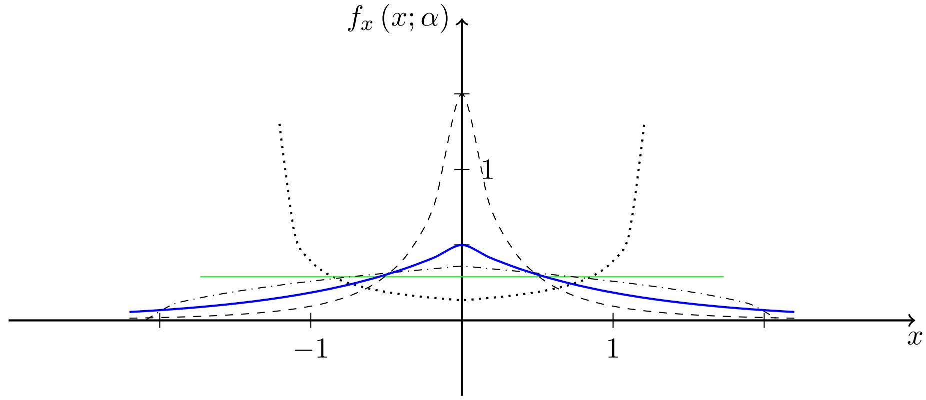


Figure 5: The density (14) depicted for $\alpha = 1.5$ (dashed); $\alpha = 1$ (solid, thick) Laplace; $\alpha = 0.6$ (dash-dotted); $\alpha = 0.5$ (solid, thin) uniform; $\alpha = 0.25$ (dotted).

changes of the dimension. By construction, the multivariate distribution is elliptical with marginal densities (14) implying that the linear conditioning rules of Lemma 3 all apply. We call this type of distribution an *elliptical distribution with linear marginal tail conditional expectations.*[22]

The following proposition states that the calculation of the equilibrium variability (Proposition 3) carries over to the more general case:

Proposition 4 *Suppose the joint distribution of* $\boldsymbol{\tau}$ *is elliptical with linear marginal tail conditional expectations and effectiveness parameter* $\alpha \in (0, 2)$. *Then, in any Class I (II) equilibrium* $\mathbb{E}[\mu^2]$ *satisfies equations* (8) *((9)), with c replaced by* αc. *Moreover, in any such equilibrium*

$$\mathbb{E}\left[\mu^2\right] \leq \frac{2-\alpha}{2-\alpha c}\sigma_\theta^2.$$

If there exists a limit equilibrium in which the sequence of thresholds $(a_1^n)_n$ *satisfies* $\lim_{n\to\infty} a_1^n = 0$, *then the upper bound on* $\mathbb{E}[\mu^2]$ *is attained.*

[22]Elliptical distributions owe their name to the fact that the level curves of their densities are elliptical. The construction via characteristic functions is standard (see, e.g., Fang et al. (1990)). However, we are not aware of any contribution in the literature that describes the subclass of elliptical distributions with linear marginal tail conditional expectations.

The linear conditioning rules allow us to get closed form solutions for the equilibrium value of communication for any problem in which equilibrium communication gets arbitrarily fine around the agreement point. The reason is that the value depends only on the product αc. If the upper bound is attained, then

$$\frac{\mathbb{E}\left[v^2\right]}{\sigma_\theta^2} = \alpha\left(1-c\right).$$

Using different methods, Alonso et al. (2008) have shown that the upper bound is attained in the uniform case.[23] Compared to the uniform case with the same variance, twice as much uncertainty is left after communication in our leading case with $\alpha = 1$.[24] Thus, the higher is α the less effective is biased communication.

We can now state our result:

Theorem 5 *Suppose that a symmetric equilibrium is played in the communication stage. Then, for* $\alpha \in (1,2)$ *the optimal information structure is unique and given by* $\sigma_\theta^{2*} = \sigma_{\omega\theta}^* = \sigma_{\omega\eta}$.

We know that the upper bound on the equilibrium variability of choices is attained for $c = 1$. To show that the optimal information structure is as stated in the theorem, it suffices to show that any information structure featuring $c < 1$ results in a lower payoff for $\alpha > 1$. Using the upper bound on payoffs, we know that the sender obtains at most a payoff of

$$\mathbb{E}u^s\left(c\mu,\eta\right) = \max_{\substack{\sigma_\theta^2,\sigma_{\omega\theta} \\ s.t.\sigma_\theta^2 \geq \sigma_{\omega\theta}}} c\left(2-c\right)\frac{2-\alpha}{2-\alpha c}\sigma_\theta^2 - \sigma^2. \tag{15}$$

For $\alpha > 1$, this value is increasing in $\sigma_{\omega\theta}$ over the set of information structures featuring $c \leq 1$. Hence, within this set, $\sigma_{\omega\theta}^* = \sigma_{\omega\eta}$. To determine whether the sender has incentives to

[23] For the uniform case, the difference equation describing equilibrium thresholds can be solved in closed form.

[24] This ratio can be computed for any equilibrium that delivers closed form values; in particular binary communication. It is easy to show that for binary communication, $\frac{\mathbb{E}\left[v^2\right]}{\sigma_\theta^2} = \frac{\alpha}{2}$. In the uniform case, binary communication eliminates 75% of the underlying uncertainty, in our leading case it eliminates 50%. In the limiting case where $\alpha \to 2$, binary communication becomes totally ineffective.

acquire more or less information about η, it suffices to understand how the ratio

$$\frac{2-c}{2-\alpha c} = \frac{2-\frac{\sigma_{\omega\eta}}{\sigma_\theta^2}}{2-\alpha\frac{\sigma_{\omega\eta}}{\sigma_\theta^2}}$$

depends on σ_θ^2. Both numerator and denominator are increasing in σ_θ^2. The increase in the numerator captures the sender's preference for intrinsically better information. The increase in the denominator captures the deterioration of equilibrium communication as the sender's information makes him more biased. If biased communication is less effective than in our leading case, the latter effect dominates, so that $\sigma_\theta^{2*} = \sigma_{\omega\eta}$. As a result, equilibrium communication about θ is conflict free. This proves the theorem, since information structures featuring $c > 1$ remain unattractive by the now familiar argument.[25]

The intuition for the result is very simple. It is always possible to communicate θ perfectly if there is no bias. In contrast, if the sender is biased then communication is the less effective the higher is α. Moreover, the marginal loss from increasing his bias is getting more pronounced the less effective the communication problem. Vice versa, if communication is very effective even in the presence of conflicts, then the sender does not mind that much that information creates conflicts, simply because he can afford to.

In sum, the receiver prefers to communicate when biased communication would be ineffective, because the shadow of such ineffective communication makes the sender acquire information that ensures unbiased communication. In other words, communication is a credible commitment to punish the sender for acquiring the wrong pieces of information. Delegation, in contrast, entails no such possibility.

9 Reversed timing

In our baseline model, the receiver chooses between the institutions of decision-making before the sender acquires information. Suppose now the receiver can choose between delegation and communication after observing what information the sender has acquired.

[25] We can also use our results to demonstrate that the threat of having to communicate may not be enough to incentivize the sender. For $\alpha < 1$, (15) is maximized for $\hat{\sigma}_{\omega\theta} = \sigma_{\omega\eta}$ and $\hat{\sigma}_\theta^2 = \frac{1}{\rho}\sigma_{\omega\eta}$. If the upper bound is attained (as it is in the uniform case) then the sender acquires his most preferred information structure and equilibrium communication features conflicts as in Alonso et al. (2008).

We find that, if interests are well aligned to begin with, then delegation is always the optimal outcome. The threat of forcing the sender to communicate thus has no bite. If interests are less well aligned, then the receiver would find it optimal to communicate with the sender if the latter acquired information about his own interests only. The sender dislikes communication and acquires information about both underlying states, to the point where the receiver becomes indifferent between communication and delegation.

Theorem 6 *Suppose the receiver chooses between communication and delegation only after the sender selects an information structure. Then, the shadow of communication partly aligns interests: for $\rho < \frac{2}{3}$, the equilibrium responsiveness of the receiver increases from $c = \rho$ to $c = \frac{2}{3}$.*

For $\rho > \frac{2}{3}$, delegation outperforms communication for any choice of information structure. Hence, in equilibrium the sender chooses his preferred information structure and the receiver delegates. For $\rho \leq \frac{2}{3}$, the receiver delegates only if the information structure satisfies $c \geq \frac{2}{3}$ and communicates otherwise. The sender has a strict preference to choose the maximally informative information structure with $c = \frac{2}{3}$ and select his preferred choice rather than choose any information structure with $c < \frac{2}{3}$ and communicate.

In equilibrium, no communication occurs. However, the shadow of communication is still helpful. The sender would face losses from communication and dislikes this. The receiver dislikes these losses as well. However, her expected utility is affected in a different way, so she can credibly threaten not to delegate if the sender acquires the wrong information structures. As in the baseline model, the threat of having to communicate with the receiver makes the sender redirect his information acquisition towards information that he would otherwise neglect, and makes him look less into things that he would otherwise look into exclusively.

10 Extensions and Conclusions

We compare two mechanisms of decision-making, delegation and communication, in a situation of delegated expertise. The expert and the decision-maker agree on the status quo but favor different actions if new information arrives. The expert chooses the precision of signals about each of the favored actions. His choice does not only impact the intrinsic usefulness of the information but also the conflicts that arise in communicating his advice. We derive a new communication model that features these endogenous state-dependent biases. Moreover, we develop a method to compute closed form expressions for the equilibrium value of information despite the fact that equilibria cannot be computed in closed form. We describe a general class of distributions for which our procedure applies. Our environment allows us to measure the amount of information that can be transmitted in equilibrium, the effectiveness of communication. We find that in environments where biased communication is ineffective, the expert chooses his information in a way that eliminates any bias in communication. Put differently, an expert who wants to be heard by the decision-maker will pay attention to things the decision-maker is interested in. This effect steers the decision-maker's choice of mechanism towards communication. The reason is that under delegation the decision-maker has no control over the expert; an expert who can choose the information and in addition can take the action will solely focus on his own interests.

Our model lends itself to many extensions, e.g. costs of information acquisition, opportunity costs of time, limited attention, simultaneous information acquisition by sender and receiver, endogenous roles of sender and receiver, and many more. We pursue some of these questions in ongoing work. We believe that the closed form expressions for the value of communication that we have obtained should prove useful in a variety of settings, for example, to study strategic communication in financial markets.

Appendix A

Proof of Lemma 1. There exists no separating equilibrium in which all sender types (s_ω, s_η) separate. The reason is that for $\rho < 1$ sender and receiver disagree on the optimal action almost surely. For quadratic utilities, players' optimal actions are given by the conditional means, $\mathbb{E}\left[\omega \mid s_\omega, s_\eta\right]$ and $\mathbb{E}\left[\eta \mid s_\omega, s_\eta\right]$, respectively. It is straightforward to show, using similar calculation as in Lemma 2, that for $\rho < 1$, we have $\mathbb{E}\left[\omega \mid s_\omega, s_\eta\right] \neq \mathbb{E}\left[\eta \mid s_\omega, s_\eta\right]$ for all $(s_\omega, s_\eta) \neq (0, 0)$.

The sender's conditionally optimal action is random before the signals are realized. Denote this random variable by $\theta \equiv \mathbb{E}\left[\eta \mid s_\omega, s_\eta\right]$. If and only if $\mathbb{E}\left[\omega \mid \theta\right] = \theta$ there exists an equilibrium where all sender types whose signals aggregate to the same level θ pool on the same strategy. Note that this equilibrium can be characterized by communication about θ only.

Consider now more generally partial pooling equilibria. Given the symmetry of the conditional distribution and the quadratic payoff, the sender's preferences over messages depend only on the distance between induced actions and θ. By strict concavity of the sender's payoff, the sender prefers actions closer to θ. We show that any equilibrium can also be characterized as one where sender types are restricted to play strategies that depend on the posterior mean θ only.

Since sender's preferences only depend on the distance to θ, we know that every set of sender types (s_ω, s_η) with $\gamma_\omega s_\omega + \gamma_\eta s_\eta = \theta$ induces at most two actions in equilibrium. Let $\underline{y}(\theta)$ denote the lower action that type θ induces and let $\overline{y}(\theta)$ denote the higher action that type θ induces. Take some particular realization $\hat{\theta}$, and suppose that the set of sender types with signal realizations $\gamma_\omega s_\omega + \gamma_\eta s_\eta = \hat{\theta}$ is divided into two subsets, S_1 and S_2, such that types in S_1 induce action $\underline{y}(\hat{\theta})$ and types in S_2 induce action $\overline{y}(\hat{\theta})$. Then, it must be that these senders are indifferent between the induced actions, $\left|\hat{\theta} - \underline{y}(\hat{\theta})\right| = \left|\overline{y}(\hat{\theta}) - \hat{\theta}\right|$. Thus, $\overline{y}(\hat{\theta}) > \hat{\theta} > \underline{y}(\hat{\theta})$. Let $\overline{y}(\hat{\theta}) - \hat{\theta} \equiv \hat{\varepsilon} > 0$.

It follows immediately that no type can induce any action $y \in \left(\hat{\theta} - \hat{\varepsilon}, \hat{\theta} + \hat{\varepsilon}\right)$. If such an action could be induced, then types with $\theta = \hat{\theta}$ could deviate and induce an action that is closer to their most preferred choice, $y = \hat{\theta}$. Likewise, all types $\theta \in \left[\hat{\theta}, \hat{\theta} + \hat{\varepsilon}\right]$ must induce action $\overline{y}(\hat{\theta})$. If some type $\tilde{\theta} \in \left[\hat{\theta}, \hat{\theta} + \hat{\varepsilon}\right]$ were supposed to induce an alternative action $\tilde{y}\left(\tilde{\theta}\right) > \hat{\theta} + \hat{\varepsilon}$ in some candidate equilibrium, then this type would prefer to induce action $\overline{y}(\hat{\theta})$ rather than $\tilde{y}\left(\tilde{\theta}\right)$. A similar argument can be given for types $\theta \in \left[\hat{\theta} - \hat{\varepsilon}, \hat{\theta}\right]$. It follows that types $\theta \in \left(\hat{\theta}, \hat{\theta} + \hat{\varepsilon}\right]$ induce a unique action $y(\theta) = \overline{y}(\hat{\theta})$ and types $\theta \in \left[\hat{\theta} - \hat{\varepsilon}, \hat{\theta}\right)$ induce a unique action $y(\theta) = \underline{y}(\hat{\theta})$. For both actions, there is a set of types who have positive mass that induce the action.

There are now two cases to distinguish: i) $\overline{y}(\hat{\theta})$ is the highest action induced in equilibrium or ii) there is $\theta^* > \hat{\theta}$ denoted as the smallest θ such that some types with $\theta = \theta^*$ induce a receiver action $\overline{y}(\theta^*)$ distinct from $\overline{y}(\hat{\theta})$. In case i), it is easy to see that the set of types that induce two actions have zero mass and are pooled with a set of types of positive mass. In case ii), indifference of types with $\theta = \theta^*$ requires by construction that $\overline{y}(\hat{\theta}) = \underline{y}(\theta^*)$. Moreover, we must have $\overline{y}(\hat{\theta}) < \theta^* < \overline{y}(\theta^*)$ and $|\theta^* - \overline{y}(\hat{\theta})| = |\overline{y}(\theta^*) - \theta^*|$. Again, for each action there is a set of types with positive mass that induce only this action. Moreover, a set of types with zero mass pools with them. Finally, note that the situation for types with $\theta = \theta^*$ corresponds exactly to the one we took as our starting point, so exactly the same arguments that have been applied to type $\theta = \hat{\theta}$ can be applied here.

Consider now the receiver's strategy. The receiver's action $\overline{y}(\hat{\theta})$ must be a best reply to the set of types that induce the action. The set of sender types who induce action $\overline{y}(\hat{\theta})$ is given by the set of signal realizations such that $\gamma_\omega s_\omega + \gamma_\eta s_\eta = \theta \in \left(\hat{\theta}, \theta^*\right)$ and subsets of sender types with $\gamma_\omega s_\omega + \gamma_\eta s_\eta = \theta \in \left\{\hat{\theta}, \theta^*\right\}$. Note that $\Pr\left[\theta \in \left(\hat{\theta}, \theta^*\right)\right] > 0$ while $\Pr\left[\theta \in \left\{\hat{\theta}, \theta^*\right\}\right] = 0$. Let all these types send some message $\hat{m}$ to induce action $\overline{y}(\hat{\theta})$. The conditional distribution of ω given $\hat{m}$ does not depend on the strategies of types $\theta \in \left\{\hat{\theta}, \theta^*\right\}$, since these types are pooled with a set of positive measure that induce only one action. Hence, the receiver's optimal action given message $\hat{m}$ does not depend on the strategies of these measure zero types either.

It follows that every equilibrium is essentially equivalent to one where senders are restricted to play strategies based on θ only. Essentially means, that the receiver's equilibrium actions, the receiver's beliefs, and expected payoffs are exactly the same; sender strategies are the same with probability one. ■

Proof of Lemma 2. **i)** The random vector $\boldsymbol{\tau} = (\omega, \eta, \varepsilon_\omega, \varepsilon_\eta)$ follows a joint Laplace distribution. Since the Laplace distribution is a member of the class of elliptically contoured distributions, the following well-known properties apply:

The sender's conditional mean θ can be calculated as $\mathbb{E}\left[\eta \mid s_\omega, s_\eta\right] = \gamma_\omega s_\omega + \gamma_\eta s_\eta$ with $\gamma_\omega = \frac{\sigma^2_{\varepsilon_\eta}\rho\sigma^2}{(\sigma^2+\sigma^2_{\varepsilon_\omega})(\sigma^2+\sigma^2_{\varepsilon_\eta})-\rho^2\sigma^4}$, and $\gamma_\eta = \frac{\sigma^2\sigma^2_{\varepsilon_\omega}-\sigma^4(1-\rho^2)}{(\sigma^2+\sigma^2_{\varepsilon_\omega})(\sigma^2+\sigma^2_{\varepsilon_\eta})-\rho^2\sigma^4}$; the weights γ_ω and γ_η are constants, independent of the realized signals. The equation follows from the fact that conditional expectations are linear functions for elliptically contoured distributions (see, e.g., Fang et al. (1990) Theorem 2.18).

The random vector (ω, η, θ) is Laplace, since affine transformations of random vectors that follows an elliptical distribution with a given characteristic generator follow a distribution with the same characteristic generator (see, e.g., Fang et al. (1990) Theorem 2.16).

The first moment of θ is zero, because the mean of τ is the zero vector. Plugging in the weights γ_ω and γ_η, the second moments of (ω, η, θ) can straightforwardly be calculated:

$$\begin{aligned}\sigma_\theta^2 &= \gamma_\omega^2 Var\left(s_\omega\right) + \gamma_\eta^2 Var\left(s_\eta\right) + 2\gamma_\omega\gamma_\eta Cov\left(s_\omega, s_\eta\right) \\ &= \gamma_\omega^2\left(\sigma^2 + \sigma_{\varepsilon_\omega}^2\right) + \gamma_\eta^2\left(\sigma^2 + \sigma_{\varepsilon_\eta}^2\right) + 2\gamma_\omega\gamma_\eta Cov\left(\omega, \eta\right) = \sigma^2 \frac{\frac{\sigma_{\varepsilon_\omega}^2}{\sigma^2} + \frac{\sigma_{\varepsilon_\eta}^2}{\sigma^2}\rho^2 + 1 - \rho^2}{\left(1 + \frac{\sigma_{\varepsilon_\omega}^2}{\sigma^2}\right)\left(1 + \frac{\sigma_{\varepsilon_\eta}^2}{\sigma^2}\right) - \rho^2} \\ &= Cov\left(\eta, \theta\right),\end{aligned}$$

and

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

ii) Letting $a \equiv \frac{\sigma_{\varepsilon_\omega}^2}{\sigma^2}$ and $b \equiv \frac{\sigma_{\varepsilon_\eta}^2}{\sigma^2}$, we can rewrite $\sigma_{\omega\theta}$ and σ_θ^2 as

$$\sigma_{\omega\theta} = \rho\sigma^2 \frac{a + b + 1 - \rho^2}{\left(1 + a\right)\left(1 + b\right) - \rho^2},$$

and

$$\sigma_\theta^2 = \sigma^2 \frac{a + b\rho^2 + 1 - \rho^2}{\left(1 + a\right)\left(1 + b\right) - \rho^2}.$$

Consider first the set of feasible levels of $\sigma_{\omega\theta} = C$. Note that for $a = 0$ or $b = 0$, the covariance is constant and equal to $\rho\sigma^2 = \sigma_{\omega\eta}$. Moreover, the covariance is decreasing in a for given b and decreasing in b for given a. By l'Hôpital's rule, we have

$$\lim_{b\to\infty} \frac{a + b + 1 - \rho^2}{\left(1 + a\right)\left(1 + b\right) - \rho^2} = \frac{1}{1 + a},$$

and

$$\lim_{a\to\infty} \frac{a + b + 1 - \rho^2}{\left(1 + a\right)\left(1 + b\right) - \rho^2} = \frac{1}{1 + b}.$$

So, letting both a and b (in whatever order) go to infinity results in a covariance of zero. By continuity, any $C \in (0, \sigma_{\omega\eta}]$ can be generated by finite levels a, b. Including the case where no signal is observed at all, we can generate all $C \in [0, \sigma_{\omega\eta}]$.

Consider next the set of feasible σ_θ^2 for any given level $\sigma_{\omega\theta} = C$. Distinguish two cases, i) $C = \sigma_{\omega\eta}$ and ii) $C \in [0, \sigma_{\omega\eta})$.

Case i) requires that $a = 0$ or $b = 0$. If $b = 0$, then $\frac{a+b\rho^2+1-\rho^2}{(1+a)(1+b)-\rho^2} = 1$ and thus $\sigma_\theta^2 = \sigma^2$ for all a. If $a = 0$, then

$$\sigma_\theta^2 = \sigma_\eta^2 \frac{b\rho^2 + 1 - \rho^2}{(1+b) - \rho^2}$$

is decreasing in b and attains value $\sigma_\theta^2 = \sigma^2$ for $b = 0$. Moreover,

$$\lim_{b\to\infty} \frac{b\rho^2 + 1 - \rho^2}{(1+b) - \rho^2} = \rho^2.$$

Hence, for $C = \sigma_{\omega\eta}$, $\sigma_\theta^2 \in [\rho^2\sigma^2, \sigma^2]$; the lower limit is included because we allow for the case where only one signal is observed.

Case ii) $C \in [0, \sigma_{\omega\eta})$ requires that $a > 0$ and $b > 0$. Let $\delta \equiv \frac{C}{\sigma_{\omega\eta}} \in [0, 1)$. The combinations of a and b that generate C satisfy

$$\frac{a + b + 1 - \rho^2}{(1+a)(1+b) - \rho^2} = \delta.$$

Solving for a as a function of b, we obtain

$$a(b;\delta) = \frac{(1-\delta)(1+b-\rho^2)}{\delta b - (1-\delta)} = \frac{(1+b-\rho^2)}{\frac{\delta}{1-\delta}b - 1}.$$

The function $a(b;\delta)$ is decreasing in b and has the limit

$$\lim_{b\to\infty} \frac{1+b-\rho^2}{\frac{\delta}{1-\delta}b - 1} = \frac{1-\delta}{\delta}.$$

In the limit as $b \to \frac{1-\delta}{\delta}$, we obtain $a \to \infty$. Hence, C can be generated for $b > \frac{1-\delta}{\delta}$ and $a = \frac{(1+b-\rho^2)}{\frac{\delta}{1-\delta}b-1}$. Substituting for $\frac{(1+b-\rho^2)}{\frac{\delta}{1-\delta}b-1}$ into σ_θ^2, we obtain

$$\sigma_\theta^2(b, a(b;\delta), \delta) = \sigma^2 \frac{\frac{(1+b-\rho^2)}{\frac{\delta}{1-\delta}b-1} + b\rho^2 + 1 - \rho^2}{\left(1 + \frac{(1+b-\rho^2)}{\frac{\delta}{1-\delta}b-1}\right)(1+b) - \rho^2} = \sigma^2 \frac{b\delta\rho^2 + 1 - \rho^2}{1 + b - \rho^2}.$$

The derivative of this expression in b is $\frac{\left(\delta\rho^2-1\right)\left(1-\rho^2\right)}{\left(1+b-\rho^2\right)^2} < 0$, so $Var\left(\theta; b, a\left(b;\delta\right), \delta\right)$ is continuous and monotone decreasing in b. In the limit as b tends to infinity, we obtain

$$\lim_{b\to\infty} \sigma^2 \frac{b\delta\rho^2 + 1 - \rho^2}{1 + b - \rho^2} = \sigma^2 \delta \rho^2 = \sigma^2 \frac{C}{\sigma_{\omega\eta}} \rho^2 = \rho C.$$

In the limit as $b \to \frac{1-\delta}{\delta}$, we obtain

$$\lim_{b\to\frac{1-\delta}{\delta}} \sigma^2 \frac{b\delta\rho^2 + 1 - \rho^2}{1 + b - \rho^2} = \sigma^2 \frac{\frac{1-\delta}{\delta}\delta\rho^2 + 1 - \rho^2}{1 + \frac{1-\delta}{\delta} - \rho^2} = \delta\sigma^2 = \frac{1}{\rho} C.$$

Hence, we have shown that for any given $C \in \left[0, \sigma_{\omega\eta}\right)$, $\sigma_\theta^2 \in \left[\rho C, \frac{1}{\rho} C\right]$. We include the lower limit, because the case where $b \to \infty$ is equivalent to the case with one signal only. ■

Lemma A1 For the Laplace distribution, for $0 \leq \underline{\theta} < \overline{\theta}$ we can write

$$\mathbb{E}\left[\theta \middle| \theta \in \left[\underline{\theta}, \overline{\theta}\right]\right] = \mathbb{E}\left[\theta \middle| \theta \geq 0\right] + \overline{\theta} - g\left(\overline{\theta} - \underline{\theta}\right), \tag{16}$$

where $g\left(q\right) = \frac{q}{1-\exp(-\lambda q)}$ and $\frac{1}{\lambda} = \mathbb{E}\left[\theta \middle| \theta \geq 0\right]$. The function $g\left(q\right)$ *satisfies* $\lim_{q\to 0} g\left(q\right) = \frac{1}{\lambda}$ *and has limits* $\lim_{q\to\infty} g\left(q\right) = \infty$, *and* $\lim_{q\to\infty}\left(q - g\left(q\right)\right) = 0$. *Moreover, the function is increasing and convex, with a slope satisfying* $\lim_{q\to 0} g'\left(q\right) = \frac{1}{2}$ *and attaining the limit* $\lim_{q\to\infty} g'\left(q\right) = 1$.

Proof of Lemma A1. Recall that the marginal density of the Laplace distribution is $f_\theta\left(\theta\right) = \lambda e^{-\lambda|\theta|}$. For the Laplace distribution for $0 \leq \underline{\theta} < \overline{\theta}$, an integration by parts gives

$$\begin{aligned}
\mathbb{E}\left[\theta \middle| \theta \in \left[\underline{\theta}, \overline{\theta}\right]\right] &= \int_{\underline{\theta}}^{\overline{\theta}} \theta\lambda \frac{\exp^{-\lambda\theta}}{\exp^{-\lambda\underline{\theta}} - \exp^{-\lambda\overline{\theta}}} d\theta = -\,\theta \frac{\exp^{-\lambda\theta}}{\exp^{-\lambda\underline{\theta}} - \exp^{-\lambda\overline{\theta}}}\Bigg|_{\underline{\theta}}^{\overline{\theta}} + \int_{\underline{\theta}}^{\overline{\theta}} \frac{\exp^{-\lambda\theta}}{\exp^{-\lambda\underline{\theta}} - \exp^{-\lambda\overline{\theta}}} d\theta. \\
&= \overline{\theta} - \frac{\left(\overline{\theta} - \underline{\theta}\right)}{1 - \exp^{-\lambda\left(\overline{\theta} - \underline{\theta}\right)}} - \frac{1}{\lambda} \frac{\exp^{-\lambda\theta}}{\exp^{-\lambda\underline{\theta}} - \exp^{-\lambda\overline{\theta}}}\Bigg|_{\underline{\theta}}^{\overline{\theta}} \\
&= \frac{1}{\lambda} + \overline{\theta} - \frac{\left(\overline{\theta} - \underline{\theta}\right)}{1 - \exp^{-\lambda\left(\overline{\theta} - \underline{\theta}\right)}}.
\end{aligned}$$

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

The slope of the function is

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

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

Differentiating $g(q)$ twice, we obtain

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

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

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

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

Proof of Lemma 3. **i)** Equation (3) follows immediately from applying again Fang et al. (1990) Theorem 2.18.

ii) By the law of iterated expectations,

$$\mathbb{E}\left[\omega|\,\theta\in\left[\underline{\theta},\overline{\theta}\right]\right]=\mathbb{E}\left[\mathbb{E}\left[\omega|\theta\right]|\,\theta\in\left[\underline{\theta},\overline{\theta}\right]\right]=\mathbb{E}\left[\left.\frac{\sigma_{\omega\theta}}{\sigma_\theta^2}\theta\,\right|\theta\in\left[\underline{\theta},\overline{\theta}\right]\right]=\frac{\sigma_{\omega\theta}}{\sigma_\theta^2}\cdot\mathbb{E}\left[\theta|\,\theta\in\left[\underline{\theta},\overline{\theta}\right]\right].$$

iii) The marginal distribution of θ is a classical Laplace distribution with density of the form $f_\theta(\theta)=\lambda e^{-\lambda|\theta|}$ by the same argument as given in Lemma 2 i). Since by Lemma A1 $\mathbb{E}\left[\theta|\,\theta\in\left[\underline{\theta},\overline{\theta}\right]\right]=\frac{1}{\lambda}+\overline{\theta}-\frac{\left(\overline{\theta}-\underline{\theta}\right)}{1-\exp^{-\lambda\left(\overline{\theta}-\underline{\theta}\right)}}=\frac{1}{\lambda}+\frac{\underline{\theta}}{1-\exp^{-\lambda\left(\overline{\theta}-\underline{\theta}\right)}}-\frac{\exp^{-\lambda\left(\overline{\theta}-\underline{\theta}\right)}\overline{\theta}}{1-\exp^{-\lambda\left(\overline{\theta}-\underline{\theta}\right)}}$ and $\lim_{\overline{\theta}\to\infty}\exp^{-\lambda\left(\overline{\theta}-\underline{\theta}\right)}\overline{\theta}=0$, we have

$$\lim_{\overline{\theta}\to\infty}\mathbb{E}\left[\theta|\,\theta\in\left[\underline{\theta},\overline{\theta}\right]\right]=\mathbb{E}\left[\theta|\,\theta\geq 0\right]+\underline{\theta}.$$

For a discussion of the parameter α see the proof of Proposition 3. ■

Appendix B

Characterization of partitional equilibria

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. In our description here, we focus on symmetric equilibria. This is without loss, since for the case $c \leq 1$ symmetric equilibria are the only ones that exist. We do prove their existence, and for logconcave densities, equilibria are unique (see Szalay (2012)). For the case $c > 1$, we prove our results also allowing for asymmetric equilibria.

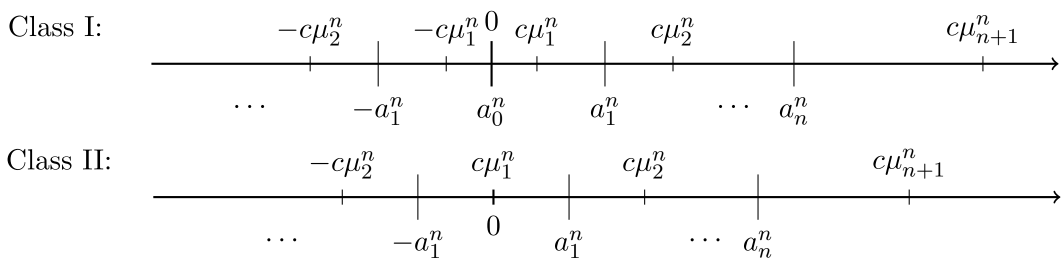


Figure 6: Class I equilibrium and Class II equilibrium.

Symmetric equilibria come in two classes; see Figure 6 for an illustration. Class I has zero as a threshold, $a_0^n = 0$, and in addition $n \geq 0$ thresholds $a_1^n, \ldots, a_n^n$ above the prior mean. By symmetry, types $-a_n^n, \ldots, -a_1^n$ 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. In this case, we eliminate a_0^n altogether. Such an equilibrium induces $2n+1$ actions. Consider Class I equilibria first.

For $n \geq 1$, let

$$\mu_i^n \equiv \mathbb{E}\left[\theta \middle| \theta \in \left[a_{i-1}^n, a_i^n\right)\right] \qquad \text{for } i = 1, \ldots, n \tag{17}$$

and $\mu_{n+1}^n \equiv \mathbb{E}\left[\theta \middle| \theta \geq a_n^n\right]$. By convention, we take all intervals as closed from below and open from above. Clearly, given quadratic loss functions and Part ii) of Lemma 3, the receiver's best reply if sender types in the interval $\left[a_{i-1}^n, a_i^n\right)$ pool is to choose

$$y\left(a_{i-1}^n, a_i^n\right) = c \cdot \mu_i^n \qquad \text{for } i = 1, \ldots, n \tag{18}$$

and $y(a_n^n, \infty) = c \cdot \mu_{n+1}^n$ if sender types with $\theta \geq a_n^n$ pool. 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 $a_1^n, \ldots, a_n^n$:

$$a_i^n - c \cdot \mu_i^n = c \cdot \mu_{i+1}^n - a_i^n, \quad \text{for } i = 1, \ldots, n. \tag{19}$$

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, (19). However, in this case conditions (17) and (18) apply only for $i = 2, \ldots, n$, while we let $\mu_1^n \equiv \mathbb{E}\left[\theta | \theta \in \left[a_{-1}^n, a_1^n\right)\right] = 0$ and $y\left(a_{-1}^n, a_1^n\right) = c \cdot \mu_1^n = 0$.

Equation (19) defines a nonlinear difference equation for any given n. The qualitative features of the equilibrium set - in particular, the maximum number n such that a solution to (19) exists - depend crucially on the magnitude of the regression coefficient.

For $c \leq 1$, the number of possible induced actions is infinite (see Proposition refprop:limit). One way to understand an equilibrium is as a combination of a "forward solution" and a "closure condition". A forward solution starting at zero takes the length of the first interval as given, say x, and computes the "next" threshold, $a_2(x)$, as a function of the preceding two, and likewise for the following thresholds. The closure condition for an equilibrium with n positive thresholds requires that x is such that type $a_n^n(x)$ is exactly indifferent between pooling downwards and upwards. Using this construction, we prove the existence of an equilibrium for arbitrary n and show that the limit as n goes to infinity is an equilibrium. 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 all intervals is increasing in the distance from the agreement point to the first threshold. Moreover, the level of the last threshold is bounded from above.

The case $c > 1$ is different in very essential ways, as shown in Proposition 2. Again, any equilibrium must feature intervals that increase in length the farther they are located from the agreement point. This is intuitive, since the extent of disagreement increases in $|\theta|$. However, the forward solution only has this feature if the length of the first interval is bounded away from zero and n is bounded.

Proof of Proposition 1. Before proving Parts i) to iii) of the proposition by a sequence of claims, we make the equilibrium conditions for the Laplace in Claim 0) explicit.

Claim 0) A Class I equilibrium is a set of marginal types that satisfies the conditions

$$cg\left(a_i^n - a_{i-1}^n\right) = 2\frac{c}{\lambda} + c\left(a_{i+1}^n - a_i^n\right) - cg\left(a_{i+1}^n - a_i^n\right) + 2\left(c-1\right)a_i^n. \tag{20}$$

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

$$cg\left(a_n^n - a_{n-1}^n\right) = 2\frac{c}{\lambda} + 2\left(c-1\right)a_n^n, \tag{21}$$

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

$$a_1^n = \frac{c}{\lambda} + c\left(a_2^n - a_1^n\right) - cg\left(a_2^n - a_1^n\right) - \left(1-c\right)a_1^n, \tag{22}$$

and in addition (20) for $i = 2, \ldots, n-1$, and (21) .

Proof: Recall the proof of Lemma A1; we write the conditional mean for the Laplace $\mu_{i+1} = \mathbb{E}\left[\theta|\,\theta \in [\theta_i, \theta_{i+1}]\right] = \mathbb{E}\left[\theta|\,\theta \geq 0\right] + \theta_{i+1} - g\left(\theta_{i+1} - \theta_i\right)$, where $0 \geq \theta_i < \theta_{i+1}$, $g\left(q\right) = \frac{q}{1-\exp(-\lambda q)}$, and $\frac{1}{\lambda} = \mathbb{E}\left[\theta|\,\theta \geq 0\right]$. For a discussion of the g function see Lemma A1. In combination with the sender's indifference conditions (19), $a_i^n - c \cdot \mu_i^n = c \cdot \mu_{i+1}^n - a_i^n$, this implies the claim.

Part i) We use the combination of forward solution and closure condition (21) to show equilibrium existence. Formally, for an initial value x a forward solution $a_2(x)$ is defined as the value of a_2 that solves

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

The forward solution for $a_i(x)$ for $i \geq 3$ is defined recursively by

$$cg\left(a_{i-1}\left(x\right) - a_{i-2}\left(x\right)\right) - \frac{c}{\lambda} - \frac{c}{\lambda} - c\left(a_i - a_{i-1}\left(x\right)\right) + cg\left(a_i - a_{i-1}\left(x\right)\right) - 2\left(c-1\right)a_{i-1}\left(x\right) = 0.$$

We prove existence of Class I equilibria first. The argument is structured as follows. In Claims i.1) to i.3), we investigate the forward solution, addressing first properties of solutions (Claims i.1) and i.2)) and then existence (Claim i.3)). In Claim i.4), we address existence and uniqueness of a fixed point. Finally, in Claim i.5) the extension to the case of Class II equilibria is presented.

Claim i.1) The forward solution features increasing intervals,

$$a_{i+1}^n - a_i^n > a_i^n - a_{i-1}^n.$$

Proof: Consider

$$\Delta\left(a_2 - x, x\right) \equiv cg\left(x\right) - \frac{c}{\lambda} + cg\left(a_2 - x\right) - c\left(a_2 - x\right) - \frac{c}{\lambda} - 2\left(c-1\right)x.$$

The forward solution for a_2 given x is the value of a_2 that solves $\Delta\left(a_2 - x, x\right) = 0$. Take $a_2 - x = x$, then

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

Since $\lim_{x\to 0} g(x) = \lim_{x\to 0} \frac{x}{1-e^{-\lambda x}} = \frac{1}{\lambda}$, we have $\lim_{x\to 0}\Delta\left(x,x\right) = 0$. Moreover,

$$\frac{\partial}{\partial x}\Delta\left(x, x\right) = 2cg'\left(x\right) - c - 2\left(c-1\right),$$

$$\frac{\partial^2}{\partial x^2}\Delta\left(x, x\right) = 2cg''\left(x\right) > 0.$$

Observe that

$$\lim_{x\to 0}\frac{\partial}{\partial x}\Delta\left(x, x\right) = c - c - 2\left(c-1\right) = -2\left(c-1\right) \geq 0$$

with a strict inequality if $c < 1$. It follows that $\Delta\left(x,x\right) > 0$ for all $x > 0$. Since for all finite a_2

$$\frac{\partial}{\partial a_2}\Delta\left(a_2 - x, x\right) = cg'\left(a_2 - x\right) - c < 0,$$

the forward solution for a_2 given x, satisfies $a_2 - x > x$.

Consider the forward solution for a_i

$$cg\left(a_{i-1}\left(x\right) - a_{i-2}\left(x\right)\right) - \frac{c}{\lambda} - \frac{c}{\lambda} - c\left(a_i - a_{i-1}\left(x\right)\right) + cg\left(a_i - a_{i-1}\left(x\right)\right) - 2\left(c-1\right)a_{i-1}\left(x\right) = 0.$$

Let $z = a_i\left(x\right) - a_{i-1}\left(x\right) = a_{i-1}\left(x\right) - a_{i-2}\left(x\right).$ Define

$$\Delta\left(z, z; a_{i-1}\right) \equiv 2\left(cg\left(z\right) - \frac{c}{\lambda}\right) - cz - 2\left(c-1\right)a_{i-1}\left(x\right).$$

Then

$$\lim_{z\to 0}\Delta\left(z, z; a_{i-1}\right) = -2\left(c-1\right)a_{i-1}\left(x\right) > 0$$

for any $a_{i-1}\left(x\right) > 0$. Since $2cg'\left(z\right) - c \geq 0$ with strict inequality for $z > 0$, we have $\Delta\left(z, z; a_{i-1}\right) > 0$ for all $z > 0$. Since the left-hand side of the equation defining the forward solution is decreasing in a_i, for any $a_{i-2}\left(x\right), a_{i-1}\left(x\right) > 0$ the solution of the forward equation must satisfy $a_i\left(x\right) - a_{i-1}\left(x\right) > a_{i-1}\left(x\right) - a_{i-2}\left(x\right).$

Claim i.2) The forward solution $a_2(x)$ satisfies $\lim_{x\to 0}(a_2(x) - x) = 0$ and $\frac{da_2}{dx} > 1$, implying that $a_2(x) - x$ is increasing in x. Moreover, the forward solutions $a_i(x) - a_{i-1}(x)$ for $i = 3, \ldots, n$ all satisfy $\lim_{x\to 0}(a_i(x) - a_{i-1}(x)) = 0$ 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.

Proof: Consider the equation determining the forward solution for $a_2(x)$, that is condition (20) 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 $x \to 0$, we obtain $\lim_{x\to 0} a_2(x) = 0$ from the fact that $\lim_{q\to 0} g(q) = \frac{1}{\lambda}$ (Lemma A1). 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 $\lim_{x\to 0}(a_2(x) - x) = 0$ 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 $\lim_{x\to 0} a_2(x) = 0$ and $\lim_{x\to 0}(a_2(x) - x) = 0$, we also have $\lim_{x\to 0} a_3(x) = 0$ and $\lim_{x\to 0}(a_3(x) - a_2(x)) = 0$. 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 $\lim_{x\to 0}(a_3(x) - a_2(x)) = 0$ 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 $\lim_{x\to 0}(a_i(x) - a_{i-1}(x)) = 0$, $\lim_{x\to 0} a_i(x) = 0$, 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 $a_{i+1}(x)$,

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

The inductive assumptions for $a_i(x)$ and $a_{i-1}(x)$ imply that $\lim_{x\to 0}(a_{i+1}(x) - a_i(x)) = 0$, so that $\lim_{x\to 0} a_{i+1}(x) = 0$. Totally differentiating, we obtain

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

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 $i = 1, \ldots, n$.

Claim i.3) For each $i = 2, \ldots, n$, 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 $x \to x_i^*$, $\lim_{x\to x_i^*} a_i(x) = \infty$. Furthermore, $x_{i+1}^* < x_i^*$.

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

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

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

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

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

$$cg(x) - \frac{c}{\lambda} < \frac{c}{\lambda} + 2(c-1)\,x. \tag{23}$$

Since $cg\left(x\right)-\frac{c}{\lambda}$ is nonnegative and increasing in x and $\frac{c}{\lambda}+2\left(c-1\right)x$ is positive for $x=0$ and nonincreasing in x, there exists a unique value x_2^* such that (23) is satisfied with equality. Hence, a finite forward solution $a_2\left(x\right)$ exists for all $x\in\left[0,x_2^*\right)$. In the limit as $x\to x_2^*$, we have $\lim_{x\to x_2^*}a_2\left(x\right)=\infty$.

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

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

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

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

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

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

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

$$cg\left(a_{i-1}\left(x\right)-a_{i-2}\left(x\right)\right)-\frac{c}{\lambda}<\frac{c}{\lambda}+2\left(c-1\right)a_{i-1}\left(x\right). \tag{24}$$

Given the derived properties of the forward solution, we have that $cg\left(a_{i-1}\left(x\right)-a_{i-2}\left(x\right)\right)-\frac{c}{\lambda}$ is nonnegative and increasing in x and $\frac{c}{\lambda}+2\left(c-1\right)a_{i-1}\left(x\right)$ is positive for $x=0$ and nonincreasing in x. Therefore, there exists a unique value $x=x_i^*$ such that (24) is satisfied with equality. Hence a finite forward solution $a_i\left(x\right)$ exists for all $x\in\left[0,x_i^*\right)$. In the limit as $x\to x_i^*$, we have $\lim_{x\to x_i^*}a_i\left(x\right)=\infty$.

Define

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

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

Claim i.4) For any n there exist a unique fixed point.

Proof: Take the forward solution for $a_i(x)$ for $i = 2, \ldots, n$ and consider the difference between the left-hand and the right-hand side of (21), 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 $i = 2, \ldots, n+1$ exist and are finite for all $x \in [0, \tilde{x}_n)$. Since $x^*_{n+1} < x^*_n$, the forward solutions for $i = 2, \ldots, n$ exist and are finite at $x = \tilde{x}_n$. Hence, this completes the proof that there exists exactly one fixed point.

Claim i.5) For all n, there exists a unique Class II equilibrium.

Proof: 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 (20) for $i = 2, \ldots, n-1$ and condition (21).

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 $\lim_{a_2 \to x}\left(\frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) - (1-c)x\right) = -(1-c)x$ and $\lim_{a_2 \to \infty}\left(\frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) - (1-c)x\right) = \frac{c}{\lambda} - (1-c)x$. 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 $\lim_{x \to \frac{c}{\lambda(2-c)}} a_2(x) = \infty$. 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 $i = 3, \ldots, n$ as well as the fixed point condition (21) are unchanged, all the remaining arguments are unchanged.

Part ii) Before analyzing the limits of equilibrium thresholds as $n \to \infty$ in Claims ii.2) and ii.3), claim ii.1) establishes some important monotonicity properties of equilibrium thresholds.

Claim ii.1) The sequence $(a_1^n)_n$ is monotone decreasing, while the sequence $(a_n^n)_n$ is monotone increasing. Moreover, equilibrium thresholds are nested,

$$a_1^{n+1} < a_1^n < a_2^{n+1} < \cdots a_n^{n+1} < a_n^n < a_{n+1}^{n+1} \qquad \forall n. \tag{25}$$

Proof: Using the notation from Part i), 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 Part i) that $a_i^{n+1} < a_i^n$ for $i = 1, \ldots, n$. 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 $a_{i+1}^{n+1} > a_i^n$ for $i = 1, \ldots, n$.

We start with two preliminary observations. Firstly, the "next" solution of the forward equation, $a_{i+1}^k(x)$ for $i = 1, \ldots, k-1$, $k = n, n+1$ 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 $i = 3, \ldots, k-1$ and $k = n, n+1$, satisfy:

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

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

and

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

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

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

and

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

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 $j = l$.

Suppose $a_{j+1}^{n+1}(\tilde{x}_{n+1}) > a_j^n(\tilde{x}_n)$ for all $j = 1, \ldots, l-1$ 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 $a_{l+2}^{n+1}(\tilde{x}_{n+1}) \geq a_{l+1}^n(\tilde{x}_n) + \left(a_{l+1}^{n+1}(\tilde{x}_{n+1}) - a_l^n(\tilde{x}_n)\right)$. 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.

Claim ii.2) Equilibrium thresholds converge for $n \to \infty$.

Proof: We know from Part i) that $(a_1^n)_n$ is monotone decreasing in n. Since the sequence is bounded by zero it must converge. Similarly, by Claim ii.1) the sequence $(a_n^n)_n$ is monotone increasing in n. The fixed point condition, (21), implies that it is bounded by $\frac{c}{1-c}\frac{1}{2\lambda}$, hence it converges. Since equilibrium thresholds are nested (cf. (25)) all sequences of thresholds must converge for $n \to \infty$.

Claim ii.3) The limit of the sequences of thresholds and actions is an equilibrium.

Proof: The limit is an equilibrium if $\lim_{n\to\infty} c\mu_i^n \leq \lim_{n\to\infty} a_i^n \leq \lim_{n\to\infty} c\mu_{i+1}^n$. Therefore, we have to show that equilibrium thresholds remain ordered in the limit, $\lim_{n\to\infty} a_i^n < \lim_{n\to\infty} a_{i+1}^n$. For all finite n, thresholds are ordered in equilibrium, $a_i^n < a_{i+1}^n$, since they are ordered for any forward equation. By Claim ii.2) equilibrium thresholds converge; denote the limits by $\overline{a}_i = \lim_{n\to\infty} a_i^n$ for all i. By convergence, for any ε there is a N such that for all $n > N$: $a_i^n \geq \overline{a}_i - \frac{\varepsilon}{2}$ and $a_{i+1}^n \leq \overline{a}_{i+1} + \frac{\varepsilon}{2}$. Suppose for contradiction that $\overline{a}_i \geq \overline{a}_{i+1} + \delta$ for some $\delta > 0$; this implies

$$a_i^n \geq \overline{a}_i - \frac{\varepsilon}{2} \geq \overline{a}_{i+1} + \delta - \frac{\varepsilon}{2} \geq a_{i+1}^n - \frac{\varepsilon}{2} + \delta - \frac{\varepsilon}{2} > a_{i+1}^n,$$

for all $\varepsilon < \delta$. Hence thresholds remain ordered in the limit and the limit is an equilibrium.

Part iii) In the limit as $n \to \infty$, we have $\lim_{n\to\infty} \tilde{x}_n = 0$.

Proof: The fixed point argument in the proof of Part i) implies that the sequence $(\tilde{x}_n)_n$ is monotone decreasing. Since it is bounded from below by zero it converges. As before, we use the notation $a_1^n = \tilde{x}_n = x_{n+1}^*$.

Recall that $x_{n+1}^* < x_n^*$ and that the forward solution for $a_n(x)$ exists for $x \leq x_n^*$, where x_n^* satisfies

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

Monotonicity of the forward solutions, $a_k(x) > a_{k-1}(x)$, and increasing length of the intervals, $a_k(x) - a_{k-1}(x) > a_{k-1}(x) - a_{k-2}(x)$, imply for $c \leq 1$ the following. For any $x > 0$ there is a k such that

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

and

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

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 $\lim_{n\to\infty} x_n^* = 0$, implying that the length of the first interval goes to zero, $\lim_{n\to\infty} \tilde{x}_n = 0$.

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

■

Proof of Proposition 2. Before proving that actions are bounded away from zero for Class I equilibria in Claim 1) and for Class II equilibria in Claim 2), Claim 0) shows a monotonicity condition. Finally, Claim 3) proves finiteness of equilibria.

Claim 0) If a Class I equilibrium exists, it features increasing intervals for all $i = 1, \ldots, n-1$,

$$a_{i+1}^n - a_i^n > a_i^n - a_{i-1}^n; \tag{26}$$

If a Class II equilibrium exists, it always shares this feature for $i = 2, \ldots, n-1$.

Proof: Consider first Class I equilibria for given $n \geq 2$. For $n < 2$, the question is meaningless. Define

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

For $c \geq 2$, no equilibrium of the considered kind exists this is shown in Claim 1) below. Now take $c \in (1, 2)$. For $n = 2$, the indifference condition of type a_2^n and a_1^n are, in that order,

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

and

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

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

$$z_2^n = \frac{c}{2-c}g\left(z_1^n\right).$$

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

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

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

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

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

$$cg\left(z_{n-1}^n\right) = \left(2-c\right)z_n^n,$$

and hence

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

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

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

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

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

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

and

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

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 $z_{i-1}^n < z_i^n$.

Class II equilibria have the same indifference conditions for the marginal types a_i^n for $i = 2, \ldots, n-1$. 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.

Claim 1) In any Class I equilibrium the receiver's induced actions are bounded away from zero.

Proof: The proof proceeds as follows. Any equilibrium must be a solution to the forward equation. This requires that the solution of the forward equation exists and features increasing intervals. This is possible only if the length of the first interval is bounded away from zero.

The forward equation for a_2 is given by

$$cg(x) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) + 2(c-1)x. \tag{27}$$

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

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

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

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

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: i) $c \in \left(1, \frac{4}{3}\right]$, ii) $c \in \left(\frac{4}{3}, 2\right)$, and iii) $c \geq 2$.

i) For $c \in \left(1, \frac{4}{3}\right]$, there exists a solution $a_2(x)$ for $x < \overline{x}$ where $\overline{x}$ is the unique value of x that satisfies $cg(\overline{x}) - \frac{c}{\lambda} = \frac{c}{\lambda} + 2(c-1)\overline{x}$. To see this, note that 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, $2(c-1)x < cg(x) - \frac{c}{\lambda}$ is satisfied for all $x > 0$. $cg(x) - \frac{c}{\lambda} < \frac{c}{\lambda} + 2(c-1)x$ holds for x small since $\lim_{x\to 0} cg(x) - \frac{c}{\lambda} = 0 < \frac{c}{\lambda}$. 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 $x \to \infty$.

ii) For $c \in \left(\frac{4}{3}, 2\right)$, there exists a solution $a_2(x)$ for $x \in (\underline{x}, \overline{x})$ where $\underline{x}$ is the uniqe value of x that satisfies $2(c-1)\underline{x} < cg(\underline{x}) - \frac{c}{\lambda}$. Note that for $c \in \left(\frac{4}{3}, 2\right)$ we have $\frac{c}{2} < 2(c-1) < c$. Since $\lim_{x\to 0} g'(x) = \frac{1}{2}$, 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.

iii) For $c \geq 2$ we have $2(c-1) \geq c$ and therefore $2(c-1) \geq cg'(x)$ for all x. Hence, $2(c-1)x \geq cg(x) - \frac{c}{\lambda}$ for all x so that 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 $c \in \left(1, \frac{4}{3}\right]$. Since equilibrium thresholds have to satisfy the increasing interval property (26), 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 right-hand side and the left-hand side of condition (27) 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 (27) is increasing in a_2. We have $\lim_{x\to 0} D(x) = 0$. Moreover, the slope of $D(x)$ at $x = 0$ is $D'(x)|_{x=0} = 2(c-1) > 0$. Hence, for x small, we would get $a_2(x) - x < x$, violating the increasing interval property (26). 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.

Claim 2) In any Class II equilibrium all but at most one of the receiver's induced actions are bounded away from zero.

Proof: Given x, $a_2(x)$ is the value of a_2 that solves

$$cg(a_2 - x) - \frac{c}{\lambda} = c(a_2 - x) + (c-2)x. \tag{28}$$

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

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

for any $c \geq 2$ and any $x \geq 0$. Therefore, we consider $1 < c < 2$ from now on. Equation (28) has a solution for $x < \frac{c}{\lambda(2-c)}$, which satisfies $\lim_{x\to 0} a_2(x) = 0$ and moreover,

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

Rearranging (28) 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

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

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

$$0=\frac{c}{\lambda}+c(a_3-a_2(x))-cg(a_3-a_2(x))+4\frac{c-1}{2-c}(a_2(x)-x)-\frac{c}{2-c}\left(cg(a_2(x)-x)-\frac{c}{\lambda}\right). \tag{30}$$

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 show that the expression is strictly positive for $a_3-a_2(x)=a_2(x)-x$, to get $a_3(x)-a_2(x)<a_2(x)-x$, in contradiction to the increasing interval property (26).

Note that the right-hand side of (30) 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 (30) 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

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

where we use that $g'(z)\big|_{z=0}=\frac{1}{2}$. 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 (30) 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 (26) 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 $\Pr(\theta\in(0,x])=\frac{1}{2}\left(1-e^{-\lambda x}\right)$ and $\Pr(\theta\in(-y,0])=\Pr(\theta\in[0,y))=\frac{1}{2}\left(1-e^{-\lambda y}\right)$. Let $\delta(x,y)\equiv\frac{\left(1-e^{-\lambda x}\right)}{\left(1-e^{-\lambda x}\right)+\left(1-e^{-\lambda y}\right)}$, 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, $\omega(x,y) \gtreqqless 0$ for $x \gtreqqless y$. The forward solution $a_2(x,y)$ is the value of a_2 that solves

$$-c\omega(x,y) = \frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) + (c-2)x. \tag{31}$$

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,

$$cg(a_2(x,y) - x) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_3 - a_2(x,y)) - cg(a_3 - a_2(x,y)) + 2(c-1)a_2(x,y). \tag{32}$$

Rearranging (31), we can write

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

Adding up with (32),

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

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.

To conclude, we have shown that in a Class I equilibrium, $\mu_1^n > 0$, in a Class II equilibrium, $\mu_2^n > 0$ (by definition, we have $\mu_1^n = 0$). Finally, in any asymmetric equilibrium, the lengths of the intervals that are adjacent to the interval containing the prior mean are bounded away from zero.

Claim 3) Only a finite number of distinct receiver actions are induced in equilibrium.

Proof: Consider a Class I equilibrium first. We show that the solution of the forward equation violates the increasing interval property (26) for n large enough.

Consider the forward equation for a_n with length x of the first interval,

$$a_{n-1}(x)-c\left(\frac{1}{\lambda}+a_{n-1}(x)-g\left(a_{n-1}(x)-a_{n-2}(x)\right)\right)=c\left(\frac{1}{\lambda}+a_n-g\left(a_n-a_{n-1}(x)\right)\right)-a_{n-1}(x).$$

There is a unique value $a_n(x)$ of a_n that solves this equation. Let a_n be such that $a_n - a_{n-1}(x) = a_{n-1}(x) - a_{n-2}(x) \equiv z$, for some $z > 0$. Let $D(z;x)$ denote the difference between the right-hand side and the left-hand side of the forward equation evaluated at z,

$$D(z;x) = 2(c-1)\,a_{n-1}(x) + c\left(\frac{2}{\lambda}+z-2g(z)\right).$$

If $D(z;x) > 0$, then a_n needs to decrease to satisfy the forward equation. Note that $\frac{2}{\lambda} + z - 2g(z)$ is strictly negative for $z > 0$ and $2(c-1)\,a_{n-1}(x)$ is strictly positive. From the first part of the proposition, we know that x is bounded away from zero. Moreover, x has to satisfy the increasing interval property (26) for $a_2(x) - x > x$. Suppose that the increasing interval property is satisfied up to the interval $a_{n-1}(x) - a_{n-2}(x)$. (If not, then we are done already.) If all intervals up to $a_{n-1}(x) - a_{n-2}(x)$ satisfy the increasing interval property, then $a_{n-1}(x) \geq (n-1)\,x$. Note that x does not depend on n. Hence, for any finite z, there is a $n(z,x)$ such that $D(z;x) > 0$ for all $n \geq n(z,x)$, implying that the increasing interval property is violated.

For the Class II equilibrium, note that the forward equation for a_n (for $n \geq 3$) is the same as above. The only difference is the value of $a_{n-1}(x)$ and the lower bound on x. However, $a_{n-1}(x) \geq x + (n-2)(a_2(x) - x)$. Note again that x and $a_2(x)$ do not depend on n.

The same argument can be given for the asymmetric case. Hence, the same conclusions obtain. ■

Appendix C

Proof of Lemma 4. We have

$$\mathbb{E}_{\mu\omega} u^r (c\mu, \omega) = -\mathbb{E}_{\mu\omega} \left[(c\mu - \omega)^2 \right] = -\mathbb{E}_{\mu\omega} \left[c^2 \mu^2 - 2c\omega\mu - \omega^2 \right] = c^2 \mathbb{E}_{\mu} [\mu]^2 - \sigma^2.$$

The last equality follows from the fact that $\mathbb{E}_{\mu\omega} [\omega\mu] = c\mathbb{E}_{\mu} [\mu^2]$, which we now demonstrate. Let $j = 1, \ldots, J$ label the partition intervals in the natural order. Let Θ_j denote a generic interval, μ_j the mean over that interval, and define $\Pr(\Theta_j) \equiv \Pr(\theta \in \Theta_j)$. Moreover, let $f_{\omega\theta}(\omega, \theta)$ denote the joint density of ω and θ. We can write

$$\begin{aligned} \mathbb{E}_{\mu\omega} [\omega\mu] = \mathbb{E}_{\mu} \left[\mathbb{E}_{\omega|\mu} [\omega\mu | \mu] \right] &= \sum_j \Pr(\Theta_j) \left[\mathbb{E}_{\omega|\mu=\mu_j} \left[\omega\mu | \mu = \mu_j \right] \right] \\ &= \sum_j \Pr(\Theta_j) \mu_j \int \omega f_{\omega|\Theta_j} (\omega | \theta \in \Theta_j) \, d\omega, \end{aligned}$$

where

$$f_{\omega|\Theta_j} (\omega | \theta \in \Theta_j) = \int_{\Theta_j} \frac{f_{\omega\theta}(\omega, \theta)}{\Pr(\Theta_j)} d\theta.$$

Interchanging the order of integration (Fubini's theorem) gives us,

$$\sum_j \Pr(\Theta_j) \mu_j \int \omega \int_{\Theta_j} \frac{f_{\omega\theta}(\omega, \theta)}{\Pr(\Theta_j)} d\theta d\omega = \sum_j \Pr(\Theta_j) \mu_j \int_{\Theta_j} \int \omega \frac{f_{\omega\theta}(\omega, \theta)}{\Pr(\Theta_j)} d\omega d\theta.$$

Dividing and multiplying by $f(\theta)$, recognizing that $\frac{f_{\omega\theta}(\omega,\theta)}{f(\theta)} = f_{\omega|\theta}(\omega | \theta)$, and applying (4) (Lemma 3 ii)), we have

$$\begin{aligned} \sum_j \Pr(\Theta_j) \mu_j \int_{\Theta_j} \int \omega \frac{f_{\omega\theta}(\omega, \theta)}{\Pr(\Theta_j)} d\omega d\theta &= \sum_j \Pr(\Theta_j) \mu_j \int_{\Theta_j} \int \omega f_{\omega|\theta} (\omega | \theta) \, d\omega \frac{f(\theta)}{\Pr(\Theta_j)} d\theta \\ &= \sum_j \Pr(\Theta_j) \mu_j \int_{\Theta_j} c\theta \frac{f(\theta)}{\Pr(\Theta_j)} d\theta \\ &= c \sum_j \Pr(\Theta_j) \mu_j^2. \end{aligned}$$

Substituting back and simplifying delivers the result.

■

Proof of Proposition 3. Preliminaries on Probabilities Recall that $f(\theta)$ and $F(\theta)$ denote the pdf and cdf of θ. For $k = 2, \ldots, n$, define $\hat{p}_{k-1}$ as the probability that $\theta \in [a_{k-2}, a_{k-1}]$ conditional on $\theta \geq a_{k-2}$,

$$\hat{p}_{k-1} \equiv \frac{F(a_{k-1}) - F(a_{k-2})}{1 - F(a_{k-2})}.$$

Accordingly, $1 - \hat{p}_{k-1} = \frac{1-F(a_{k-1})}{1-F(a_{k-2})}$ is the probability that $\theta \geq a_{k-1}$, conditional on $\theta \geq a_{k-2}$. We can write these probabilities as

$$\hat{p}_{k-1} = \frac{\mathbb{E}[\theta|\,\theta \geq a_{k-1}] - \mathbb{E}[\theta|\,\theta \geq a_{k-2}]}{\mathbb{E}[\theta|\,\theta \geq a_{k-1}] - \mu_{k-1}} \tag{33}$$

and

$$1 - \hat{p}_{k-1} = \frac{\mathbb{E}[\theta|\,\theta \geq a_{k-2}] - \mu_{k-1}}{\mathbb{E}[\theta|\,\theta \geq a_{k-1}] - \mu_{k-1}}.$$

To see this, note that

$$\begin{aligned}
(F(a_{k-1}) - F(a_{k-2}))\,\mu_{k-1} &= \int_{a_{k-2}}^{a_{k-1}} \theta f(\theta)\, d\theta = \int_{a_{k-2}}^{\infty} \theta f(\theta)\, d\theta - \int_{a_{k-1}}^{\infty} \theta f(\theta)\, d\theta \\
&= (1 - F(a_{k-2}))\,\mathbb{E}[\theta|\,\theta \geq a_{k-2}] - (1 - F(a_{k-1}))\,\mathbb{E}[\theta|\,\theta \geq a_{k-1}].
\end{aligned}$$

Hence

$$\hat{p}_{k-1}\mu_{k-1} = \mathbb{E}[\theta|\,\theta \geq a_{k-2}] - (1 - \hat{p}_{k-1})\,\mathbb{E}[\theta|\,\theta \geq a_{k-1}].$$

Solving for $\hat{p}_{k-1}$ delivers the desired conclusion.

Observe that $(1 - \hat{p}_{k-2}) \cdot \hat{p}_{k-1}$ is the probability of the event $\theta \in [a_{k-2}, a_{k-1}]$ conditional on $\theta \geq a_{k-3}$, and $(1 - \hat{p}_{k-2}) \cdot (1 - \hat{p}_{k-1})$ is the probability of the event $\theta \geq a_{k-1}$ conditional on $\theta \geq a_{k-3}$. To see this, note that $1 - \hat{p}_{k-2} = \Pr[\theta \geq a_{k-2}|\,\theta \geq a_{k-3}] = \frac{1-F(a_{k-2})}{1-F(a_{k-3})}$ and recall that $\hat{p}_{k-1} = \frac{F(a_{k-1})-F(a_{k-2})}{1-F(a_{k-2})}$.

Induction

Induction Basis:

Recall that $\mu_+ \equiv \mathbb{E}[\theta|\,\theta \geq 0]$. Let the distribution satisfy $\mathbb{E}\left[\theta|\,\theta \geq \overline{\theta}\right] = \mu_+ + \alpha \cdot \overline{\theta}$ for all $\overline{\theta} \geq 0$ and for some constant α. Note that for the Laplace distribution, $\alpha = 1$. Finally, define

$$\hat{c} \equiv \alpha c.$$

Assume that $\hat{c} \in (0,2)$. Let

$$X_n^n\left(a_{n-1}^n\right) \equiv \hat{p}_n^n\left(\hat{c}\mu_n^n - \hat{c}\mu_+\right)^2 + \left(1-\hat{p}_n^n\right)\left(\hat{c}\mu_{n+1}^n - \hat{c}\mu_+\right)^2.$$

$X_n^n\left(a_{n-1}^n\right)$ is equal to $\hat{c}^2$ times the expected squared deviation of the truncated means from μ_+, conditional on $\theta \geq a_{n-1}^n$. Substituting for $\hat{p}_n^n$ from (33), and multiplying and dividing by $\hat{c}$ for convenience, we can write

$$\begin{aligned} X_n^n\left(a_{n-1}^n\right) &= \frac{\hat{c}\mu_{n+1}^n - \hat{c}\mathbb{E}\left[\theta \mid \theta \geq a_{n-1}^n\right]}{\hat{c}\mu_{n+1}^n - \hat{c}\mu_n^n}\left(\hat{c}\mu_n^n - \hat{c}\mu_+\right)^2 \\ &\quad + \frac{\hat{c}\mathbb{E}\left[\theta \mid \theta \geq a_{n-1}^n\right] - \hat{c}\mu_n^n}{\hat{c}\mu_{n+1}^n - \hat{c}\mu_n^n}\left(\hat{c}\mu_{n+1}^n - \hat{c}\mu_+\right)^2. \end{aligned}$$

Expanding the numerators of the probabilities by $\pm\hat{c}\mu_+$, reorganizing according to common factors, and simplifying (using lengthy but straightforward computations), we can write

$$X_n^n\left(a_{n-1}^n\right) = A_n^n + B_n^n,$$

where

$$A_n^n \equiv \left(\hat{c}\mu_{n+1}^n - \hat{c}\mu_+\right)\left(\hat{c}\mu_+ - \hat{c}\mu_n^n\right)$$

and

$$B_n^n \equiv \left(\hat{c}\mathbb{E}\left[\theta \mid \theta \geq a_{n-1}^n\right] - \hat{c}\mu_+\right)\left(\left(\hat{c}\mu_n^n + \hat{c}\mu_{n+1}^n\right) - 2\hat{c}\mu_+\right).$$

We can further simplify the terms A_n^n and B_n^n, using the indifference condition of the marginal type a_n^n (multiplied by α), $\hat{c}\mu_n^n + \hat{c}\mu_{n+1}^n = 2\alpha a_n^n$ and the linearity of the tail conditional expectation, $\alpha a_n^n = \mu_{n+1}^n - \mu_+$. Substituting the latter condition into the former one, and solving for μ_{n+1}^n, we obtain

$$\frac{\hat{c}\mu_n^n + 2\mu_+}{2-\hat{c}} = \mu_{n+1}^n.$$

Substituting back into A_n^n and B_n^n, and simplifying, we have shown that

$$\begin{aligned} X_n^n\left(a_{n-1}^n\right) &= \frac{\hat{c}}{2-\hat{c}}\left(\hat{c}\mu_n^n + \hat{c}\mu_+\right)\left(\hat{c}\mu_+ - \hat{c}\mu_n^n\right) \\ &\quad + 2\left(\hat{c}\mathbb{E}\left[\theta \mid \theta \geq a_{n-1}^n\right] - \hat{c}\mu_+\right)\left(\frac{\hat{c}}{2-\hat{c}}\left(\mu_n^n + \mu_+\right) - \hat{c}\mu_+\right). \end{aligned}$$

Induction hypothesis:

$$X_k^n\left(a_{k-1}^n\right)=\frac{\hat{c}}{2-\hat{c}}\left(\hat{c}\mu_+ + \hat{c}\mu_k^n\right)\left(\hat{c}\mu_+ - \hat{c}\mu_k^n\right)+2\left(\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-1}^n\right]-\hat{c}\mu_+\right)\left(\frac{\hat{c}}{2-\hat{c}}\left(\mu_+ + \mu_k^n\right)-\hat{c}\mu_+\right).$$

Inductive step:

By definition

$$X_{k-1}^n\left(a_{k-2}^n\right)=\hat{p}_{k-1}^n\left(\hat{c}\mu_{k-1}^n-\hat{c}\mu_+\right)^2+\left(1-\hat{p}_{k-1}^n\right)X_k^n\left(a_{k-1}^n\right).$$

Substituting for the probability distribution from (33) and using the inductive hypothesis, we have

$$\begin{aligned}X_{k-1}^n\left(a_{k-2}^n\right) \quad = \quad & \frac{\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-1}^n\right]-\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-2}^n\right]}{\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-1}^n\right]-\hat{c}\mu_{k-1}^n}\left(\hat{c}\mu_{k-1}^n-\hat{c}\mu_+\right)^2 \\ & +\frac{\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-2}^n\right]-\hat{c}\mu_{k-1}^n}{\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-1}^n\right]-\hat{c}\mu_{k-1}^n}\left(\begin{array}{c}\frac{\hat{c}}{2-\hat{c}}\left(\hat{c}\mu_+ + \hat{c}\mu_k^n\right)\left(\hat{c}\mu_+ - \hat{c}\mu_k^n\right) \\ +2\left(\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-1}^n\right]-\hat{c}\mu_+\right)\left(\frac{\hat{c}}{2-\hat{c}}\left(\mu_+ + \mu_k^n\right)-\hat{c}\mu_+\right)\end{array}\right).\end{aligned}$$

Expanding the numerators of the probabilities by $\pm\hat{c}\mu_+$ and reorganizing according to common factors, we can write

$$X_{k-1}^n\left(a_{k-2}^n\right)=A_{k-1}^n+B_{k-1}^n$$

with

$$\begin{aligned}A_{k-1}^n \equiv & \frac{\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-1}^n\right]-\hat{c}\mu_+}{\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-1}^n\right]-\hat{c}\mu_{k-1}^n}\left(\hat{c}\mu_{k-1}^n-\hat{c}\mu_+\right)^2+\frac{\hat{c}\mu_+ - \hat{c}\mu_{k-1}^n}{\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-1}^n\right]-\hat{c}\mu_{k-1}^n} \\ & \cdot\left(\frac{\hat{c}}{2-\hat{c}}\left(\hat{c}\mu_+ + \hat{c}\mu_k^n\right)\left(\hat{c}\mu_+ - \hat{c}\mu_k^n\right)+2\left(\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-1}^n\right]-\hat{c}\mu_+\right)\left(\frac{\hat{c}}{2-\hat{c}}\left(\mu_+ + \mu_k^n\right)-\hat{c}\mu_+\right)\right)\end{aligned}$$

and

$$\begin{aligned}B_{k-1}^n \equiv & \frac{\hat{c}\mu_+ - \hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-2}^n\right]}{\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-1}^n\right]-\hat{c}\mu_{k-1}^n}\left(\hat{c}\mu_{k-1}^n-\hat{c}\mu_+\right)^2+\frac{\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-2}^n\right]-\hat{c}\mu_+}{\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-1}^n\right]-\hat{c}\mu_{k-1}^n} \\ & \cdot\left(\frac{\hat{c}}{2-\hat{c}}\left(\hat{c}\mu_+ + \hat{c}\mu_k^n\right)\left(\hat{c}\mu_+ - \hat{c}\mu_k^n\right)+2\left(\hat{c}\mathbb{E}\left[\theta\right|\theta\geq a_{k-1}^n\right]-\hat{c}\mu_+\right)\left(\frac{\hat{c}}{2-\hat{c}}\left(\mu_+ + \mu_k^n\right)-\hat{c}\mu_+\right)\right).\end{aligned}$$

We consider each term in sequence. We first show that

$$A_{k-1}^n=\frac{\hat{c}}{2-\hat{c}}\left(\hat{c}\mu-\hat{c}\mu_{k-1}^n\right)\left(\hat{c}\mu+\hat{c}\mu_{k-1}^n\right).$$

The indifference condition of type a_{k-1}^n,$\hat{c}\mu_k^n = 2\alpha a_{k-1}^n - \hat{c}\mu_{k-1}^n$, allows us to substitute for $\hat{c}\mu_k^n$. Hence,

$$\begin{aligned} A_{k-1}^n =& \frac{\hat{c}\mathbb{E}\left[\theta|\, \theta \geq a_{k-1}^n\right] - \hat{c}\mu_+}{\hat{c}\mathbb{E}\left[\theta|\, \theta \geq a_{k-1}^n\right] - \hat{c}\mu_{k-1}^n} \left(\hat{c}\mu_{k-1}^n - \hat{c}\mu_+\right)^2 + \frac{\hat{c}\mu_+ - \hat{c}\mu_{k-1}^n}{\hat{c}\mathbb{E}\left[\theta|\, \theta \geq a_{k-1}^n\right] - \hat{c}\mu_{k-1}^n} \\ & \cdot \left(\frac{\hat{c}}{2-\hat{c}} \left(\hat{c}\mu_+ + 2\alpha a_{k-1}^n - \hat{c}\mu_{k-1}^n\right) \left(\hat{c}\mu_+ - \left(2\alpha a_{k-1}^n - \hat{c}\mu_{k-1}^n\right)\right) \right. \\ & \left. \quad +2\left(\hat{c}\mathbb{E}\left[\theta|\, \theta \geq a_{k-1}\right] - \hat{c}\mu_+\right) \left(\frac{1}{2-\hat{c}} \left(\hat{c}\mu + 2\alpha a_{k-1} - \hat{c}\mu_{k-1}^n\right) - \hat{c}\mu_+ \right) \right). \end{aligned}$$

Collecting terms with the common factor $\frac{\hat{c}\mathbb{E}\left[\theta|\theta\geq a_{k-1}^n\right]-\hat{c}\mu_+}{\hat{c}\mathbb{E}\left[\theta|\theta\geq a_{k-1}^n\right]-\hat{c}\mu_{k-1}^n}\left(\hat{c}\mu_{k-1}^n - \hat{c}\mu_+\right)$ and simplifying, we get

$$\begin{aligned} A_{k-1}^n =& \frac{\hat{c}\mu_+ - \hat{c}\mu_{k-1}^n}{\hat{c}\mathbb{E}\left[\theta|\, \theta \geq a_{k-1}^n\right] - \hat{c}\mu_{k-1}^n} \left(\frac{\hat{c}}{2-\hat{c}} \left(\left(\hat{c}\mu_+ - \hat{c}\mu_{k-1}^n\right)\left(\hat{c}\mu_+ + \hat{c}\mu_{k-1}^n\right) + \left(-4\left(\alpha a_{k-1}^n\right)^2 + 4\alpha a_{k-1}^n \hat{c}\mu_{k-1}^n\right)\right) \right. \\ & \left. + \left(\hat{c}\mathbb{E}\left[\theta|\, \theta \geq a_{k-1}^n\right] - \hat{c}\mu_+\right) \left(\frac{\hat{c}}{2-\hat{c}} \left(\hat{c}\mu_{k-1}^n + \hat{c}\mu_+\right) + \frac{4}{2-\hat{c}} \left(\alpha a_{k-1}^n - \hat{c}\mu_{k-1}^n\right) \right) \right). \end{aligned}$$

It is easy to see that

$$\begin{aligned} & \frac{\hat{c}\mu_+ - \hat{c}\mu_{k-1}^n}{\hat{c}\mathbb{E}\left[\theta|\, \theta \geq a_{k-1}^n\right] - \hat{c}\mu_{k-1}^n} \frac{\hat{c}}{2-\hat{c}} \left(\hat{c}\mu_+ - \hat{c}\mu_{k-1}^n\right)\left(\hat{c}\mu_+ + \hat{c}\mu_{k-1}^n\right) \\ + & \frac{\hat{c}\mu_+ - \hat{c}\mu_{k-1}^n}{\hat{c}\mathbb{E}\left[\theta|\, \theta \geq a_{k-1}^n\right] - \hat{c}\mu_{k-1}^n} \left(\hat{c}\mathbb{E}\left[\theta|\, \theta \geq a_{k-1}^n\right] - \hat{c}\mu_+\right) \frac{\hat{c}}{2-\hat{c}} \left(\hat{c}\mu_{k-1}^n + \hat{c}\mu_+\right) \\ = & \frac{\hat{c}}{2-\hat{c}} \left(\hat{c}\mu_+ - \hat{c}\mu_{k-1}^n\right)\left(\hat{c}\mu_+ + \hat{c}\mu_{k-1}^n\right). \end{aligned}$$

Moreover, since $\hat{c}\mathbb{E}\left[\theta|\, \theta \geq a_{k-1}\right] - \hat{c}\mu_+ = \alpha\hat{c}a_{k-1}^n$, all the terms involving a_{k-1}^n exactly cancel out. Hence, the desired conclusion follows.

The term B_{k-1}^n is simplified using the same essential steps: the indifference condition of the marginal type to substitute for $\hat{c}\mu_k^n$, collecting terms with common factors and terms that add up conveniently, and the linear tail conditional expectation. Hence we can conclude that

$$B_{k-1}^n = 2\left(\hat{c}\mathbb{E}\left[\theta|\, \theta \geq a_{k-2}^n\right] - \hat{c}\mu_+\right)\left(\frac{\hat{c}}{2-\hat{c}} \left(\mu_+ + \mu_{k-1}^n\right) - \hat{c}\mu_+ \right).$$

This completes the induction.

Building on the characterization, we can compute $\mathbb{E}\left[\mu^2\right]$ in any equilibrium.

Finite Class I: In a Class I equilibrium, $a_0^n = 0$. Hence,

$$X_1^n\left(a_0^n\right) = \frac{\hat{c}}{2-\hat{c}}\left(\hat{c}\mu_+ - \hat{c}\mu_1^n\right)\left(\hat{c}\mu_+ + \hat{c}\mu_1^n\right).$$

Recalling the definition of $X_{k-1}^n\left(a_{k-2}^n\right)$, we also have

$$X_1^n\left(a_0^n\right) = \hat{c}^2\sum_{i=1}^{n+1}\hat{p}_i^n\left(\mu_i^n - \mu_+\right)^2 = \hat{c}^2\sum_{i=1}^{n+1}\hat{p}_i^n\left(\mu_i^n\right)^2 - \hat{c}^2\mu_+^2,$$

where the second equality follows from the fact that $\sum_{i=1}^{n+1}\hat{p}_i^n\left(\mu_i^n - \mu_+\right) = 0$. Solving for $\sum_{i=1}^{n+1}\hat{p}_i^n\left(\mu_i^n\right)^2$ between these equations, we get

$$\sum_{i=1}^{n+1}\hat{p}_i^n\left(\mu_i^n\right)^2 = \frac{X_1^n\left(a_0^n\right)}{\hat{c}^2} + \mu_+^2 = \frac{2}{2-\hat{c}}\mu_+^2 - \frac{\hat{c}}{2-\hat{c}}\left(\mu_1^n\right)^2.$$

For the uni-dimensional Laplace distribution with density

$$f\left(\theta\right) = \frac{1}{2}\lambda e\left(-\lambda\left|\theta\right|\right),$$

the scale parameter λ determines all the relevant moments of the distribution. In particular, $\mu_+ = \frac{1}{\lambda}$ and $\sigma_\theta^2 = \frac{2}{\lambda^2} = 2\mu_+^2$. Moreover, $\alpha = 1$. Hence, we have $\sum_{i=1}^{n+1}\hat{p}_i\left(\mu_i^n\right)^2 = \frac{1}{2-c}\sigma_\theta^2 - \frac{c}{2-c}\left(\mu_1^n\right)^2$. By the symmetry of the distribution, $\Pr\left[\theta \geq 0\right] = \Pr\left[\theta \leq 0\right] = \frac{1}{2}$ and $\mathbb{E}\left[\mu^2 \middle| \theta \geq 0\right] = \mathbb{E}\left[\mu^2 \middle| \theta \leq 0\right]$, so that

$$\mathbb{E}\left[\mu^2\right] = \frac{\mathbb{E}\left[\mu^2 \middle| \theta \geq 0\right] + \mathbb{E}\left[\mu^2 \middle| \theta \leq 0\right]}{2} = \mathbb{E}\left[\mu^2 \middle| \theta \geq 0\right].$$

Hence, we have shown that in a Class I equilibrium

$$\mathbb{E}\left[\mu^2\right] = \frac{1}{2-c}\sigma_\theta^2 - \frac{c}{2-c}\left(\mu_1^n\right)^2.$$

Finite Class II: In a Class II equilibrium, a_0 is eliminated. We have

$$\begin{aligned}
X_2^n\left(a_1^n\right) &= \frac{\hat{c}}{2-\hat{c}}\left(\hat{c}\mu_+ - \hat{c}\mu_2^n\right)\left(\hat{c}\mu_+ + \hat{c}\mu_2^n\right) + 2\left(\hat{c}\mathbb{E}\left[\theta \middle| \theta \geq a_1^n\right] - \hat{c}\mu_+\right)\left(\frac{\hat{c}}{2-\hat{c}}\left(\mu_+ + \mu_2^n\right) - \hat{c}\mu_+\right) \\
&= \frac{\hat{c}}{2-\hat{c}}\left(\hat{c}\mu_+ - \hat{c}\mu_2^n\right)\left(\hat{c}\mu_+ + \hat{c}\mu_2^n\right) + 2\alpha\hat{c}a_1^n\left(\frac{\hat{c}}{2-\hat{c}}\left(\mu_+ + \mu_2^n\right) - \hat{c}\mu_+\right).
\end{aligned}$$

Using the definition of X_2^n and then the fact that $\sum_{i=2}^{n+1} \hat{p}_i^n \mu_i^n = \mu_+ + \alpha a_1^n$ for a distribution with an linear tail conditional expectation, we get

$$\begin{aligned} \frac{X_2^n(a_1^n)}{\hat{c}^2} &= \sum_{i=2}^{n+1} \hat{p}_i^n (\mu_i^n)^2 - 2\mu_+ \sum_{i=2}^{n+1} \hat{p}_i^n \mu_i^n + \mu_+^2 \\ &= \sum_{i=2}^{n+1} \hat{p}_i^n (\mu_i^n)^2 - \mu_+^2 - 2\alpha a_1^n \mu_+. \end{aligned}$$

Hence

$$\begin{aligned} \sum_{i=2}^{n+1} \hat{p}_i^n (\mu_i^n)^2 &= \frac{X_2^n(a_1^n)}{\hat{c}^2} + \mu_+^2 + 2\alpha a_1^n \mu_+ \\ &= \frac{2}{2-\hat{c}}\mu_+^2 - \frac{\hat{c}}{2-\hat{c}}(\mu_2^n)^2 + \frac{2\alpha a_1^n}{2-\hat{c}}\left(\mu_+ + \mu_2^n\right). \end{aligned}$$

Now, we may write

$$\begin{aligned} \mathbb{E}\left[\mu^2\right] &= \Pr\left[\theta \geq a_1^n\right] \cdot \mathbb{E}\left[\mu^2 \middle| \theta \geq a_1^n\right] + \Pr\left[\theta \leq -a_1^n\right] \cdot \mathbb{E}\left[\mu^2 \middle| \theta \leq -a_1^n\right] \\ &= 2\Pr\left[\theta \geq a_1^n\right] \cdot \mathbb{E}\left[\mu^2 \middle| \theta \geq a_1^n\right] \\ &= \left(1 - \Pr\left[\theta \in [-a_1^n, a_1^n)\right]\right) \cdot \mathbb{E}\left[\mu^2 \middle| \theta \geq a_1^n\right]. \end{aligned}$$

The first equality uses the fact that $\mu_1^n = 0$ in a Class II equilibrium, and the other two equalities use the symmetry of the distribution, which implies that $\mathbb{E}\left[\mu^2 \middle| \theta \geq a_1^n\right] = \mathbb{E}\left[\mu^2 \middle| \theta \leq -a_1^n\right]$ and $\Pr\left[\theta \geq a_1^n\right] = \Pr\left[\theta \leq -a_1^n\right]$. Hence, we have shown that

$$\mathbb{E}\left[\mu^2\right] = \left(1 - \Pr\left[\theta \in [-a_1^n, a_1^n)\right]\right)\left[\frac{2}{2-\hat{c}}\mu_+^2 - \frac{\hat{c}}{2-\hat{c}}(\mu_2^n)^2 + \frac{2\alpha a_1^n}{2-\hat{c}}\left(\mu_+ + \mu_2^n\right)\right].$$

The indifference condition of the marginal type a_1^n requires that $c\mu_2^n - a_1^n = a_1^n$. Substituting for $2a_1^n = c\mu_2^n$, noting that $\hat{c} = \alpha c$, and simplifying, we obtain

$$\mathbb{E}\left[\mu^2\right] = \left(1 - \Pr\left[\theta \in \left[-\frac{c\mu_2^n}{2}, \frac{c\mu_2^n}{2}\right)\right]\right)\left[\frac{2}{2-\hat{c}}\mu_+^2 + \frac{\hat{c}}{2-\hat{c}}\mu_2^n \mu_+\right],$$

which coincides with expression (9) for $\alpha = 1$ and $\sigma_\theta^2 = 2\mu_+^2$, the Laplace case.

Limit: In a limit equilibrium resulting from the limit of a Class I equilibrium, the sequence $(\mu_1^n)_n$ satisfies $\lim_{n\to\infty} \mu_1^n = 0$. In a limit equilibrium resulting from the limit of a

Class II equilibrium, the sequences $(a_1^n)_n$ and $(\mu_2^n)_n$ satisfy $\lim_{n\to\infty} a_1^n = 0$ and $\lim_{n\to\infty} \mu_2^n = 0$. Hence, in the limit

$$\mathbb{E}\left[\mu^2\right] = \frac{2}{2-\hat{c}}\mu_+^2.$$

Substituting for the Laplace case, $\alpha = 1$ and $\sigma_\theta^2 = 2\mu_+^2$, gives expression (10).

In the limit equilibrium resulting from the limit of finite Class I and Class II equilibria, if it exists, $\mathbb{E}\left[\mu^2\right]$ is maximized. The right-hand side of (10) exceeds the right-hand side of (8) for all finite n, since $\mu_1^n > 0$ for finite n. We now show that the right-hand side of (10) also exceeds the right-hand side of (9) for all finite n. Noting that $\left(1 - \Pr\left[\theta \in \left[-\frac{c\mu_2^n}{2}, \frac{c\mu_2^n}{2}\right)\right]\right) = \exp\left(-\lambda\frac{c\mu_2^n}{2}\right) = \exp\left(-\frac{c\mu_2^n}{2\mu_+}\right)$,

$$\frac{2}{2-\hat{c}}\mu_+^2 > \left(1 - \Pr\left[\theta \in \left[-\frac{c\mu_2^n}{2}, \frac{c\mu_2^n}{2}\right)\right]\right)\left[\frac{2}{2-\hat{c}}\mu_+^2 + \frac{\hat{c}}{2-\hat{c}}\mu_2^n\mu_+\right]$$

is equivalent to

$$1 - \exp\left(-\frac{c\mu_2^n}{2\mu_+}\right) > \exp\left(-\frac{c\mu_2^n}{2\mu_+}\right)\frac{c\mu_2^n}{2\mu_+}.$$

This is true for all $\mu_2^n > 0$ since the function $\exp(-x)(1+x)$ satisfies $\exp(-x)(1+x) < 1$ for all $x > 0$. ■

Proof of Lemma 5. We derive here the density of the marginal distribution of θ. Let $\hat{f}(\theta;\alpha)$ and $\hat{F}(\theta;\alpha)$ denote the density and cdf of the distribution, conditional on $\theta \geq 0$. After an integration by parts, (5) is equivalent to

$$\mu_+ + \alpha\theta = \theta + \frac{\int_\theta^{\overline{\theta}} \left(1 - \hat{F}(t;\alpha)\right) dt}{1 - \hat{F}(\theta;\alpha)}. \tag{34}$$

Define $q(\theta) = \int_\theta^{\overline{\theta}} \left(1 - \hat{F}(t;\alpha)\right) dt$ and note that $\dot{q} \equiv \frac{\partial q(\theta)}{\partial\theta} = -\left(1 - \hat{F}(\theta;\alpha)\right)$. In terms of these functions, we can write (34) as the ordinary differential equation

$$\frac{\dot{q}}{q} = \frac{1}{(1-\alpha)\theta - \mu_+}$$

with initial condition $q(0) = \mu_+$. The solution is

$$q(\theta) = \left(\mu_+\right)^{-\frac{\alpha}{1-\alpha}} \left(\mu_+ - \theta (1-\alpha)\right)^{\frac{1}{1-\alpha}}.$$

To satisfy $\lim_{\theta \to \overline{\theta}} \hat{F}(\theta; \alpha) = 1$, we have $\overline{\theta} = \frac{\mu_+}{1-\alpha}$ for $\theta < 1$. For $\alpha \geq 1$, the support is $\mathbb{R}^+$. Differentiating twice, we obtain the density

$$\hat{f}(\theta; \alpha) = \alpha \left(\mu_+\right)^{-\frac{\alpha}{1-\alpha}} \left(\mu_+ - \theta (1-\alpha)\right)^{\frac{2\alpha-1}{1-\alpha}}. \tag{35}$$

For future reference, the cdf is

$$\hat{F}(\theta; \alpha) = 1 - \left(\mu_+\right)^{-\frac{\alpha}{1-\alpha}} \left(\mu_+ - \theta (1-\alpha)\right)^{\frac{\alpha}{1-\alpha}}.$$

The density is square integrable since $\alpha < 2$. Straightforward integration reveals that the variance v_+^2 of the distribution is $v_+^2 = \frac{\alpha}{2-\alpha}\mu_+^2$.

Consider now the density on the whole support. By symmetry and the variance decomposition, $\sigma_\theta^2 = v_+^2 + \mu_+^2$, so

$$\sigma_\theta^2 = \frac{2}{2-\alpha}\mu_+^2.$$

Hence, we get expression (14). ■

Proof of Proposition 4. Note that the first part is a corollary to Proposition 3. So, we only need to verify the upper bound on $\mathbb{E}[\mu^2]$ in any symmetric equilibrium. For Class I equilibria this is obvious, so consider Class II equilibria. Note that $\Pr[\theta \in [-a_1^n, a_1^n)] = \hat{F}(a_1^n; \alpha)$. Moreover,

$$\hat{F}(\theta; \alpha) = 1 - \left(\mu_+\right)^{-\frac{\alpha}{1-\alpha}} \left(\mu_+ - \theta (1-\alpha)\right)^{\frac{\alpha}{1-\alpha}}.$$

Hence,

$$\frac{2}{2-\hat{c}}\mu_+^2 > \left(1 - \Pr\left[\theta \in \left[-\frac{c\mu_2^n}{2}, \frac{c\mu_2^n}{2}\right)\right]\right) \left[\frac{2}{2-\hat{c}}\mu_+^2 + \frac{\hat{c}}{2-\hat{c}}\mu_2^n \mu_+\right]$$

is equivalent to

$$\left(1 - \left(\mu_+\right)^{-\frac{\alpha}{1-\alpha}} \left(\mu_+ - \frac{c\mu_2^n}{2}(1-\alpha)\right)^{\frac{\alpha}{1-\alpha}}\right) \frac{2}{2-\hat{c}}\mu_+^2 > \left(\mu_+\right)^{-\frac{\alpha}{1-\alpha}} \left(\mu_+ - \frac{c\mu_2^n}{2}(1-\alpha)\right)^{\frac{\alpha}{1-\alpha}} \frac{\hat{c}}{2-\hat{c}}\mu_2^n \mu_+.$$

Simplifying, we obtain

$$1 > \left(1 - \frac{c\mu_2^n}{2\mu_+}(1-\alpha)\right)^{\frac{\alpha}{1-\alpha}} \left(1 + \alpha\frac{c\mu_2^n}{2\mu_+}\right).$$

To see this is always satisfied, consider the function $h(x) \equiv (1 - x(1-\alpha))^{\frac{\alpha}{1-\alpha}}(1+\alpha x)$. Note that $h(0) = 1$. Moreover, $h'(x) < 0$ for $x > 0$. ■

References

Aghion, P. and Tirole, J. (1997). Formal and real authority in organizations, *Journal of Political Economy* **105**(1): 1–29.

Alonso, R., Dessein, W. and Matouschek, N. (2008). When does coordination require centralization?, *American Economic Review* **98**(1): 145–179.

Alonso, R. and Matouschek, N. (2008). Optimal delegation, *The Review of Economic Studies* **75**(1): 259–293.

Alonso, R. and Rantakari, H. (2013). The art of brevity, *SSRN discussion paper 2306824* .

Amador, M. and Bagwell, K. (2013). The theory of optimal delegation with an application to tariff caps, *Econometrica* **81**(4): 1541–1599.

Antic, N. and Persico, N. (2017). Communication among shareholders. Working paper.

Argenziano, R., Severinov, S. and Squintani, F. (2016). Strategic information acquisition and transmission, *American Economic Journal: Microeconomics* **8**(3): 119–155.

Austen-Smith, D. (1994). Strategic transmission of costly information, *Econometrica* **62**(4): 955–963.

Battaglini, M. (2002). Multiple referrals and multidimensional cheap talk, *Econometrica* **70**(4): 1379–1401.

Blume, A., Board, O. J. and Kawamura, K. (2007). Noisy talk, *Theoretical Economics* **2**(4): 395–440.

Chakraborty, A. and Harbaugh, R. (2007). Comparative cheap talk, *Journal of Economic Theory* **132**(1): 70–94.

Chen, Y., Kartik, N. and Sobel, J. (2008). Selecting cheap-talk equilibria, *Econometrica* **76**(1): 117–136.

Crawford, V. and Sobel, J. (1982). Strategic information transmission, *Econometrica* **50**(6): 1431–1451.

Deimen, I. and Szalay, D. (2018). Information, authority, and smooth communication in organizations. Working paper.

Demski, J. S. and Sappington, D. E. (1987). Delegated expertise, *Journal of Accounting Research* **25**(1): 68–89.

Dessein, W. (2002). Authority and communication in organizations, *The Review of Economic Studies* **69**(4): 811–838.

Dimitrakas, V. and Sarafidis, Y. (2005). Advice from an expert with unknown motives, *SSRN working paper 797844* .

Eső, P. and Szalay, D. (2017). Incomplete language as an incentive device. Working paper.

Fang, K.-T., Kotz, S. and Ng, K. W. (1990). *Symmetric multivariate and related distributions*, Chapman and Hall.

Goltsman, M., Hörner, J., Pavlov, G. and Squintani, F. (2009). Mediation, arbitration and negotiation, *Journal of Economic Theory* **144**(4): 1397–1420.

Gordon, S. (2010). On infinite cheap talk equilibria. Working paper.

Grossman, S. J. and Hart, O. D. (1986). The costs and benefits of ownership: A theory of vertical and lateral integration, *Journal of political economy* **94**(4): 691–719.

Hart, O. and Moore, J. (1990). Property rights and the nature of the firm, *Journal of political economy* **98**(6): 1119–1158.

Holmstrom, B. (1982). Moral hazard in teams, *The Bell Journal of Economics* **13**(2): 324–340.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion, *American Economic Review* **101**(6): 2590–2615.

Kotz, S., Kozubowski, T. and Podgórski, K. (2001). *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*, Birkhäuser.

Levy, G. and Razin, R. (2007). On the limits of communication in multidimensional cheap talk: a comment, *Econometrica* **75**(3): 885–893.

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

Mailath, G. J. and Nöldeke, G. (2008). Does competitive pricing cause market breakdown under extreme adverse selection?, *Journal of Economic Theory* **140**(1): 97–125.

Melumad, N. D. and Shibano, T. (1991). Communication in settings with no transfers, *The RAND Journal of Economics* **22**(2): 173–198.

Morgan, J. and Stocken, P. C. (2003). An analysis of stock recommendations, *The RAND Journal of Economics* **34**(1): 183–203.

Moscarini, G. (2007). Competence implies credibility, *American Economic Review* **97**(1): 37–63.

Ottaviani, M. and Sørensen, P. N. (2006). Professional advice, *Journal of Economic Theory* **126**(1): 120–142.

Pei, H. (2015). Communication with endogenous information acquisition, *Journal of Economic Theory* **160**: 132–149.

Rantakari, H. (2008). Governing adaptation, *The Review of Economic Studies* **75**(4): 1257–1285.

Stein, J. (1989). Cheap talk and the fed: A theory of imprecise policy announcements, *American Economic Review* **79**(1): 32–42.

Szalay, D. (2005). The economics of clear advice and extreme options, *The Review of Economic Studies* **72**(4): 1173–1198.

Szalay, D. (2012). Strategic information transmission and stochastic orders. Working paper.