# Delegated Expertise, Authority, and Communication[*]

Inga Deimen

University of Arizona and CEPR

and

Dezső Szalay

University of Bonn and CEPR

August 13, 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, endogenous conflicts, delegated expertise, strategic information transmission, delegation, effectiveness of biased communication

1

# 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 as well, 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. Moreover, if the losses from strategic communication are sufficiently sensitive to the expert's bias, then the expert has incentives to avoid this situation and to eliminate the bias. The optimal information acquisition strategy has the following features in this case. The expert acquires perfect information about the preferred action from the decision-maker's perspective. Moreover, 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 his advice.

Put differently, an expert who wishes to have a pronounced 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 biased advice is less effective if extreme disagreement between expert and decision-maker is likely. The reason is that communication based on expert-optimal information 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 receiver relevant information.

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. We develop a new set of linear models that allow for closed form expressions of the value of information. 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 time horizons may create wedges of the sort envisioned here. For example, a consultant will care relatively more about the short-term impact of his advice than the advised firm does.[5] Plain and

---

[3]In particular, our approach differs from Goltsman et al. (2009) and Alonso and Rantakari (2013) where payoffs arising from mediated talk are used to determine upper bounds on the value of communication.

[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.

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

simple, our analysis predicts that communication serves to align incentives in these contexts, when news likely induces pronounced revisions of plans. For example, in financial decision-making this is the case if return distributions have fat tails; in corporate decisions, this is the case if the firm is likely to expand substantially into the new market or retreat completely from it.

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. We characterize the environments with endogenous biases in which communication always outperforms delegation.[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, it is shown 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]

---

[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 and 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.

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. Similar to our findings they show that communication creates better incentives for information acquisition than delegation. Roughly speaking, both papers use the threat of bad communication off the equilibrium path to discipline the sender's information acquisition. Argenziano et al. (2016) show that this threat implies a higher marginal value of information under communication than under delegation so that the sender acquires *more* information under communication. Intuitively, if the receiver does not listen in case the sender does not acquire the *right amount* of information, then communication can create very strong incentives for information acquisition in the overt mode. However, this effect is robust to allowing for different beliefs off equilibrium path and also holds when information acquisition is covert, albeit for a different reason.[8] In the present paper, we study overt information acquisition and argue that the threat of bad communication serves to discipline the sender to acquire the *right pieces* of information. When biased communication is sufficiently ineffective the sender prefers to avoid having a bias by acquiring information that correlates better with the decision-maker's preferred choice. While both papers emphasize a virtue of communication, our argument is based on fundamentally different forces. We face a different decision-problem with conflicts that arise endogenously ex post – as a function of the information that is acquired – but are absent ex ante. Moreover, we require a multi-dimensional approach: a two-dimensional state and signal space to understand the sender's incentives to learn about one or the other dimension of the state space. Finally, to make this problem tractable, we develop a general theory of location experiments with linear posteriors which is new to the literature.[9]

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

---

[8]Under covert information acquisition, the set of actions that the sender can induce is determined by how much information acquisition the receiver expects. This can make the life of a sender who shirks on information acquisition particularly unpleasant. 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.

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 and explain the convenient linearity features of our environment. We characterize equilibria at the communication stage 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 biased 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.

---

[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.

# 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 the realization $\omega$ of a random variable $\tilde{\omega}$. More precisely, the payoff of the receiver is

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

Unfortunately, the receiver does not know $\omega$. 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 $\eta$ is the realization of a random variable $\tilde{\eta}$. We denote $(\omega, \eta)$ the state of the world. We assume that the random variables $\tilde{\omega}$ and $\tilde{\eta}$ each 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, $\tilde{\omega}$ and $\tilde{\eta}$ have identical variances $\sigma^2$ and are positively correlated with coefficient of correlation $\rho = \frac{\sigma_{\omega\eta}}{\sigma^2} \in (0, 1)$.[12] The sender does not know the state either, but he can observe signals $s_\omega = \omega + \varepsilon_\omega$ and $s_\eta = \eta + \varepsilon_\eta$ that reveal the state with noise. We assume that the noise terms are uncorrelated with each other and with the state. We write $\boldsymbol{\tau} \equiv (\omega, \eta, \varepsilon_\omega, \varepsilon_\eta)$ with typical element $\tau$ and let $\boldsymbol{\Sigma}$ denote the covariance matrix.[13]

We introduce a class of distributions that is particularly suited to study updating combined with strategic communication. Our leading case is the joint Laplace distribution, with well known marginal density $f_\tau (\tau) = \frac{1}{2}\lambda e^{-\lambda |\tau|}$ for $\lambda = \frac{\sqrt{2}}{\sigma_\tau}$ on support $\mathbb{R}$ and characteristic function $\Phi (t) = \frac{1}{1 + \frac{1}{2}t^2 \sigma_\tau^2}$. Following e.g. Kotz et al. (2001), the joint Laplace distribution is defined by the same characteristic function, $\Phi (\mathbf{t}) = \frac{1}{1 + \frac{1}{2}\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}$. It belongs to the class of

---

[12]Our analysis easily extends to not too asymmetric prior variances. The model is interesting only if $\rho > 0$, because no meaningful communication is possible for $\rho \leq 0$.
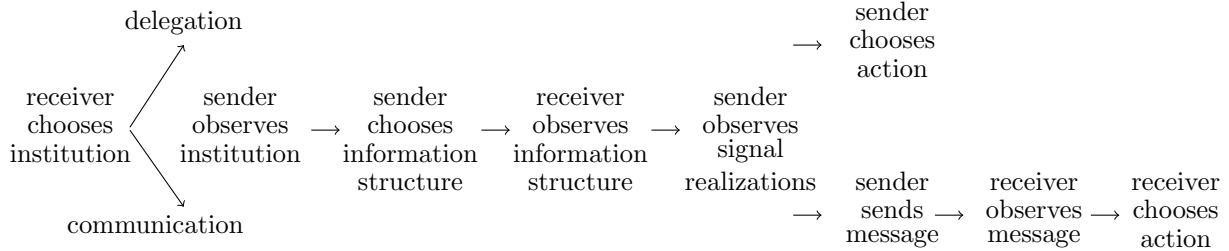
[13]Formally, we assume

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma^2 & \sigma_{\omega\eta} & 0 & 0 \\ \sigma_{\omega\eta} & \sigma^2 & 0 & 0 \\ 0 & 0 & \sigma_{\varepsilon_\omega}^2 & 0 \\ 0 & 0 & 0 & \sigma_{\varepsilon_\eta}^2 \end{pmatrix} .$$

elliptical distributions.[14] We explain in detail in Section 3 why this environment is useful and generalize our analysis beyond the Laplace case in Section 8.

## 2.2 Timing

Timeline:



The strategic interaction unfolds as indicated in the timeline.[15] 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 the action 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, after observing the receiver's choice of institution, the sender chooses what information to acquire. Formally, the sender chooses the variances of the noise terms in the signals, $\sigma^2_{\varepsilon_\omega}$ and $\sigma^2_{\varepsilon_\eta}$. 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 the selected institution of decision-making. Under delegation, the sender picks his preferred action policy. Under communication, the sender communicates

---

[14]Elliptical distributions owe their name to the fact that the level curves of their densities are elliptical. The construction via the same characteristic function, independently of the dimension, is standard (see, e.g., Fang et al. (1990)). See Mailath and Nöldeke (2008) for a model using elliptically contoured distributions in information economics.

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

9

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 $(\sigma_{\varepsilon_\omega}^2, \sigma_{\varepsilon_\eta}^2)$ in the covariance matrix $\Sigma$. Secondly, given $d = communication$, given the information structure $\Sigma$, 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 $\mathcal{D} \to \mathbb{R}_+^2$, $d \mapsto (\sigma_{\varepsilon_\omega}^2, \sigma_{\varepsilon_\eta}^2)$ and $M : \mathbb{R}^2 \times \mathbb{R}_+^2 \to \mathbb{M}$, $(s_\omega, s_\eta, \sigma_{\varepsilon_\omega}^2, \sigma_{\varepsilon_\eta}^2) \mapsto m$. 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, $Y : \mathbb{R}^2 \times \mathbb{R}_+^2 \to \mathbb{R}$, $(s_\omega, s_\eta, \sigma_{\varepsilon_\omega}^2, \sigma_{\varepsilon_\eta}^2) \mapsto y$.

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, $Y : \mathbb{M} \times \mathbb{R}_+^2 \to \mathbb{R}$, $(m, \sigma_{\varepsilon_\omega}^2, \sigma_{\varepsilon_\eta}^2) \mapsto y$. As is well known, the receiver never mixes over actions, due to the strict concavity of her payoff function in $y$.

There is commitment to $\mathcal{D}$ and $\Sigma$, 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 over sender types who might have 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.[16]

# 3 The Informational Environment

The purpose of this section is twofold. Firstly, we simplify our problem by showing that it is sufficient to focus on communication about the sender's posterior optimal action. Secondly, we explain the linearity properties of our environment and discuss how information choices shape biases in decision making.

## 3.1 A useful reduction

Suppose the sender has observed the signal realizations $(s_\omega, s_\eta)$. Due to quadratic losses, the ideal policy from his perspective is to match the action to his posterior mean conditional on the observed signal realizations. Let $\theta$ be the level of the sender's posterior mean,

$$\theta \equiv \mathbb{E}\left[\tilde{\eta}|\tilde{s}_\omega = s_\omega, \tilde{s}_\eta = s_\eta\right],$$

and let $\tilde{\theta}$ be the ex ante random level of the posterior mean. Sender types with distinct signal realizations that aggregate to the same posterior mean share the same ideal action and, more generally, share exactly the same preferences over any pair of actions. This makes

---

[16]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 could account for any kind of off-path threats. In this case, communication with appropriate off path threats would always dominate delegation.

it essentially impossible – that is, except on measure zero sets – to elicit the underlying signals from the sender.

**Lemma 1** *Any equilibrium under communication is essentially equivalent to one where the sender's message strategy is a function of $\theta$ only and all sender types $(s_\omega, s_\eta)$ such that $\mathbb{E}\left[\tilde{\eta} | \tilde{s}_\omega = s_\omega, \tilde{s}_\eta = s_\eta\right] = \theta$ induce the same action. Moreover, any equilibrium can either be characterized by an interval partition of the support of $\tilde{\theta}$ or is fully revealing about $\theta$.*

The lemma is intuitive. As usual, we can think of communication as of recommending actions. The value of his posterior mean is all the sender needs to know to compute his ideal action. The receiver would appreciate to know the individual signal realizations to compute her ideal action. However, the sender is almost surely not kind enough to reveal them. Lemma 1 allows us to eliminate the signals from the analysis and to focus on the sender's posterior mean, $\theta$.

## 3.2   Information and biases

Due to the elliptical joint distribution, the sender's posterior mean is a linear function of the realized signals, $\mathbb{E}\left[\tilde{\eta} | \tilde{s}_\omega = s_\omega, \tilde{s}_\eta = s_\eta\right] = \gamma_\omega s_\omega + \gamma_\eta s_\eta$, with weights $\gamma_\omega, \gamma_\eta$ that are independent of the signal realizations and moreover, $(\tilde{\omega}, \tilde{\eta}, \tilde{\theta})$ follows the same joint distribution as $\tilde{\tau}$. The exact weights are provided in the proof of the following Lemma 2.

Everything is as if the sender observed the realization $\theta$ of an aggregated signal that contains information of use to the sender as well as to the receiver. The covariances of the aggregated signal, $\theta$, with the underlying state, $(\omega, \eta)$, naturally measure the informational content of the aggregated signal. The exact expressions of the covariances are[17]

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

and

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

---

[17]The equality $\sigma_{\eta\theta} = \sigma_\theta^2$ follows from the linearity of the updating rules.

Intuitively, a higher precision of any signal increases both covariances and hence corresponds to a higher informational content. The set of all second moments that can be generated by some combination of noise variances is denoted by $\Gamma$ and depicted in the left panel of Figure 1. The extreme points of the feasible set are easy to understand. If the sender's underlying signals are infinitely noisy, then both covariances are zero (the origin in the left panel of Figure 1). If the signal $\tilde{s}_\omega$ is perfectly precise and the signal $\tilde{s}_\eta$ is infinitely noisy, the covariance with $\tilde{\omega}$ is maximal, reaching a level $\sigma_{\omega\eta}$, and the covariance with $\tilde{\eta}$ takes value $\rho^2\sigma^2$ (the top left corner of $\Gamma$). If the sender observes $\eta$ perfectly, then both covariances are maximal, $\sigma_\theta^2 = \sigma^2$ and $\sigma_{\omega\theta} = \sigma_{\omega\eta}$ (the top right corner of $\Gamma$).
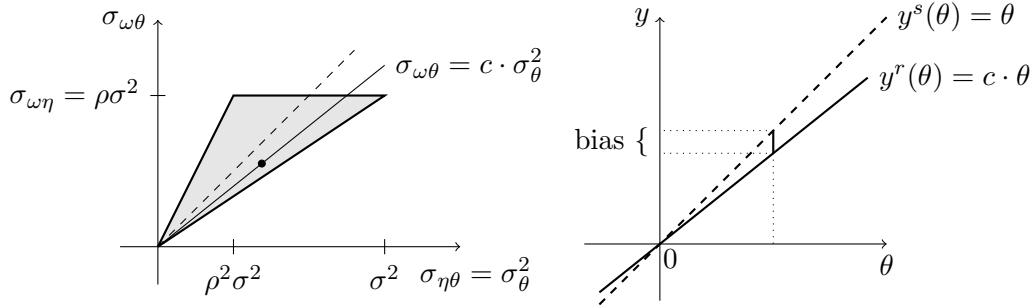


Figure 1: Left panel: the set $\Gamma$ of feasible second moments of the joint distribution of $\left(\tilde{\omega}, \tilde{\eta}, \tilde{\theta}\right)$. Right panel: ideal choices as a function of the underlying information for a particular choice of information structure (dot in the left panel).

For convenience of the reader, we summarize these insights in the following lemma.

**Lemma 2** *For any given $\left(\sigma_{\varepsilon_\omega}^2, \sigma_{\varepsilon_\eta}^2\right) \in \mathbb{R}_+^2$, the vector of random variables $\left(\tilde{\omega}, \tilde{\eta}, \tilde{\theta}\right)$ follows a joint Laplace distribution with first moments $\mathbb{E}\left[\tilde{\omega}\right] = \mathbb{E}\left[\tilde{\eta}\right] = \mathbb{E}[\tilde{\theta}] = 0$ and second moments given by (1) and (2). Moreover, a joint distribution of $\left(\tilde{\omega}, \tilde{\eta}, \tilde{\theta}\right)$ can be generated through Bayesian updating from signals $(s_\omega, s_\eta)$ if and only if $\sigma_{\omega\theta} \in [0, \sigma_{\omega\eta}]$ and for any given $\sigma_{\omega\theta} = C$, $\sigma_\theta^2 \in \left[\rho C, \frac{1}{\rho}C\right]$.*

Consider next how information $\theta$ shapes ideal decisions if it is directly observable. By construction, the sender's ideal policy is $y^s(\theta) = \theta$. Again due to elliptical distributions, the

13

receiver's ideal decision conditional on $\theta$ is $y^r(\theta) = \mathbb{E}\left[\tilde{\omega}|\tilde{\theta} = \theta\right] = c \cdot \theta$, where

$$c \equiv \frac{\sigma_{\omega\theta}}{\sigma_\theta^2} = \frac{\rho\left(\sigma_{\varepsilon_\omega}^2 + \sigma_{\varepsilon_\eta}^2 + \sigma^2\left(1 - \rho^2\right)\right)}{\sigma_{\varepsilon_\omega}^2 + \sigma_{\varepsilon_\eta}^2 \rho^2 + \sigma^2\left(1 - \rho^2\right)}. \tag{3}$$

The coefficient $c \in \left[\rho, \frac{1}{\rho}\right]$ is decreasing in the amount of noise contained in $\tilde{s}_\omega$ and increasing in the amount of noise in $\tilde{s}_\eta$. The Blackwell-best information structure from each player's perspective maximizes the variance of his/her ideal decision rule. The sender's Blackwell-best signal maximizes $\sigma_\theta^2$ and hence corresponds to the top right corner in $\Gamma$. The receiver's preferred signal combination maximizes $Var(c \cdot \theta) = \frac{\sigma_{\omega\theta}^2}{\sigma_\theta^2}$, corresponding to the top left corner in $\Gamma$.

Since $\theta$ is privately observed by the sender, the relative informativeness of information structures determines the way the sender is inclined to misrepresent his information. 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 (downplay negative realizations). In contrast, if $c > 1$, then the argument is reversed. Finally, for $c = 1$, there is no bias with respect to using the aggregated signal $\theta$. Thus, the bias is endogenous in our model. Depending on what pieces of information the sender acquires, he has different incentives to lie about what he observed ex post. We illustrate the bias resulting from a particular choice of information structure (the dot in the left panel) in the right panel of Figure 1.

Due to the possibly partitional structure of equilibria in the communication game shown in Lemma 1, we also need to know how the receiver updates about $\omega$ if she learns an interval around the sender's posterior mean only. Conveniently, the linear updating rules extend to truncations on intervals:

$$\mathbb{E}\left[\tilde{\omega}|\tilde{\theta} \in \left[\underline{\theta}, \overline{\theta}\right]\right] = c \cdot \mathbb{E}\left[\tilde{\theta}\middle|\tilde{\theta} \in \left[\underline{\theta}, \overline{\theta}\right]\right] \qquad \forall \underline{\theta} \leq \overline{\theta}, \tag{4}$$

where $c$ is defined in (3).

Consider now the marginal distribution of $\theta$. It is again a Laplace distribution, because all elliptical distributions are closed under linear combinations. More generally, all properties shown so far hold for the entire class of joint elliptical distributions (hence also for, e.g., the joint Normal distribution). The reason to assume the Laplace distribution is that it satisfies

14

– in contrast to the Normal – the following condition (5). The updating rule for $\tilde{\theta}$ conditional on truncations to the tails of the distribution is linear in the truncation point.

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

The Laplace satisfies (5) with $\alpha = 1$. We embed the Laplace in the generalized subclass of elliptical distributions that satisfy (5) in Section 8. These combined linearity features enable us to provide a closed form solution for the value of communication.

**Lemma 3** *i) The joint distribution of $\left(\tilde{\omega}, \tilde{\eta}, \tilde{\theta}\right)$ features linear conditional expectations, $\mathbb{E}\left[\tilde{\omega}\middle|\tilde{\theta} = \theta\right] = c \cdot \theta$ with c defined in (3).*
*ii) The linear updating rules extend to conditioning on truncations, (4).*
*iii) The marginal distribution of $\tilde{\theta}$ is a one-dimensional Laplace distribution, and the tail conditional expectation of $\tilde{\theta}$ satisfies (5) for $\alpha = 1$.*

We are now equipped to study strategic communication.

# 4 Equilibria in the Communication Game

We now investigate equilibria in the communication game for a given information structure that the sender has chosen. Due to Lemma 1, 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 the discussion after Theorem 2). Our results are in line with the literature (Gordon (2010)). However, since we assume an unbounded state space, we have to prove everything from scratch. We state only our main results here and refer to Appendix B for details. As is 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*

15

*(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 $\lim_{n \to \infty} a_1^n = 0$.*

$c \leq 1:$
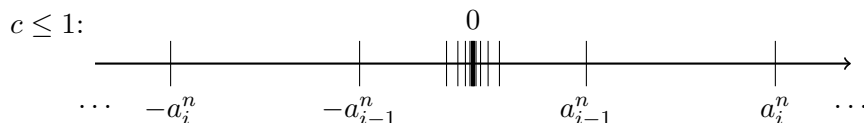
Figure 2: Intervals around the agreement point $\theta = 0$ get arbitrarily small as $n \to \infty$.

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

The sender acquires information with a view to using it in the chosen institution. While it is straightforward to compute expected payoffs under delegation, assessing the value of communication requires some work. To fix ideas, suppose first that the sender would naïvely transmit information $\theta$ 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 $\mathbb{E}[-(c\tilde{\theta} - \tilde{\omega})^2] = c^2 \sigma_\theta^2 - \sigma^2$, while the sender's expected utility would be $\mathbb{E}[-(c\tilde{\theta} - \tilde{\eta})^2] = c(2-c)\sigma_\theta^2 - \sigma^2$. However, honest communication is not necessarily an equilibrium. In a partitional equilibrium, the sender reveals only that $\theta$ belongs to some interval. Let $a_0^n = 0$ and define

$\mu_i^n \equiv \mathbb{E}\left[\tilde{\theta} \middle| \tilde{\theta} \in \left[a_{i-1}^n, a_i^n\right)\right]$ for $i = 1, \ldots, n$ and $\mu_{n+1}^n \equiv \mathbb{E}\left[\tilde{\theta} \middle| \tilde{\theta} \geq a_n^n\right]$. By the linearity of conditional expectations (Lemma 3), the support of the receiver's actions in a Class I equilibrium is $c \cdot \left\{\mu_{-(n+1)}^n, \mu_{-n}^n, \ldots, \mu_{-1}^n, \mu_1^n, \ldots, \mu_n^n, \mu_{n+1}^n\right\}$. With a minor modification, a similar description holds for a Class II equilibrium (see Appendix B for details). Before knowing which interval $\theta$ belongs to, the receiver's posterior expectation of $\tilde{\theta}$ is a random variable, $\tilde{\mu}$. Likewise, let $\tilde{v}^2$ denote the random variable that describes the receiver's posterior variance, conditional on the interval that the sender reveals. The distributions of these random variables are derived from the marginal distribution of $\tilde{\theta}$. The following lemma states that the functional forms of the expected utilities under naïve and strategic communication are the same.

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

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

*The sender's expected equilibrium utility is*

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

Intuitively, the amount of information under strategic communication is $\mathbb{E}\left[\tilde{\mu}^2\right]$, instead of $\sigma_\theta^2$ under naïve communication. To see this, note that by a standard variance decomposition, $\mathbb{E}\left[\tilde{\mu}^2\right] = \sigma_\theta^2 - \mathbb{E}\left[\tilde{v}^2\right]$. We can understand the receiver's and the sender's expected utilities as an intrinsic value of information net of a loss equal to the residual variance after strategic communication, $\mathbb{E}\left[\tilde{v}^2\right]$.

The difficulty at this point is that the equilibrium cannot be solved for in closed form. However, the convenient property of distributions with linear tail conditional expectations is that this is not necessary to obtain the value of communication in a limit equilibrium. Let $\mu_+ \equiv \mathbb{E}\left[\tilde{\theta} \middle| \tilde{\theta} \geq 0\right]$.

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

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

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

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

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

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

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

The proof uses a "dynamic programming" idea, where the sender's indifference conditions replace the usual "optimality conditions". We first compute the expected variability over the last two intervals in the tail of the distribution, conditional on $\tilde{\theta} \geq a_{n-1}^n$. The value depends on $\mu_+, \mu_n^n$, and on the truncation point $a_{n-1}^n$; $\mu_{n+1}^n$ can be eliminated using the linearity of the tail conditional expectations (5). Next, we treat the resulting value as a continuation value and use the indifference condition of the marginal type $a_{n-1}^n$ to compute the expected variability conditional on truncation to $\tilde{\theta} \geq a_{n-2}^n$. The functional form of the continuation values remains the same. By an induction argument, we show that this property holds for an arbitrary number of steps. Expressions (8) and (9) provide the resulting values for finite equilibria. In limit equilibria, $\mu_1^n$ and $\mu_2^n$, respectively, go to zero and we obtain the closed form representation (10).

The amount of information transmitted in a limit equilibrium, which exists if and only if $c \leq 1$, is simply a constant fraction $\frac{1}{2-c}$ of the total amount, $\sigma_\theta^2$. The fraction depends on the information chosen by the sender through its impact on the conflict between the sender and the receiver. The fraction is increasing in $c$ and maximal if there is no conflict, $c = 1$. Since both $c = \frac{\sigma_{\omega\theta}}{\sigma_\theta^2}$ and $\sigma_\theta^2$ are endogenous, the sender faces a trade-off between intrinsically more useful information structures (high $\sigma_\theta^2$) and transmittable information structures (high $\frac{\sigma_{\omega\theta}}{\sigma_\theta^2}$). We next analyze this trade-off.

# 6   Information Acquisition

We now turn to the sender's choice of the information structure. We assume that sender and receiver coordinate on an equilibrium that gives them the highest possible expected utility

in the set of all equilibria. Clearly, selecting the most efficient equilibria is exactly in the tradition of the communication literature following Crawford and Sobel (1982).[18]

A simple argument shows that we can restrict attention to information structures such that $c = \frac{\sigma_{\omega\theta}}{\sigma_\theta^2} \leq 1$. In this case a limit equilibrium exists and gives maximal utility. In particular, note that the information structure with the highest $\sigma_\theta^2$ in the subset of $\Gamma$ such that $c \geq 1$ satisfies $\sigma_\theta^2 = \sigma_{\omega\theta} = \sigma_{\omega\eta}$, hence $c = 1$. From $(7)$, the utility gain from this information structure is higher than for any information structure with $c > 1$. Formally, $c\,(2-c)\,\mathbb{E}\left[\tilde{\mu}^2\right] \leq c\,(2-c)\,\sigma_\theta^2 < \sigma_\theta^2 < \sigma_{\omega\eta}$ for $c > 1$. The first inequality is due to the informational loss due to strategic communication. The second inequality follows from the fact that the sender would prefer to take decisions himself. Finally, the most informative information structure for the sender within the feasible set for $c \geq 1$ is given by $\sigma_\theta^2 = \sigma_{\omega\eta}$. Intuitively, information structures where the receiver would overreact $(c > 1)$, have strategic disadvantages *and* are intrinsically suboptimal from the sender's point of view.[19]

Substituting $\mathbb{E}\left[\tilde{\mu}^2\right]$ from (3) into $(7)$, we obtain the sender's maximization problem

$$\max_{\substack{\sigma_{\omega\theta},\sigma_\theta^2 \\ s.t.\ \sigma_{\omega\theta},\sigma_\theta^2 \in \Gamma,\ c \leq 1}} c\,(2-c)\,\frac{1}{2-c}\sigma_\theta^2 - \sigma^2. \tag{11}$$

The solution to problem (11) shows that the sender cannot gain from acquiring information that is intrinsically more useful to him but that cannot be transmitted without loss.

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

Based on the previous propositions, the proof is obvious. There are three channels of influence. Firstly, the sender values information structures that carry a higher informational

---

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

[19]This insight hinges on the selection of the most informative equilibrium on and off path. If we allowed for selecting a babbling equilibrium off path, we could sustain any choice of information structure in equilibrium, including the receiver's most preferred one. However, since there exists no limit equilibrium in this case, the analysis becomes intractable and it is not clear whether the receiver prefers this at all. As alluded to in Footnote 16, we could deal with this case by decomposing the optimal actions of sender and receiver into a common component and a private component, and allow the sender to observe each of the components with noise. All feasible information structures would satisfy $c \leq 1$. The receiver optimal information structure would then coincide with the one we get in Theorem 2 below.

content to him (higher $\sigma_\theta^2$). It is immediate that for a fixed value of $c$, the highest feasible variance maximizes the sender's problem. Secondly, the information structure impacts the way the receiver discounts the sender's advice through the term $c\left(2-c\right)$. Thirdly, due to strategic communication, only a fraction $\frac{1}{2-c}$ of the total amount of information from the sender's perspective can be transmitted in equilibrium. For future reference, the fraction $\frac{1}{2-c}$ measures the effectiveness of biased communication.

The combination of these effects results in a maximum expected utility for the sender given by

$$\mathbb{E}u^s\left(c\tilde{\mu},\tilde{\eta}\right)=\sigma_{\omega\eta}-\sigma^2.$$

Since this is independent of $\sigma_\theta^2$ the sender is indifferent between all the information structures in the theorem. For an illustration see the solid line in Figure 3. This indifference is a special feature of the Laplace distribution. The robust insight that we exploit in what follows is that the sender *cannot gain* from acquiring information that he would individually prefer, because the added value of such information would be lost in transmission. As shown in Section 8, the sender strictly loses from acquiring intrinsically more useful information if the fraction he can transmit is smaller than in the Laplace case, or, as we term it, biased communication is less effective than in the Laplace case.

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 a limit equilibrium is

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

a decreasing function of $\sigma_\theta^2$. Clearly, the receiver suffers if the sender chooses an information structure with a higher $\sigma_\theta^2$; at the same time, the sender derives no benefit from such behavior.

**Theorem 2** *The set of sender optimal information structures contains a 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 $\theta$ truthfully to the receiver, who follows the sender's proposal one-for-one.*

For convenience, we depict the theorem and the discussion preceding it graphically in Figure 3. At the Pareto optimal information structure within the set of sender-optimal ones,
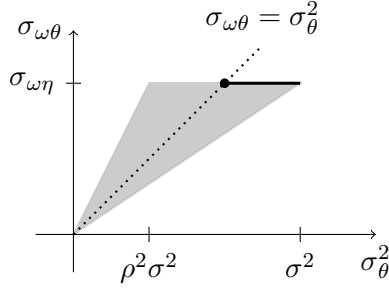
Figure 3: The solid line represents the sender-optimal information structures; the dot represents the Pareto optimum in this set.

we have $c = 1$ and the bias with respect to communicating $\theta$ is eliminated. Hence, it is an equilibrium for the sender to follow the message strategy $m(\theta) = \theta$ for all $\theta$, 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 $\theta$ 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 $c = 1$.[20] In terms of the underlying noise, the sender is perfectly informed about the receiver's ideal action, $\sigma^2_{\varepsilon_\omega} = 0$. However, he remains partially ignorant about his preferred choice, $\sigma^2_{\varepsilon_\eta} = \frac{1-\rho^2}{\rho}\sigma^2$, to convince the receiver of his unbiasedness. There is no harm if the sender knows the ideal action of the receiver. By contrast, if the sender knew his ideal action, then the receiver would discount his advice and information would be lost.

We next 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.

---

[20]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).

# 7 Delegation versus Communication

Depending on the receiver's choice of institution, the sender acquires different pieces of information. Under communication he cannot gain from choosing his privately preferred information structure and he is happy to choose information that is equally useful to himself and the receiver. Hence, 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

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

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 $\theta$. This results in an expected utility for the sender of

$$\mathbb{E}u^s\left(\tilde{\theta}, \tilde{\eta}\right) = -\mathbb{E}\left(\tilde{\theta} - \tilde{\eta}\right)^2 = \sigma_\theta^2 - \sigma^2,$$

where we have used the fact that $\sigma_{\eta\theta} = \sigma_\theta^2$ by construction of $\theta$ (see Section 3). Obviously, the sender just acquires information that is privately optimal for him, i.e. he learns $\eta$ perfectly, so that the variance of his posterior expectation is maximal in the feasible set. Formally, he chooses $\hat{\sigma}_{\omega\theta} = \sigma_{\omega\eta}$ and $\hat{\sigma}_\theta^2 = \sigma^2 = \frac{1}{\rho}\sigma_{\omega\eta}$. The receiver's expected utility under delegation is

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

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 $\rho \in (0,1)$. Formally:

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

The surprising insight is that this result holds for *any* underlying correlation of interests. The driving force behind our result is that the receiver's choice of institution – delegation versus communication – results in the sender acquiring different pieces of information. Anticipating that he will have to talk to the receiver, the sender understands that he rather

pays attention to the receiver's interests. If he paid too much attention to his own interests, the receiver would simply discount his advice, neutralizing any potential gain to the sender. By contrast, under delegation the receiver gives away any influence and the sender has no reason to pay attention to what the receiver's cares about. As a result, the receiver prefers communication over delegation.

The novelty is that the nature of the acquired information matters, under communication the receiver can punish the sender for acquiring the wrong pieces of information. This explains why our result differs from what is known for the case of exogenously given information structures and biases.[21] Indeed, if we assume an exogenously given information structure that coincides with the one chosen under delegation, $\hat{\sigma}_{\omega\theta} = \sigma_{\omega\eta}$ and $\hat{\sigma}_{\theta}^2 = \frac{1}{\rho}\sigma_{\omega\eta}$, our results are qualitatively in line with the literature: communication is strictly preferred to delegation for $\rho \in \left(0, \frac{2}{3}\right)$ and delegation is strictly preferred for $\rho \in \left(\frac{2}{3}, 1\right)$. This is due to the familiar trade-off: while communication entails a loss of information due to strategic communication, delegation results in a choice of action that is not ideal from the receiver's point of view.[22]

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

# 8    The Effectiveness of Biased Communication

We now characterize environments where the sender strictly loses under communication when acquiring information that he individually prefers. This is the case if the fraction

---

[21]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.

[22]These two scenarios correspond to the most and the least pronounced effect of institution-choice on information acquisition. For cases in between the extremes, qualitatively similar results hold. 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 under communication.

of the information he can pass on is strictly smaller than for the Laplace, or as we term it, biased strategic communication is strictly less *effective* than in the Laplace case. The reason is that the effectiveness of biased communication is more sensitive to conflicts in these environments, and as a result, the sender has incentives to acquire information that eliminates the bias. Hence the choice of institution impacts the sender's choice of information and communication trumps delegation.

## 8.1 Generalized informational environment

To formalize these thoughts, we introduce a parameter $\alpha \in (0, 2)$ that captures the informational content in strategic communication in our model. We assume that the marginal distributions of $\tilde{\boldsymbol{\tau}}$ have densities of the form,

$$f_\tau(\tau; \alpha) = \frac{1}{2\sigma_\tau}\sqrt{\frac{2\alpha^2}{(2-\alpha)}}\left(1 - \sqrt{\frac{2}{2-\alpha}}(1-\alpha)\frac{|\tau|}{\sigma_\tau}\right)^{\frac{2\alpha-1}{1-\alpha}} \tag{14}$$

on support $T$.[23] The density (14) nests the Laplace case with $\alpha = 1$. The support for all $\alpha \geq 1$ is $T = \mathbb{R}$. For $\alpha < 1$, the support is the symmetric interval around zero such that the term in brackets is non-negative. An example with interval support is the uniform distribution with $\alpha = \frac{1}{2}$. The joint distribution is elliptical and constructed in the same way as the joint Laplace case via the characteristic function of the marginal distribution (14). As we explain in more detail below all linear conditioning rules of Section 3 apply (see Lemma 5). We hence call our class of distributions *joint elliptical distributions with linear marginal tail conditional expectations*.[24]

---

[23]The density can be derived for arbitrary, positive $\alpha$. However, in general it cannot be expressed in terms of the variance, which only exists for $\alpha < 2$. Since expected utilities are only defined for finite variance in a quadratic loss model, we restrict attention to these cases.

[24]By symmetry, the characteristic function is a function of $t^2\sigma_x^2$ only, $\Phi_\alpha(t) = \phi_\alpha(t^2\sigma_x^2)$. We define the characteristic function of the multivariate distribution as $\Phi_\alpha(\mathbf{t}) = \phi_\alpha(\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t})$, that is, we take the characteristic function as invariant with respect to changes of the dimension. While the construction via characteristic functions is standard for elliptical distributions (see, e.g., Fang et al. (1990)), we are not aware of any contribution in the literature that describes the subclass of elliptical distributions with linear marginal tail conditional expectations.

We illustrate the density (14) for different values of $\alpha$ and variances equal to one in Figure 4. Relative to the Laplace distribution, densities with $\alpha > 1$ have a higher variability in the one-sided distributions, that is, more mass around zero and in the tails. Intuitively, more mass in the tails means that events with extreme disagreement between the sender and the receiver become more likely. Naturally, this drives the value of biased communication down. More variability in the one-sided distributions renders the effectiveness of strategic communication more sensitive to conflicts, as we explain in detail in Subsection 8.3 below.
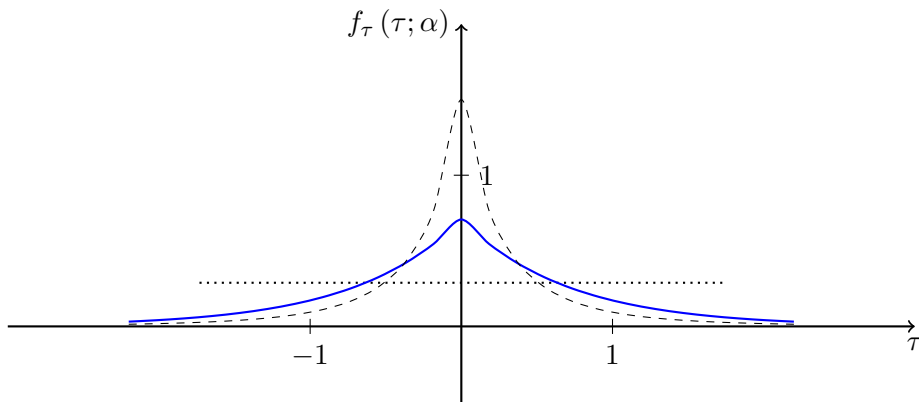


Figure 4: The density (14) depicted for $\alpha = 1.5$ (dashed); $\alpha = 1$ (solid) Laplace; $\alpha = 0.5$ (dotted) uniform. Note that the solid and the dashed line intersect again farther out in the support.

## 8.2 Communication trumps delegation

Our main result generalizes as follows:

**Theorem 4** *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 $\sigma_\theta^{2*} = \sigma_{\omega\theta}^* = \sigma_{\omega\eta}$. The ensuing communication continuation game has a smooth communication equilibrium and the receiver strictly prefers to communicate rather than to delegate decision-making to the sender.*

The intuition for this result is extremely simple: deviating to an intrinsically sender pre-ferred information structure hurts the sender because he can transmit less of his information. Moreover, relative to the Laplace case, transmitting all the information becomes more im-portant to the sender. We now develop the formal details behind this argument carefully. The key insight is in the construction of our class of distributions.

**Lemma 5** *If the distribution of $\tilde{\boldsymbol{\tau}}$ follows a joint elliptical distribution with linear marginal tail conditional expectations, then $\left(\tilde{\omega}, \tilde{\eta}, \tilde{\theta}\right)$ follows a distribution in the same class and the linear updating rules of Section 3 all apply. The density of $\tilde{\theta}$ is given by equation $(14)$ with $\alpha \in (0, 2)$. The variance of $\tilde{\theta}$ is finite and related to $\mu_+$ via $\sigma_\theta^2 = \frac{2\mu_+^2}{2-\alpha}$.*

The defining feature of our class of distributions is the linearity of their tail conditional expectations, $\mathbb{E}\left[\tilde{\theta}\middle| \tilde{\theta} \geq \bar{\theta}\right] = \mu_+ + \alpha \cdot \bar{\theta}, \forall \bar{\theta} \geq 0$. This condition can be restated as a differential equation that can be solved for the distribution as a function of $\mu_+$ and $\alpha$. Substituting for $\mu_+$ from the variance $\sigma_\theta^2 = \frac{2\mu_+^2}{2-\alpha}$ results in the density $(14)$. Our value characterization (Proposition 3) extends to all distributions with these linear updating rules. With this at hand, our analysis generalizes almost effortlessly.

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

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

*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 $\mathbb{E}\left[\tilde{\mu}^2\right]$ is attained.*

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 in a limit equilibrium depends only on the product $\alpha c$, $\mathbb{E}\left[\tilde{\mu}^2\right] \leq \frac{2}{2-\alpha c}\mu_+^2$. Using the connection between the variance and $\mu_+^2$ in Lemma 5 gives the upper bound in the proposition. Exactly as in the Laplace case, the sender can transmit at most a fraction of the total amount of information $\sigma_\theta^2$,

$$\frac{2-\alpha}{2-\alpha c}. \tag{15}$$

We define the *effectiveness of biased communication* as the fraction (15). For $c < 1$ (15) is decreasing in $\alpha$. Hence the sender can transmit less information in an infinite equilibrium if $\alpha$ is higher.

It is now straightforward to address the sender's incentives to acquire information. We know that the upper bound on the equilibrium variability of choices is attained if the sender acquires information of equal use to the receiver and himself, so that $c = 1$. To prove the theorem, it suffices to show that the sender strictly loses from acquiring any information that is intrinsically relatively more useful to himself, hence featuring $c < 1$ for any $\alpha > 1$. Using the upper bound on payoffs, we know that the sender obtains at most a payoff of

$$\mathbb{E}u^s\left(c\tilde{\mu}, \tilde{\eta}\right) = \max_{\substack{\sigma_\theta^2, \sigma_{\omega\theta} \\ s.t. \sigma_\theta^2 \geq \sigma_{\omega\theta}}} c\left(2 - c\right) \frac{2 - \alpha}{2 - \alpha c} \sigma_\theta^2 - \sigma^2, \tag{16}$$

where everything is as in the Laplace case except for the fraction of information that can be transmitted. To determine the upper bound we proceed sequentially. First, we solve for the optimal covariance for a given variance. For $\alpha > 1$, the sender's expected payoff 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 achieve this, the sender becomes perfectly informed about the receiver's ideal action, $\omega$. Second, we optimize over the variance given the optimal covariance. The objective at this stage is

$$\left(2 - \frac{\sigma_{\omega\eta}}{\sigma_\theta^2}\right) \frac{2 - \alpha}{2 - \alpha \frac{\sigma_{\omega\eta}}{\sigma_\theta^2}} \sigma_{\omega\eta} - \sigma^2.$$

The first term is increasing in $\sigma_\theta^2$. All else equal, the sender prefers intrinsically more useful information structures. The second term is decreasing in $\sigma_\theta^2$, so information that is more useful to the sender is harder to transmit. For $\alpha > 1$ the second effect strictly dominates the first one. Hence the sender abstains from acquiring perfect information about $\eta$. As a result, equilibrium communication about $\theta$ is conflict free.

In other words, if biased communication would perform badly, the sender has incentives to be unbiased. As a consequence, the receiver prefers communication over delegation when biased communication would be less effective than in the Laplace world. This proves the theorem, since information structures featuring $c > 1$ remain unattractive by the same arguments as used in our leading case.[25]

---

[25]Note that, in contrast to the Laplace case, the theorem makes a statement about symmetric equilibria

## 8.3 Comparing environments

Communication trumps delegation in all environments where the effectiveness of biased communication is more sensitive to conflicts than in the Laplace case. In these environments (with $\alpha > 1$), communication is notoriously difficult. To illustrate this, we plot the share of the sender's information, $\sigma_\theta^2$, that reaches the receiver if the sender acquires perfect information about his ideal choice $\eta$, $\frac{2-\alpha}{2-\alpha\rho}$, for $\rho \in (0,1)$.
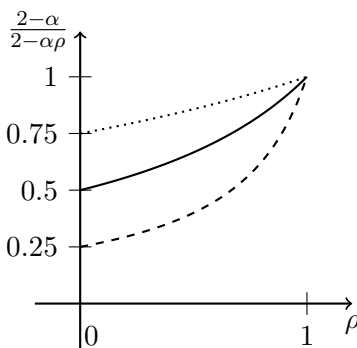


Figure 5: The effectiveness of biased communication in a limit equilibrium for $\alpha = 1.5$ (dashed), $\alpha = 1$ (solid) Laplace, and $\alpha = 0.5$ (dotted) uniform distribution.

We do so for the three cases illustrated in Figure 4.[26] In all three cases, information transmission about the sender's posterior mean becomes perfect as $\rho$ tends to one. Even though there is a countable infinity of induced actions, expected utilities approach the level they reach if $\theta$ is communicated truthfully. At the other extreme, as $\rho$ approaches zero, the amount of information transmitted in an infinite equilibrium tends to the amount that can be transmitted with only two messages, indicating whether $\theta$ is positive or negative.

---

only. The reason is that densities with $\alpha > 1$ are not logconcave, a property of the Laplace that we exploit to prove uniqueness of equilibria inducing a given number of receiver actions.

[26]For the Laplace case and the uniform case the value corresponds to the actual value of communication in a limit equilibrium. We have established existence of a limit equilibrium for the Laplace case, Alonso et al. (2008) have demonstrated the existence of such an equilbrium in the uniform case. For the case $\alpha = 1.5$, we do not claim that a limit equilibrium exists. Our argument is that such an information choice is *not* part of an equilibrium *even if* an equilibrium attaining this value exists.

To verify, a simple variance decomposition delivers the informational content of the binary equilibrium. By symmetry, $\mu_+^2 + v_+^2 = \sigma_\theta^2$. The first term captures the meaning conveyed by the messages, the second one the residual variance after communication. Using the variance expression in Lemma 5, we find that the informational content of binary communication is given by

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

reproducing the intercepts in the figure as claimed: two messaged eliminate 75% of the underlying uncertainty in the uniform case, 50% in the Laplace case, and only 25% for $\alpha = 1.5$. The informational content of binary communication is an exact measure of the sensitivity of the effectiveness of biased communication to conflicts in our model. Since perfect information transmission is possible for $c = 1$ in all environments, environments are more sensitive to conflicts if and only if the effectiveness of biased communication tends to a lower value as $\rho$ tends to zero.

It is now easy to understand when the shadow of communication has a disciplining effect on the sender's information acquisition and when not. In the uniform case, the sender is assured to convey at least 75% of his information to the receiver, no matter what information he acquires. Vice versa, paying attention to the receiver's interest can raise the amount that is transmitted at most by the remaining 25%. It is thus natural that the sender simply acquires information about what he is interested in, $\eta$, and just lives with the consequences. In contrast, eliminating conflicts raises the amount transmitted by a substantial amount for $\alpha > 1$. It is thus the environments where conveying meaning is statistically difficult where the sender can be creadibly punished for acquiring the wrong pieces of information. In these environments, the receiver prefers to communicate, because delegation entails no possibility to punish the sender.

# 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.

Adopting again the Laplace specification, we find that delegation is always the optimal outcome if interests are well aligned to begin with. 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 ideal action only. The sender dislikes communication and acquires information about both ideal actions, to the point where the receiver becomes indifferent between communication and delegation.

**Theorem 5** *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 biased 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.** For quadratic utilities, players' optimal actions are given by the conditional means, $\mathbb{E}\left[\tilde{\omega}\middle| \tilde{s}_\omega = s_\omega, \tilde{s}_\eta = s_\eta\right]$ and $\mathbb{E}\left[\tilde{\eta}\middle| \tilde{s}_\omega = s_\omega, \tilde{s}_\eta = s_\eta\right]$, respectively. We denote $\theta \equiv \mathbb{E}\left[\tilde{\eta}\middle| \tilde{s}_\omega = s_\omega, \tilde{s}_\eta = s_\eta\right]$. For the sender $\mathbb{E}\left[-(y - \tilde{\eta})^2\middle|\tilde{s}_\omega = s_\omega, \tilde{s}_\eta = s_\eta\right] = -y^2 + 2y\mathbb{E}\left[\tilde{\eta}\middle|\tilde{s}_\omega = s_\omega, \tilde{s}_\eta = s_\eta\right] - \mathbb{E}\left[\tilde{\eta}^2\middle|\tilde{s}_\omega = s_\omega, \tilde{s}_\eta = s_\eta\right]$, hence the sender's preferences satisfy the single crossing condition in $y$ and $\theta$.

We now argue that any equilibrium is either fully separating in $\theta$ or involves partial pooling in $\theta$ everywhere. Partial pooling everywhere follows straightforwardly from single

crossing. Suppose there exists an equilibrium where the receiver is fully responsive to $\theta$ on some interval $[\underline{\theta}, \overline{\theta}]$ and involves pooling around the interval. This implies $\mathbb{E}\left[\tilde{\omega} \mid \tilde{\theta} = \theta\right] = \theta$ on $[\underline{\theta}, \overline{\theta}]$. Now take some type $\hat{\theta} = \overline{\theta} + \delta$ for some $\delta > 0$ that induces the pooling action $y_p^r$ strictly above $y^r(\overline{\theta}) = \overline{\theta}$. For $\delta$ sufficiently small $\hat{\theta} - y^r(\overline{\theta}) < y_p^r - \hat{\theta}$, implying that $\mathbb{E}u^s(y^r(\overline{\theta})) > \mathbb{E}u^s(y_p^r)$. Hence sender types close to $\overline{\theta}$ have an incentive to deviate, so the receiver's choice of actions does not constitute an equilibrium. Since the same argument holds for $\underline{\theta}$, the receiver cannot be fully responsive to $\theta$ on a bounded interval. Hence, an equilibrium that involves separation in $\theta$ somewhere must involve separation in $\theta$ everywhere.

Clearly, these two classes of equilibria can be characterized by communication about $\theta$ only. Consider now an equilibrium where some sender types with the same conditional expectation play different strategies. By the single crossing property, types with the same $\theta$ are indifferent between at most two distinct actions. It follows immediately from single crossing that this corresponds to the second class of equilibria considered above, where the strategies are changed on measure zero sets. Since this does not change the receiver's equilibrium actions, we can characterize such equilibria – up to the strategies of sender types on measure zero sets – by communication about $\theta$ only. ∎

**Proof of Lemma 2.**    i) The random vector $\tilde{\boldsymbol{\tau}} = (\tilde{\omega}, \tilde{\eta}, \tilde{\varepsilon}_\omega, \tilde{\varepsilon}_\eta)$ 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 $\theta$ can be calculated as $\mathbb{E}\left[\tilde{\eta} \mid \tilde{s}_\omega = s_\omega, \tilde{s}_\eta = s_\eta\right] = \gamma_\omega s_\omega + \gamma_\eta s_\eta$ with $\gamma_\omega = \frac{\sigma_{\varepsilon_\eta}^2 \rho \sigma^2}{(\sigma^2 + \sigma_{\varepsilon_\omega}^2)(\sigma^2 + \sigma_{\varepsilon_\eta}^2) - \rho^2 \sigma^4}$, and $\gamma_\eta = \frac{\sigma^2 \sigma_{\varepsilon_\omega}^2 - \sigma^4(1 - \rho^2)}{(\sigma^2 + \sigma_{\varepsilon_\omega}^2)(\sigma^2 + \sigma_{\varepsilon_\eta}^2) - \rho^2 \sigma^4}$; the weights $\gamma_\omega$ and $\gamma_\eta$ 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 $\left(\tilde{\omega}, \tilde{\eta}, \tilde{\theta}\right)$ 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 $\tilde{\theta}$ is zero, because the mean of $\tilde{\boldsymbol{\tau}}$ is the zero vector. Plugging in the

weights $\gamma_\omega$ and $\gamma_\eta$, the second moments of $\left(\tilde{\omega}, \tilde{\eta}, \tilde{\theta}\right)$ can straightforwardly be calculated:

$$\sigma_\theta^2 = \gamma_\omega^2 Var\left(\tilde{s}_\omega\right) + \gamma_\eta^2 Var\left(\tilde{s}_\eta\right) + 2\gamma_\omega\gamma_\eta Cov\left(\tilde{s}_\omega, \tilde{s}_\eta\right)$$

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

$$= \sigma_{\eta\theta},$$

and

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

**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 $\sigma_\theta^2$ as

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

and

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

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

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

and

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

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

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

34

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 $C \in [0, \sigma_{\omega\eta})$, $\sigma_\theta^2 \in \left[\rho C, \frac{1}{\rho}C\right]$. 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 $0 \le \underline{\theta} < \overline{\theta}$ we can write

$$\mathbb{E}\left[\tilde{\theta} \middle| \tilde{\theta} \in [\underline{\theta}, \overline{\theta}]\right] = \mathbb{E}\left[\tilde{\theta} \middle| \tilde{\theta} \ge 0\right] + \overline{\theta} - g\left(\overline{\theta} - \underline{\theta}\right), \tag{17}$$

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

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

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

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

The slope of the function is

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

35

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

Differentiating $g(q)$ twice, we obtain

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

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

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

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,

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

**iii)** The marginal distribution of $\tilde{\theta}$ 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 $\mathbb{E}\left[ \tilde{\theta} \middle| \theta \in [\underline{\theta}, \overline{\theta}] \right] = \frac{1}{\lambda} + \overline{\theta} - \frac{(\overline{\theta} - \underline{\theta})}{1 - \exp^{-\lambda(\overline{\theta} - \underline{\theta})}} = \frac{1}{\lambda} + \frac{\underline{\theta}}{1 - \exp^{-\lambda(\overline{\theta} - \underline{\theta})}} - \frac{\exp^{-\lambda(\overline{\theta} - \underline{\theta})} \overline{\theta}}{1 - \exp^{-\lambda(\overline{\theta} - \underline{\theta})}}$ and $\lim_{\overline{\theta} \to \infty} \exp^{-\lambda(\overline{\theta} - \underline{\theta})} \overline{\theta} = 0$, we have

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

For a discussion of the parameter $\alpha$ 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.
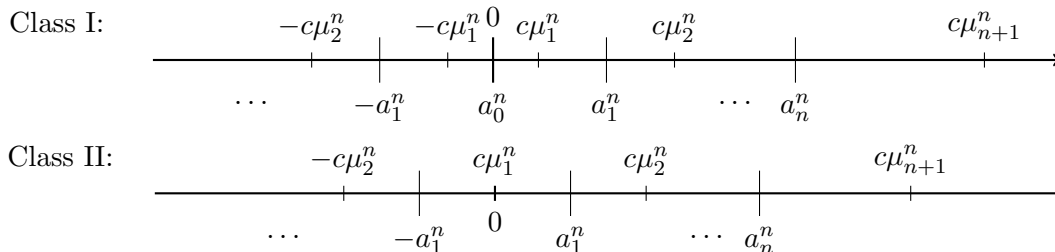


Figure 6: Class I equilibrium and Class II equilibrium.

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

$$\mu_i^n \equiv \mathbb{E}\left[\tilde{\theta}\,\Big|\,\tilde{\theta} \in \left[a_{i-1}^n, a_i^n\right)\right] \quad \text{for } i = 1, \ldots, n \tag{18}$$

and $\mu_{n+1}^n \equiv \mathbb{E}\left[\tilde{\theta}\,\Big|\,\tilde{\theta} \geq a_n^n\right]$. 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

$$y\left(a_{i-1}^n, a_i^n\right) = c \cdot \mu_i^n \quad \text{for } i = 1, \ldots, n \tag{19}$$

and $y\left(a_n^n, \infty\right) = c \cdot \mu_{n+1}^n$ 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 $a_1^n, \ldots, a_n^n$:

$$a_i^n - c \cdot \mu_i^n = c \cdot \mu_{i+1}^n - a_i^n, \quad \text{for } i = 1, \ldots, n. \tag{20}$$

37

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

Equation (20) 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 (20) exists - depend crucially on the magnitude of the regression coefficient, $c$.

For $c \leq 1$, there is no bound on the number of induced actions (see Proposition 1). 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. Recall the definition of the $g$ function from Lemma A1.

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

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

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

$$cg\left(a_n^n - a_{n-1}^n\right) = 2\frac{c}{\lambda} + 2\left(c - 1\right) a_n^n, \tag{22}$$

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

$$a_1^n = \frac{c}{\lambda} + c\left(a_2^n - a_1^n\right) - cg\left(a_2^n - a_1^n\right) - \left(1 - c\right) a_1^n, \tag{23}$$

and in addition (21) for $i = 2, \ldots, n - 1$, and (22).

Proof: Recall the proof of Lemma A1; we write the conditional mean for the Laplace $\mu_{i+1} = \mathbb{E}\left[\tilde{\theta}\middle|\tilde{\theta} \in [\theta_i, \theta_{i+1}]\right] = \mathbb{E}\left[\tilde{\theta}\middle|\tilde{\theta} \geq 0\right] + \theta_{i+1} - g\left(\theta_{i+1} - \theta_i\right)$, where $0 \leq \theta_i < \theta_{i+1}$, $g\left(q\right) = \frac{q}{1 - \exp(-\lambda q)}$, and $\frac{1}{\lambda} = \mathbb{E}\left[\tilde{\theta}\middle|\tilde{\theta} \geq 0\right]$. In combination with the sender's indifference conditions (20), $a_i^n - c \cdot \mu_i^n = c \cdot \mu_{i+1}^n - a_i^n$, this implies the claim.

**Part i)** We use the combination of forward solution and condition (22) 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

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

39

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 $a_2\left(x\right)$ satisfies $\lim_{x\to 0}\left(a_2\left(x\right) - x\right) = 0$ and $\frac{da_2}{dx} > 1$, implying that $a_2\left(x\right) - x$ is increasing in $x$. Moreover, the forward solutions $a_i\left(x\right) - a_{i-1}\left(x\right)$

40

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 (21) 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 $c \le 1$. 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 $a_{i+1}(x)$,

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

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

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

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

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

Since $cg(x) - \frac{c}{\lambda}$ is nonnegative and increasing in $x$ and $\frac{c}{\lambda} + 2(c-1)x$ is positive for $x = 0$ and nonincreasing in $x$, there exists a unique value $x_2^*$ such that (24) is satisfied with equality. Hence, a finite forward solution $a_2(x)$ exists for all $x \in [0, x_2^*)$. In the limit as $x \to x_2^*$, we have $\lim_{x \to x_2^*} a_2(x) = \infty$.

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

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

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

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

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

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

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

$$cg(a_{i-1}(x) - a_{i-2}(x)) - \frac{c}{\lambda} < \frac{c}{\lambda} + 2(c-1)a_{i-1}(x). \tag{25}$$

Given the derived properties of the forward solution, we have that $cg(a_{i-1}(x) - a_{i-2}(x)) - \frac{c}{\lambda}$ is nonnegative and increasing in $x$ and $\frac{c}{\lambda} + 2(c-1)a_{i-1}(x)$ is positive for $x = 0$ and nonincreasing in $x$. Therefore, there exists a unique value $x = x_i^*$ such that (25) is satisfied with equality. Hence a finite forward solution $a_i(x)$ exists for all $x \in [0, x_i^*)$. In the limit as $x \to x_i^*$, we have $\lim_{x \to x_i^*} a_i(x) = \infty$.

Define

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

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

**Claim i.4)** For any $n$ there exists a unique value of $\tilde{x}_n$ such that condition (22) holds for $a_{n-1}$ and $a_n$ defined as solutions to the forward equation.

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 the condition (22), 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

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

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

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 the fixed point equation

$$\Delta_n(x) = 0.$$

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, $\tilde{x}_n$. So in equilibrium $a_1^n = \tilde{x}_n$.

**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 (21) for $i = 2, \ldots, n-1$ and condition (22).

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

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

44

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 $c \le 1$, 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
(22) 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,

$$a_1^{n+1} < a_1^n < a_2^{n+1} < \cdots a_n^{n+1} < a_n^n < a_{n+1}^{n+1} \qquad \forall n. \tag{26}$$

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:

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

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

45

and

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

The conclusion follows from the fact that $a_i^k\left(x\right)$ 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 $a_i^k\left(x\right) - a_{i-1}^k\left(x\right)$.

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

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

and

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

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

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

However, if $a_{l+2}^{n+1}\left(\tilde{x}_{n+1}\right) < a_{l+1}^n\left(\tilde{x}_n\right)$ and $a_{l+2}^{n+1}\left(\tilde{x}_{n+1}\right) - a_{l+1}^{n+1}\left(\tilde{x}_{n+1}\right) < a_{l+1}^n\left(\tilde{x}_n\right) - a_l^n\left(\tilde{x}_n\right)$, then $a_{l+3}^{n+1}\left(\tilde{x}_{n+1}\right) < a_{l+2}^n\left(\tilde{x}_n\right)$ and so forth. Hence, we would have $a_{j+1}^{n+1}\left(\tilde{x}_{n+1}\right) < a_j^n\left(\tilde{x}_n\right)$ and $a_{j+1}^{n+1}\left(\tilde{x}_{n+1}\right) - a_j^{n+1}\left(\tilde{x}_{n+1}\right) < a_j^n\left(\tilde{x}_n\right) - a_{j-1}^n\left(\tilde{x}_n\right)$ 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 $\left(a_1^n\right)_n$ is monotone decreasing in $n$. Since the sequence is bounded by zero it must converge. Similarly, by Claim ii.1) the sequence $\left(a_n^n\right)_n$ is monotone

increasing in $n$. The fixed point condition, (22), implies that it is bounded by $\frac{c}{1-c}\frac{1}{2\lambda}$, hence it converges. Since equilibrium thresholds are nested (cf. (26)) 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 $\lim_{n\to\infty} c\mu_i^n \le \lim_{n\to\infty} a_i^n \le \lim_{n\to\infty} c\mu_{i+1}^n$. 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 $\bar{a}_i = \lim_{n\to\infty} a_i^n$ for all $i$. By convergence, for any $\varepsilon$ there is a $N$ such that for all $n > N$: $a_i^n \ge \bar{a}_i - \frac{\varepsilon}{2}$ and $a_{i+1}^n \le \bar{a}_{i+1} + \frac{\varepsilon}{2}$. Suppose for contradiction that $\bar{a}_i \ge \bar{a}_{i+1} + \delta$ for some $\delta > 0$; this implies

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

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 $x \le x_n^*$, where $x_n^*$ satisfies

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

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 $c \le 1$ the following. For any $x > 0$ there is a $k$ such that

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

and

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

Therefore, for a fixed length $x$ of the first interval, the forward equation has a solution only for a finite number of steps. 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, $\lim_{n\to\infty} \tilde{x}_n = 0$.

47

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. Recall the definition of the $g$ function from Lemma A1.

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

$$a_{i+1}^n - a_i^n > a_i^n - a_{i-1}^n; \tag{27}$$

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,

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

and

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

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

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

Since $g(z) > z$ 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

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

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

48

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

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

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

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

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

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,

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

and

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

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: 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

$$cg(x) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_2 - x) - cg(a_2 - x) + 2(c-1)x. \tag{28}$$

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 $x < \bar{x}$ where $\bar{x}$ is the unique value of $x$ that satisfies $cg(\bar{x}) - \frac{c}{\lambda} = \frac{c}{\lambda} + 2(c-1)\bar{x}$. 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 $x \in (\underline{x}, \bar{x})$ 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.

50

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 (27), 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 (28) 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 (28) 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 (27). 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

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

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 (29) has a solution for $x < \frac{c}{\lambda(2-c)}$, which satisfies $\lim_{x \to 0} a_2(x) = 0$ and moreover,

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

51

Rearranging (29) we can write

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

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

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

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

$$0=\frac{c}{\lambda}+c\left(a_3-a_2\left(x\right)\right)-cg\left(a_3-a_2\left(x\right)\right)+4\frac{c-1}{2-c}\left(a_2\left(x\right)-x\right)-\frac{c}{2-c}\left(cg\left(a_2\left(x\right)-x\right)-\frac{c}{\lambda}\right). \tag{31}$$

Note that the right-hand side of this equation is increasing in $a_3$ and that $a_3\left(x\right)$ is the unique value that sets the expression equal to zero. We show that the expression is strictly positive for $a_3-a_2\left(x\right)=a_2\left(x\right)-x$, to get $a_3\left(x\right)-a_2\left(x\right)<a_2\left(x\right)-x$, in contradiction to the increasing interval property (27).

Note that the right-hand side of (31) depends only on the differences $a_2\left(x\right)-x$ and $a_3-a_2\left(x\right)$. Moreover, note that $a_2\left(x\right)-x$ goes to zero as $x$ goes to zero. Let $z=a_2\left(x\right)-x$ and evaluate the rhs of (31) at $a_3-a_2\left(x\right)=z$. We obtain

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

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

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

where we use that $g'\left(z\right)|_{z=0}=\frac{1}{2}$. For $c\in\left(1,2\right)$, we have $5c-c^2-4>0$ and we know that $F\left(z\right)>0$ for $z$ small. Since, the right-hand side of (31) is increasing in $a_3$, to restore equality with zero, $a_3$ needs to decrease, which would imply that $a_3\left(x\right)-a_2\left(x\right)<a_2\left(x\right)-x$. However, this contradicts the the increasing interval property (27) 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 $\Pr\left(\tilde{\theta}\in\left(0,x\right]\right)=\frac{1}{2}\left(1-e^{-\lambda x}\right)$ and $\Pr\left(\tilde{\theta}\in\left(-y,0\right]\right)=$

$\Pr\left(\tilde{\theta} \in [0, y)\right) = \frac{1}{2}\left(1 - e^{-\lambda y}\right)$. 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

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

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

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

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

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

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,

$$cg(a_2(x, y) - x) - \frac{c}{\lambda} = \frac{c}{\lambda} + c(a_3 - a_2(x, y)) - cg(a_3 - a_2(x, y)) + 2(c - 1)a_2(x, y). \tag{33}$$

Rearranging (32), we can write

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

Adding up with $(33)$,

$$\frac{2(c - 1)}{(c - 2)}cw(x, y)$$
$$= \frac{c}{\lambda} + c(a_3(x, y) - a_2(x, y)) - cg(a_3(x, y) - a_2(x, y)) + 4\frac{c - 1}{2 - c}(a_2(x, y) - x)$$
$$- \frac{c}{2 - c}\left(cg(a_2(x, y) - x) - \frac{c}{\lambda}\right).$$

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

53

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 (27) for $n$ large enough.

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

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

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

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

If $D\left(z; x\right) > 0$, then $a_n$ needs to decrease to satisfy the forward equation. Note that $\frac{2}{\lambda} + z - 2g\left(z\right)$ is strictly negative for $z > 0$ and $2\left(c - 1\right) a_{n-1}\left(x\right)$ 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 (27) for $a_2\left(x\right) - x > x$. Suppose that the increasing interval property is satisfied up to the interval $a_{n-1}\left(x\right) - a_{n-2}\left(x\right)$. (If not, then we are done already.) If all intervals up to $a_{n-1}\left(x\right) - a_{n-2}\left(x\right)$ satisfy the increasing interval property, then $a_{n-1}\left(x\right) \geq \left(n - 1\right) x$. Note that $x$ does not depend on $n$. Hence, for any finite $z$, there is a $n\left(z, x\right)$ such that $D\left(z; x\right) > 0$ for all $n \geq n\left(z, x\right)$, 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 $a_{n-1}\left(x\right)$ and the lower bound on $x$. However, $a_{n-1}\left(x\right) \geq x + \left(n - 2\right)\left(a_2\left(x\right) - x\right)$. Note again that $x$ and $a_2\left(x\right)$ 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

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

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

$$\mathbb{E}_{\tilde{\mu}\tilde{\omega}} \left[ \tilde{\omega}\tilde{\mu} \right] = \mathbb{E}_{\tilde{\mu}} \left[ \mathbb{E}_{\tilde{\omega}|\tilde{\mu}=\mu} \left[ \tilde{\omega}\tilde{\mu} | \tilde{\mu} = \mu \right] \right] = \sum_j \Pr(\Theta_j) \left[ \mathbb{E}_{\tilde{\omega}|\tilde{\mu}=\mu_j} \left[ \tilde{\omega}\tilde{\mu} | \tilde{\mu} = \mu_j \right] \right]$$

$$= \sum_j \Pr(\Theta_j) \mu_j \int \omega f_{\tilde{\omega}|\Theta_j} \left( \omega | \tilde{\theta} \in \Theta_j \right) d\omega,$$

where

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

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

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

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

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

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 $\tilde{\theta}$. For $k = 2, \ldots, n$, define $\hat{p}_{k-1}$ as the probability that $\tilde{\theta} \in [a_{k-2}, a_{k-1}]$ conditional on $\tilde{\theta} \geq a_{k-2}$,

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

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

$$\hat{p}_{k-1} = \frac{\mathbb{E}\left[\tilde{\theta}\middle| \tilde{\theta} \geq a_{k-1}\right] - \mathbb{E}\left[\tilde{\theta}\middle| \tilde{\theta} \geq a_{k-2}\right]}{\mathbb{E}\left[\tilde{\theta}\middle| \tilde{\theta} \geq a_{k-1}\right] - \mu_{k-1}} \tag{34}$$

and

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

To see this, note that

$$(F(a_{k-1}) - F(a_{k-2})) \mu_{k-1} = \int_{a_{k-2}}^{a_{k-1}} \theta f(\theta)\, d\theta = \int_{a_{k-2}}^{\infty} \theta f(\theta)\, d\theta - \int_{a_{k-1}}^{\infty} \theta f(\theta)\, d\theta$$

$$= (1 - F(a_{k-2})) \mathbb{E}\left[\tilde{\theta}\middle| \tilde{\theta} \geq a_{k-2}\right] - (1 - F(a_{k-1})) \mathbb{E}\left[\tilde{\theta}\middle| \tilde{\theta} \geq a_{k-1}\right].$$

Hence

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

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 $\tilde{\theta} \in [a_{k-2}, a_{k-1}]$ conditional on $\tilde{\theta} \geq a_{k-3}$, and $(1 - \hat{p}_{k-2}) \cdot (1 - \hat{p}_{k-1})$ is the probability of the event $\tilde{\theta} \geq a_{k-1}$ conditional on $\tilde{\theta} \geq a_{k-3}$. To see this, note that $1 - \hat{p}_{k-2} = \Pr\left[\tilde{\theta} \geq a_{k-2}\middle| \tilde{\theta} \geq a_{k-3}\right] = \frac{1 - F(a_{k-2})}{1 - F(a_{k-3})}$ 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 $\mu_+ \equiv \mathbb{E}\left[\tilde{\theta}\middle|\tilde{\theta} \geq 0\right]$. Let the distribution satisfy $\mathbb{E}\left[\tilde{\theta}\middle|\tilde{\theta} \geq \bar{\theta}\right] = \mu_+ + \alpha \cdot \bar{\theta}$ for all $\bar{\theta} \geq 0$ and for some constant $\alpha$. 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 $\mu_+$, conditional on $\tilde{\theta} \geq a_{n-1}^n$. Substituting for $\hat{p}_n^n$ from $(34)$, and multiplying and dividing by $\hat{c}$ for convenience, we can write

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

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

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

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 $\alpha$), $\hat{c}\mu_n^n + \hat{c}\mu_{n+1}^n = 2\alpha a_n^n$ 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 $\mu_{n+1}^n$, we obtain

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

57

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

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

**Induction hypothesis:**

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

**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 (34) and using the inductive hypothesis, we have

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

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

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

58

and

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

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

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,

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

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

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

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

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

$$\left. + \left(\hat{c}\mathbb{E}\left[\tilde{\theta}\middle|\tilde{\theta} \geq a_{k-1}^n\right] - \hat{c}\mu_+\right)\left(\frac{\hat{c}}{2 - \hat{c}}\left(\hat{c}\mu_{k-1}^n + \hat{c}\mu_+\right) + \frac{4}{2 - \hat{c}}\left(\alpha a_{k-1}^n - \hat{c}\mu_{k-1}^n\right)\right)\right).$$

It is easy to see that

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

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

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

Moreover, since $\hat{c}\mathbb{E}\left[\tilde{\theta}\middle|\tilde{\theta}\geq a_{k-1}\right] - \hat{c}\mu_+ = \alpha\hat{c}a_{k-1}^n$, 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

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

This completes the induction.

Building on the characterization, we can compute $\mathbb{E}\left[\tilde{\mu}^2\right]$ 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 $\lambda$ 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, $\Pr\left[\tilde\theta \geq 0\right] = \Pr\left[\tilde\theta \leq 0\right] = \frac{1}{2}$ and $\mathbb{E}\left[\tilde\mu^2 \middle| \tilde\theta \geq 0\right] = \mathbb{E}\left[\tilde\mu^2 \middle| \tilde\theta \leq 0\right]$, so that

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

Hence, we have shown that in a Class I equilibrium

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

**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

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

Hence

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

Now, we may write

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

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 $\mathbb{E}\left[\tilde{\mu}^2\big|\tilde{\theta} \geq a_1^n\right] = \mathbb{E}\left[\tilde{\mu}^2\big|\tilde{\theta} \leq -a_1^n\right]$ and $\Pr\left[\tilde{\theta} \geq a_1^n\right] = \Pr\left[\tilde{\theta} \leq -a_1^n\right]$. Hence, we have shown that

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

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

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

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

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

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

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

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 $x > 0$. ∎

62

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

$$\mu_+ + \alpha\theta = \theta + \frac{\int_\theta^{\bar{\theta}} \left(1 - \hat{F}(t; \alpha)\right) dt}{1 - \hat{F}(\theta; \alpha)}. \tag{35}$$

Define $q(\theta) = \int_\theta^{\bar{\theta}} \left(1 - \hat{F}(t; \alpha)\right) dt$ 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 (35) 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 $\lim_{\theta \to \bar{\theta}} \hat{F}(\theta; \alpha) = 1$, we have $\bar{\theta} = \frac{\mu_+}{1-\alpha}$ for $\alpha < 1$. For $\alpha \geq 1$, the support is $\mathbb{R}^+$. Differentiating twice, we obtain the density

$$\hat{f}(\theta; \alpha) = \alpha \left(\mu_+\right)^{-\frac{\alpha}{1-\alpha}} \left(\mu_+ - \theta(1-\alpha)\right)^{\frac{2\alpha-1}{1-\alpha}}. \tag{36}$$

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 $\mathbb{E}\left[\tilde{\mu}^2\right]$ in any symmetric equilibrium. For Class

I equilibria this is obvious, so consider Class II equilibria. Note that $\Pr\left[\tilde{\theta} \in [-a_1^n, a_1^n)\right] = \hat{F}(a_1^n; \alpha)$. Moreover,

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

Hence,

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

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.