Communication in the Shadow of Catastrophe*

Inga Deimen
University of Arizona
and CEPR

and

Dezső Szalay
University of Bonn
and CEPR

August 2022

Abstract

We study the role of risk in strategic information transmission. We show that an increased likelihood of extreme states – heavier tails – decreases the amount of information transmission and makes it optimal to alter the mode of decision-making from communication to simple delegation. Moreover, the worst-case losses under communication increase relative to the worst-case losses under delegation when the tails get heavier.

JEL: D83, D82

Keywords: strategic communication, delegation, authority, organizations, risk, extreme events, Generalized Pareto distribution

---

*Deimen: University of Arizona and CEPR, Eller College of Management, University of Arizona, 1130 E. Helen St, Tucson, AZ 85721, ideimen@arizona.edu. Szalay: University of Bonn and CEPR, Institute for Microeconomics, University of Bonn, Adenauerallee 24-42, 53113 Bonn, Germany, szalay@uni-bonn.de.
We have benefited enormously from discussions with numerous colleagues and audiences at Aalto, Arizona, Austin, Chicago, Cornell, Cowles foundation, Duke, ESSET (Gerzensee), Florida State, Hamburg, HCM (Bonn), MIT, Munich, Oxford, Paris, Queen Mary, SAET, SET, SIOE, Texas A&M, Tokio, Turin, UCLA, VfS org econ meeting, VSET, Washington U. We would like to thank in particular Wouter Dessein, Amanda Friedenberg, Roger Myerson, and Ernesto Rivera Mora for helpful comments. Financial support through the Hausdorff Center of Mathematics is gratefully acknowledged. All remaining errors are our own.

1 Introduction

One question that arises naturally in organizational design is whether an expert should report to a decision maker (communication) or should the expert decide (delegation). Dessein (2002) shows that delegating decision rights to the expert dominates strategic communication, as introduced in Crawford and Sobel (1982), whenever the level of conflict is small.

This paper reconsiders the question when the decision at hand involves the potential of catastrophes. The catastrophes that we have in mind are human-made outcomes (in contrast to natural disasters) with extremely low payoffs for the decision maker. They stem from a combination of exceptional events and disagreement on what is the right action to take. Examples for such catastrophes are the oil drilling blowout on the *Deepwater Horizon*, where *BP* made decisions against expert advice (see details below); the *Challenger* space shuttle explosion, where engineers warned in vain about potential problems arising from the low temperatures; or the flooding of the Ahr valley, where officials delayed the evacuation despite experts' warnings of an extreme rise of the water level.

Our goal is to address how an increase in the likelihood of exceptional events changes the comparison between communication and delegation. As a natural first approach, we increase the variance to scale up the support (and thereby the likelihood of exceptional events). This does not change the comparison but only scales all payoffs down in our model. Then, controlling for variance, we take a novel approach by looking at *tail risk* – the shape of the distribution (essentially kurtosis). This not only enlarges the support but also alters probability weights in the distribution. We show that increasing tail risk is detrimental to communication but not to delegation. As a consequence, we advocate delegating decision rights to experts, in the shadow of catastrophic outcomes.

Intuitively, an increase in tail risk for a fixed variance has two effects: Exogenously, it results in more mass in the peak of the distribution as well as in the tails. Endogenously, it moves equilibrium choices relatively closer to the mean. Because sender and receiver are in agreement around the mean in our model, communication works very well for these states. More weight here thus does not improve the payoffs substantially. By contrast, the increased likelihood of extreme states is very costly. Naturally, it is very difficult to communicate whether a state is 'extreme' or 'very extreme'. In addition, in our model the disagreement between sender and receiver increases in the state which makes communication even harder and messages at the extremes very coarse: actions induced by communication are far away from the ideal action. Convex losses disproportionately penalize large distances, implying extreme losses for extreme states.

To formalize this intuition, we design a model of organizational decision making and communication that has the following features: i) The receiver can choose between communicating with the sender (cheap talk, Crawford and Sobel (1982)) or fully delegating the decision to the sender. ii) Sender and receiver have unique optimal actions given the realization of the state and suffer losses that are quadratic in the distance to the respective optimum. iii) Sender and receiver agree on the optimal action in one state, the farther away from that agreement point the more they disagree (state dependent, multiplicative bias). iv) The state is symmetrically distributed around the agreement point, has a logconcave density, and a finite variance. We discuss our modeling choices in Section 2.1, after formally introducing the model.

We first prove that communication equilibria – taking the typical form of an interval partition – exist and are essentially unique (Proposition 1). Then, assuming only symmetry and fixing the shape of the density, we show that an increase in variance scales communication as well as delegation payoffs down. A change in risk in terms of variance will thus never change the optimal decision mechanism in our model (Theorem 1).

To vary the tail risk of the distribution (without changing the variance), we focus on the two-sided generalized Pareto distribution which has two separate parameters for scale (variance) and shape (tail risk). Increasing the shape parameter results in more mass in the peak of the distributions as well as in the tails, varying from the uniform to the exponential distribution. The value of communication is decreasing in tail risk (Proposition 2), while the payoffs under delegation are unaffected. Moreover, relative to the support, higher tail risk moves equilibrium thresholds and actions closer to the prior mean (Proposition 3): communication becomes more precise around the agreement point but leaves large pooling intervals at the extremes. Delegation becomes relatively better than communication in environments with higher tail risk (Theorem 2).

Debates in the aftermath of a catastrophe often center around the question whether the outcome was avoidable. To address this question, we offer a worst-case analysis which complements the usual ex ante perspective in the literature. Interestingly, we find a systematic difference between ex ante and worst-case optimality in favor of delegation (Theorem 3): delegation dominates communication from a worst-case perspective in cases in which the optimal mode of decision making from an ex ante perspective is communication.

Finally, we study the value of communication for the general class of symmetric, logconcave distributions. The uniform conditional variability order (Whitt (1985)) is a partial order that allows to study risk in terms of variation in the shape of the distribution while keeping the variance fixed.[1] We prove that for sufficiently large conflicts, if the tails of two distributions are ordered by the uniform conditional variability order, the value of communication is lower for the distribution with the more variable, risky tail (Proposition 5). Thus higher risk makes delegation relatively more attractive.

The following well-documented example illustrates our link between catastrophic decision making based on communication and changes in the environment the organization operates in, in terms of risk.[2] *British Petroleum* (*BP*) was drilling for oil from the rig *Deepwater Horizon* in the *Macondo* well in the Gulf of Mexico, when in 2010 a blowout occurred. The outcome was catastrophic: 11 people lost their lives, the largest offshore oil spill in the U.S. history materialized, and *BP* paid about \$20 billion claims for compensation and lost nearly \$100 billion in stock market value.

BP, the owner of the drilling rights, relied on a subcontractor, *Transocean*, to perform the drilling. *PB* (the receiver) directed the work, *Transocean* (the sender) provided advice, the drilling rig, and the crew operating it. *BP* and *Transocean* had agreed on a budget and a timeline (p.2). Every deviation from the planned procedure

[1] Note that orders like mlrp, fosd, sosd do not allow to keep the variance fixed.

[2] The information provided here is based on the report of the National Comission on the BP Deepwater Horizon oil spill and offshore drilling (2011).

was costly to *BP* with costs increasing in the length of the resulting delay ("leasing its [the *Deepwater Horizon* rig's] services reportedly cost as much as $1 million per day." p.2) As a consequence, *BP* responded conservatively to proposed changes by *Transocean* and made decisions that deviated from the recommended actions (p.125). The report states that "Most, if not all, of the failures at *Macondo* can be traced back to underlying failures of management and communication." (p.122).[3]

Notably, the oil drilling industry prior to the incident is characterized by a huge change in production and risk. While in 1990 4.4% of the oil production in the Gulf of Mexico originated from deepwater wells, the number increased to 80% in 2009. ("Deepwater energy exploration and production, particularly at the frontiers of experience, involve risk for which neither industry nor government has been adequately prepared [...]." p.9)

A central question during the congressional hearings was whether *BP* had dealt adequately with the risk and whether the catastrophic outcome was avoidable. Our model casts new light on these questions: it is indeed possible that the ex ante optimal choice of communication produces looses that appear avoidable in hindsight; we find a systematic discrepancy in favor of delegation between ex ante optimality and minimizing worst case losses.

Dessein (2002) is the first to study the allocation of decision authority in the setup of Crawford and Sobel (1982). Dessein shows that whenever influential communication is possible at all, the receiver prefers to delegate: for small biases, the loss of control under delegation weights less severely than the loss of information through strategic communication. Considering risk in his specification with uniform distribution and additive bias implies a focus on variance. Dessein points out that for a higher variance the receiver prefers to delegate for a larger set of biases. Intuitively, an increase in the variance – by increasing the support of the distribution –

[3] A catastrophe in this paper corresponds to a catastrophically low payoff. Such catastrophes are human made, stemming from disagreement in exceptional situations. There are other situations in which a catastrophe is commonly anticipated – imagine a meteorite coming earth's way. Such a situation could arguably align interests perfectly. While this is an important and interesting situation, we focus on human made catastrophes in this paper leaving other cases to future work.

decreases the relative bias between sender and receiver.[4]

We here assume a multiplicative bias that is independent of changes in the variance; sender and receiver have an agreement point (at zero) and conflicts linearly increase in the state. Communication with linear conflicts has been investigated first by Melumad and Shibano (1991), and more recently by Alonso et al. (2008), Rantakari (2008), and Deimen and Szalay (2019).

Alonso et al. (2008) and Rantakari (2008) study organizations with multiple divisions in need of adaptation to private information and of coordination with each other. Imperfect profit sharing in their models provides a micro foundation for linear conflicts. Moreover, since all their payoffs are linearly decreasing in variance, the optimal choice of the organizational structure does not depend on the variance.

More recent work shows that these organizational theories provide rich predictions on the impact of variance when other factors are taken into account as well. Rantakari (2013) allows firms to choose the compensation and the authority structure jointly. He finds that firms that operate in volatile environments are characterized by decentralized decision making and a compensation with focus on performance at the division level. Dessein et al. (forthcoming) provide a theoretical model that predicts that an environment that is more volatile locally results in more decentralized decision making only when the need for coordination across sub-units is low. Our predictions are in line with these observations. In addition to the analysis of scaling risk (changes in variance), we provide novel insights on risk in terms of the shape of the distribution; this has so far only received attention in other fields (Artzner et al. (1999) (actuarial sciences), Whitt (1985) (operations research)).

Deimen and Szalay (2019) focuses on the allocation of authority with endogenous information: the paper looks at incentives for information acquisition by an initially uninformed sender. Conflicts between the sender and the receiver arise depending on the type of information that the sender acquires.[5] Information is noisy and multidimensional. Since communication under conflicts works very badly in fat-tailed environments (featuring logconvex tails), the sender prefers to acquire information

[4]See Liu and Migrow (2019), for an analysis of volatility in a model of disclosure.

[5]See also Antić and Persico (2020) for a communication model with endogenous preference-based conflicts.

that aligns incentives completely. In the current paper, conflicts are exogenous and the environment features logconcave tails.

Chen and Gordon (2015) study the effect of more aligned preferences in terms of bias and/ or prior. They show that information transmission is improved when ideal choices are closer. This is satisfied when the distributions are ordered by the monotone likelihood ratio order (mlrp). In our setup, distributions cannot be ordered by mlrp. Instead, the likelihood ratio on the half-supports is unimodal and the distributions are ordered by the conditional variability order (Whitt (1985)). For sufficiently pronounced conflicts, the order translates to mean-preserving spreads of the equilibrium actions on the entire support, which implies improved information transmission.

The remainder of the paper is organized as follows. We present our formal model in Section 2. Payoffs and equilibria of the communication game are derived in Section 3. In Section 4, we study the impact of variance. We introduce the generalized Pareto distribution and tail risk in Section 5, and study the impact on communication. In Section 6, we derive the optimal choice of decision making. We provide intuition in Section 7. A worst case analysis is done in Section 8. In Section 9, we analyze tail risk and communication in a more general environment. Finally, Section 10 concludes. All proofs are in the appendix.

2 Model

We consider a game with two players, a sender S and a receiver R. Sender and receiver have quadratic payoffs

$$u_S(y,\theta) = -(y-\theta)^2 \quad \text{and} \quad u_R(y,\theta,\beta) = -(y-\beta\cdot\theta)^2$$

that depend on an action $y \in \mathbb{R}$, on the realization θ of state of the world Θ, and on a parameter $\beta \in (0,1)$ that determines the conflict of interest between the players. The ideal choice functions of sender and receiver are $y_S(\theta) = \theta$ and $y_R(\theta) = \beta \cdot \theta$, respectively. The parameter β thus induces a state dependent bias of $(1-\beta)\cdot\theta$.

The state of the world Θ is a random variable with a common prior distribution F with density f on an appropriate interval support $\mathcal{S} \subseteq \mathbb{R}$. We assume that the density is symmetric, logconcave, and that the mean is zero and the variance σ^2 is finite.

The sender privately learns the realization of the state θ. The receiver can choose to rely on communication with the sender (*communication*). In this case, a sender strategy maps states into distributions over messages, $M_S : \mathcal{S} \to \Delta M$; and a receiver strategy maps messages into actions $Y_R : M \to \mathbb{R}$. Strict concavity of payoffs implies that a restriction to pure receiver strategies is without loss of generality. As a simple alternative, the receiver can choose to delegate decision-making to the sender (*delegation*) in which case a sender strategy maps states into actions, $Y_S : \mathcal{S} \to \mathbb{R}$. We solve for Bayes Nash equilibria of the game.

2.1 Discussion of modeling choices

The comparison of communication and full delegation is common in the literature. Note that optimal constraint delegation can obviously do better, in particular it can replicate any communication outcome.

Quadratic losses are standard in the communication literature as they simplify the analysis. Intuitively, assuming such losses seems reasonable in setups in which players have strong incentives to take the 'right' action.

Linear conflicts capture the idea that players have an agreement point and interests that diverge with the distance to that point, which we normalize to zero. For example, if a baseline procedure has been agreed upon in advance, players can disagree on how to adapt to changed circumstances. Motivated by potential adaptation costs defrayed by the receiver, the receiver responds more conservatively to these changes than the sender.

The combined modeling of a linear bias together with a flexible shape of the distribution allows us to capture the idea of large disagreement in extreme situations. For our generalization in Section 9, linearity of the bias is not crucial.

Logconcavity ensures that optimal choices and expected utility are well defined and that the tail of the distribution is relatively thin. In particular, we rule out distributions with tails that are heavier than exponential (Laplace), thus the likelihood of extreme states is not very high. This seems reasonable in the context of organizational decision making, as we expect regulators to ban operations in environments in which exceptional events occur with some likelihood.

Symmetry of the density implies that we can write $f(\theta) = c\frac{1}{\sigma}\psi\left(\frac{\theta^2}{\sigma^2}\right)$, where c is a normalizing constant and ψ is a (density generator) function that captures the shape of the distribution.[6] Importantly, the density depends only on the standardized variable $\frac{\theta}{\sigma}$. This representation allows us to vary the shape of the distribution and the variance independently, to study different measures of risk.

Given our focus on deviations from some agreement, symmetry implies that we treat deviations of the state in both directions equally. For example, changes in the pressure conditions in the well away from the expected value require adequate actions both for higher as well as for lower pressure values.

3 Equilibria and Payoffs

3.1 Communication equilibria

As is standard in cheap talk, communication equilibria are partitional. A partitional equilibrium is characterized by a sequence of *critical types*, $\boldsymbol{t}^n = (t_i^n)$, with $t_{i-1}^n < t_i^n$ and n relating to the number of induced actions. Sender types strictly within an interval, $\left(t_{i-1}^n, t_i^n\right)$, induce the same action; critical types, t_i^n, are indifferent between inducing the action in the interval below or the action in the interval above. As we show in Proposition 1 below, for any finite number of induced actions equilibria are symmetric in our model. For notational simplicity we, therefore, take $t_i^n \geq 0$ and denote the critical types below zero by $-t_i^n$ for all i and n. Receiving a message that indicates $\theta \in [\underline{t}, \bar{t})$, the receiver updates her belief by taking the conditional expectation $\mu(\underline{t}, \bar{t}) = \mathbb{E}\left[\Theta | \, \Theta \in [\underline{t}, \bar{t})\right]$. For equilibrium critical types $\boldsymbol{t}^n$, we define

$$\mu_i^n := \mathbb{E}\left[\Theta | \, \Theta \in \left[t_{i-1}^n, t_i^n\right)\right] \text{ for } i = 1, \ldots, n \text{ and } \mu_{n+1}^n := \mathbb{E}\left[\Theta | \, \Theta \geq t_n^n\right]. \quad (1)$$

[6]Note that any symmetric one-dimensional density is elliptical, Cambanis et al. (1981). The particular representation of elliptical densities can be found, e.g., in Gómez et al. (2003). Many distributions that are used in economics are elliptical with logconcave densities. Examples include the uniform, the Gaussian, the Laplace distribution, and many more.

Thus, the receiver's equilibrium action given a message indicating $\theta \in \left[t_{i-1}^n, t_i^n\right)$ is $\beta \cdot \mu_i^n = \arg\max_y \mathbb{E}[u_R(y, \theta, \beta) | \theta \in \left[t_{i-1}^n, t_i^n\right)]$. The indifference conditions of critical types that determine partitional equilibria are given by

$$t_i^n - \beta \cdot \mu_i^n = \beta \cdot \mu_{i+1}^n - t_i^n, \quad \text{for } i = 1, \ldots, n. \tag{2}$$

Symmetric equilibria come in two classes, depending on whether the total number of induced actions is even or odd. In an equilibrium with an even number of actions, type $\theta = 0$ must be a critical type. We call this type of equilibrium a *Class I* equilibrium, and the characterization uses $t_0^n = 0$. If the total number of induced actions is odd, then a symmetric interval around zero is part of the equilibrium. We call this a *Class II* equilibrium. In this case, we omit t_0^n from the construction. For an illustration with $n = 2$, see Figure 1. The step function depicts the receiver's actions.

Proposition 1 *Assume a symmetric distribution with a logconcave density.*
i) For all n*, there exist an essentially unique Class I equilibrium, which is symmetric and induces* $2(n+1)$ *actions, and an essentially unique Class II equilibrium, which is symmetric and induces* $2n+1$ *actions.*
ii) For $n \to \infty$*, the limits of the finite Class I and Class II equilibria exist, which induce infinitely many actions. We call any of these a limit equilibrium.*
iii) In a limit equilibrium, we have $\lim_{n\to\infty} t_1^n = 0$*.*

Proposition 1 proves the existence and uniqueness of partitional equilibria for arbitrary finite n. An analogous characterization of partitional equilibria is given in Deimen and Szalay (2019) for the special case of the Laplace distribution. Proposition 1 generalizes the result to all symmetric distributions with a logconcave density, a large and important class. Note that the support can be bounded or unbounded. Logconcavity of the distribution and the linear bias with $\beta \in (0,1)$ together imply that the solution of a certain forward difference equation is monotonic in the initial value, which we use to prove uniqueness.[7]

[7] For the proof, we take equilibria as a combination of a "forward solution" and a "closure condition." A forward solution that starts at t_0, takes the length of the first interval, say τ, as given,

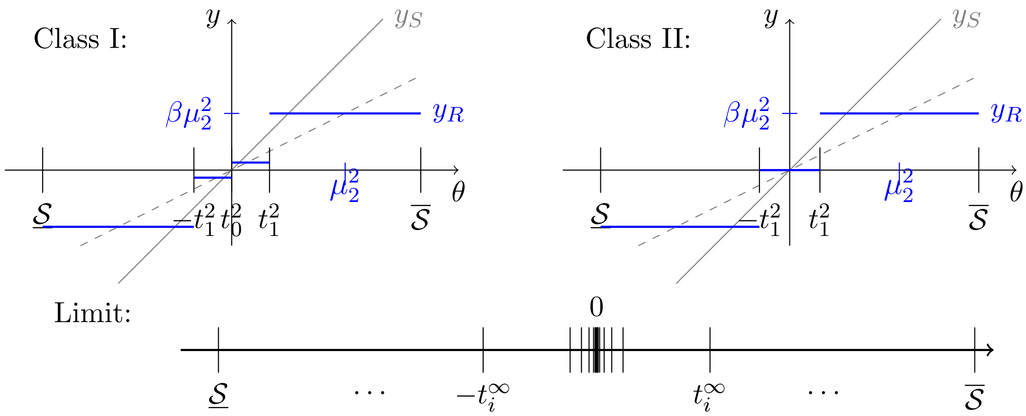

Figure 1: Partitional equilibria. Class I and Class II for $n = 2$. In a limit equilibrium, intervals around the prior mean $\mathbb{E}[\Theta] = 0$ get arbitrarily small as $n \to \infty$.

Moreover, the proposition proves that the limit as $n \to \infty$ also is an equilibrium. The limit equilibrium features an accumulation point at zero and a finite highest critical type, $\lim_{n\to\infty} t_n^n < \infty$. The boundedness of the highest critical type follows from the fact that a distribution with a logconcave density must have a decreasing mean residual life function. The highest critical type must remain finite to ensure that the distance from the highest receiver action to the critical type below it remains positive. This insight is new to the literature, which typically assumes a compact state space.

The partition of a limit equilibrium is illustrated in Figure 1, bottom panel. While the partitional form of equilibria is known from the seminal work of Crawford and Sobel (1982), the structure of the limit equilibrium is closest in spirit to Alonso et al. (2008) and Rantakari (2008). Gordon (2010) offers the first systematic account of the existence of infinite equilibria. In Gordon's taxonomy the sender is outward biased towards more extreme actions. We add to this literature by highlighting the

and computes the "next" threshold, $t_2(\tau)$, as a function of the preceding two, τ and t_0. Likewise, all following thresholds are constructed using their two predecessors. The closure condition for an equilibrium with n positive thresholds requires that τ is such that type $t_n^n(\tau)$ satisfies the indifference condition.

role of distributions and, in particular, the role of logconcavity for existence and uniqueness. Logconcavity provides a microfoundation for regularity properties that are often imposed in the literature.[8]

3.2 Communication payoff

We define the random variable μ^n of truncated expectations on the discrete support $(\pm\mu_i^n)_i$ given in equation (1). The discrete random variable μ^n is important for the calculation of the value of communication: as the next lemma illustrates, to compute expected payoffs, we need to determine the moment $var(\mu^n)$ from the marginal distribution of Θ and the equilibrium characterization. We denote the truncated expectation on the half of the support by $\mu_+ := \mathbb{E}[\Theta|\Theta \geq 0]$.

Under communication, the receiver bases her decision on the sender's message and takes the action $y_R(\mu_i^n) = \beta\mu_i^n$.

Lemma 1 *For any symmetric distribution with a given density generator $\psi(\cdot)$, the receiver's expected utility in any communication equilibrium is a linear function of the variance σ^2,*

$$\mathbb{E}u_R^{com}(y_R, \Theta, \beta) = -\beta^2\left(\sigma^2 - var(\mu^n)\right) = -\beta^2(1 - \ell(\beta, n))\sigma^2,$$

for some function $\ell(\beta, n)$ that is independent of σ^2.

The receiver's expected utility is proportional to the expected residual variance after communication. By a variance decomposition, this can be split into the difference of the prior variance (of the continuous state θ) to the expected variation in the (discrete) receiver actions. Moreover, the endogenous variance of receiver actions $var(\mu^n)$, which captures how much the receiver learns from communication, turns

[8] Following Crawford and Sobel (1982), the literature invokes condition M to ensure uniqueness. Logconcavity of the density and a receiver response with a slope less than one – not necessarily constant – is a condition on the primitives of the model that ensures that condition M is satisfied. See Lemma A.4 in the Appendix and Szalay (2012) for details. Gordon (2010) assumes a *regular* receiver response. Our results are in line with his nice characterization. We provide conditions that make a receiver response *regular*.

out to be a linear function of the exogenous state variance σ^2. This follows from the symmetry of the distribution, which allows us to write the sender's indifference conditions as functions of the standardized critical types. The receiver's payoff is thus linearly decreasing in the state variance σ^2.

3.3 Delegation payoff

As a simple alternative to directive decision-making where the receiver communicates with the sender, we consider simple, unconstrained delegation to the sender.[9] Under delegation, the informed sender takes the action $y_S = \theta$. Thus the information directly enters the decision; from the receiver's perspective, however, the decision is biased.

Lemma 2 *The receiver's expected utility under delegation is*

$$\mathbb{E}u_R^{del}\left(y_S, \Theta, \beta\right) = -\left(1-\beta\right)^2 \sigma^2.$$

Delegation is preferred by the receiver over choosing the action without communication for $\beta \geq \frac{1}{2}$*.*

4 Variance and the optimal mode of decision-making

Does an increase in risk *necessarily* require a change of the optimal mode of decision making? The answer is no. From Lemma 1 and Lemma 2, we know that expected utilities under communication and delegation are both linear in the variance. As a consequence, the difference in expected payoffs under communication and delegation is also linear in the variance. This implies:

Theorem 1 *For any symmetric distribution with a given density generator* $\psi(\cdot)$*, the choice between delegation and communication – in any equilibrium of the communication game – is independent of the variance* σ^2*.*

[9] In the literature on optimal delegation, the receiver can constrain the choice set of the sender. See, for example, Alonso and Matouschek (2008). Optimal delegation can replicate communication outcomes and is therefore always weakly better. We show that even simple, unconstrained delegation can strictly improve upon communication.

If a mode of decision-making is optimal for some variance, then the same mode of decision-making must be optimal for any level of the variance, all else equal. All else equal requires in particular, that the stochastic environment remains governed by a distribution with the same density generator, i.e., with the same shape. For example, think of two normal distributions with different variances.

If we keep the shape of the distribution fixed, then an increase in risk corresponds to a linear rescaling of the state space. Intuitively, this is a very regular increase in risk. In terms of our example one could think of oil drilling under similar circumstances in other regions, or of replicating the same activity. Our model shows that scaling risk up this way does not require a change in the organizational structure. The same mode of decision-making remains optimal.

5 Tail risk

We want to capture the idea of increased risk in terms an increased likelihood of extreme events. At the same time we want to keep the variance fixed to keep the delegation payoff constant (Theorem 1). We therefore introduce a fully parametrized model that allows to vary the shape of the distribution while keeping the variance fixed. The shape parameter of the generalized Pareto distribution measures the likelihood of extreme outcomes; we take this as a measure for the *tail risk* of the distribution. We impose the following assumption.

Assumption 1 *The state is distributed according to a two-sided generalized Pareto distribution with density*

$$f(\theta;\delta,s)=\frac{1}{2s}\left(1+\delta\frac{|\theta|}{s}\right)^{-\frac{1}{\delta}-1}\ \textit{for}\ \theta\in\left[\frac{s}{\delta},-\frac{s}{\delta}\right],$$

where $s\in(0,\infty)$ is a scale parameter and $\delta\in[-1,0]$ is a shape parameter.[10]

[10]The distribution is constructed from the well-known one-sided generalized Pareto by reflecting at zero. The location parameter is set at zero, to ensure that the mean is zero. The distribution is defined more generally for shape parameters $\delta\in(-\infty,\infty)$, but we restrict attention to the subset that features logconcave tails. We treat the case $\delta\geq 0$ in Deimen and Szalay (2019); these distributions have logconvex tails and an infinite support.

The framework naturally embodies tail risk in the shape parameter δ that determines – among other things – the kurtosis. It nests many well known distributions. In particular, the case $\delta = -1$ is the uniform distribution, $\delta = -\frac{1}{2}$ is the triangular distribution, and the limit case $\delta = 0$ is the Laplace distribution. For an illustration of these distributions, see Figure 2. The variance of the distribution is a function of scale and shape: $\sigma^2(s, \delta) = \frac{2s^2}{(1-\delta)(1-2\delta)}$. The support of the distribution is $[\frac{s}{\delta}, -\frac{s}{\delta}]$.

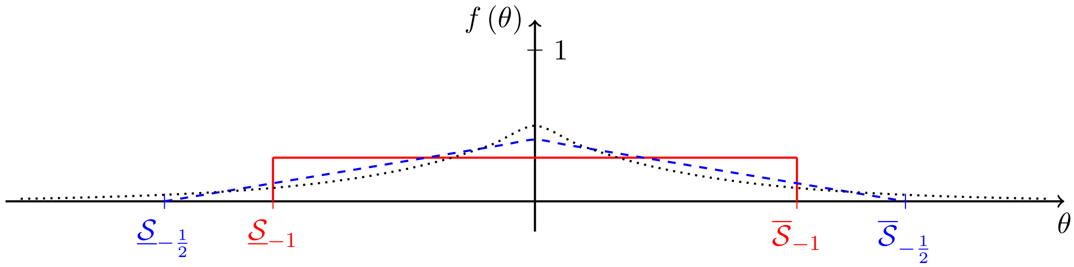

Figure 2: The uniform distribution (solid red, $\delta = -1$) and the triangular distributions (dashed blue, $\delta = -\frac{1}{2}$) and the Laplace distribution (dotted black, $\delta = 0$) all with variance $\sigma^2 = 1$.

The distributions vary in their supports and in their shape. Distributions that have larger supports also have more mass around zero – the densities on each half of the support cross twice. This is necessary to keep the variance constant.[11]

5.1 Tail risk and the performance of decision making

The generalized Pareto environment allows us to solve for the expected utilities arising from communication in closed form.

Proposition 2 (Deimen and Szalay (2019)) *For the two-sided generalized Pareto distribution with shape* $\delta \in [-1, 0]$ *and scale* $s^2 = \sigma^2 \frac{(1-\delta)(1-2\delta)}{2}$*, the variance of* μ^n *in*

[11]If the densities cross only once over the halves of the support, the distributions overall are ordered in the convex order, which implies different variances.

a Class I equilibrium[12] *is given by*

$$var(\mu^n) = \frac{2}{2-\frac{\beta}{1-\delta}}\mu_+^2 - \frac{\frac{\beta}{1-\delta}}{2-\frac{\beta}{1-\delta}}\left(\mu_1^n\right)^2. \tag{3}$$

In a limit equilibrium, we have

$$var(\mu^\infty) = \frac{2}{2-\frac{\beta}{1-\delta}}\mu_+^2 = \frac{2-\frac{1}{1-\delta}}{2-\frac{\beta}{1-\delta}}\sigma^2. \tag{4}$$

Since the generalized Pareto distribution features a linear tail conditional expectation function, the value of communication can be computed in closed form via dynamic programming (Deimen and Szalay (2019)).[13] Naturally, $var(\mu^n) \leq var(\mu^\infty) \leq var(\Theta)$; the value of partitional communication reaches the upper bound of fully revealing communication exactly if $\beta = 1$, that is, if interests are perfectly aligned. For given $\beta \in (0,1)$, the value is decreasing in δ. Higher risk in terms of δ reduces the value of communication, less information is transmitted in equilibrium.[14]

The dynamic programming approach delivers a sharp result, but offers little intuition for why tail risk is detrimental to information transmission. Based on the comparative statics of the model that we discuss below, a heuristic explanation is as follows. A higher tail risk parameter δ corresponds to an *exogenously* higher likelihood of extreme realizations of the state and around the mean – the tails and the peak of the distribution get more mass. The *endogenous* effect on communication equilibria is that relative to the support the actions move closer to the mean and the partition intervals in the tails get longer; a larger subset of the state space at each extreme is pooled. Quadratic loss functions disproportionately penalize large

[12]The variance of μ^n in a Class II equilibrium is given by $var(\mu^n) = \left(1 - \Pr\left[\Theta \in \left[-\frac{\beta\mu_2^n}{2}, \frac{\beta\mu_2^n}{2}\right)\right]\right) \cdot \left(\frac{2}{2-\frac{\beta}{1-\delta}}\mu_+^2 + \frac{\frac{\beta}{1-\delta}}{2-\frac{\beta}{1-\delta}}\mu_2^n\mu_+\right)$.

[13]In Deimen and Szalay (2019), distributions with a linear tail conditional expectation are derived from first principles as the solution to a differential equation. In that formulation, we obtain a solution that involves variance and the slope of the tail conditional expectation. Here, we observe that the generalized Pareto class can be obtain as a reparametrization – in terms of shape and scale – of the distributions with linear tail conditional expectations.

[14]For $\beta = 0$, the receiver's action equals zero for any sender strategy.

distances, so that gains from increased communication around the mean are dominated by extreme losses for extreme states. The value of communication is decreased in environments with higher tail risk.

6 Tail risk and optimal decision making – ex ante

We now readdress our question, whether a change of risk – now in the sense of tail risk – necessitates a change in the mode of decision-making.

Theorem 2 *Suppose the receiver can choose between communication and delegation. Then, delegation is better than communication – in any equilibrium of the communication game – if* $\delta \geq \frac{2-3\beta}{2-2\beta}$. *Communication in a limit equilibrium is better than delegation if* $\delta \leq \frac{2-3\beta}{2-2\beta}$.

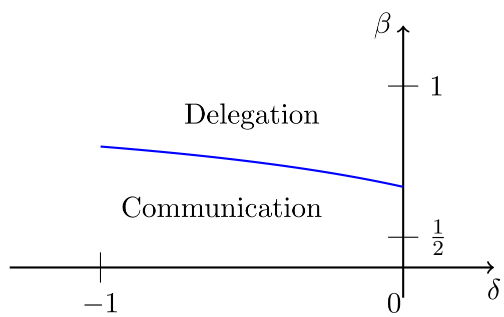

Figure 3: Delegation versus communication. On the horizontal axis, the tail risk parameter increases from −1 (uniform distribution) to 0 (Laplace distribution); on the vertical axis, the level of agreement increases from $\frac{1}{2}$ to 1.

The intuition for the result is straightforward. While the performance of delegation depends only on the variance of the environment, the performance of communication depends in addition on the shape of the distribution. The fraction of information that is transmitted in a limit equilibrium, $\frac{2-\frac{1}{1-\delta}}{2-\frac{\beta}{1-\delta}}$, is smaller in environments that feature greater tail risk as captured by the shape parameter δ. We depict the comparison in Figure 3.

Consistent with the literature, delegation dominates communication for low levels of conflict, that is, if the receiver is able to listen relatively well, $\beta \geq \frac{2-2\delta}{3-2\delta}$.[15] The comparison in terms of tail risk adds a new dimension to the literature. For $\beta \in \left(\frac{2}{3}, \frac{4}{5}\right)$, for a distribution with low tail risk communication is optimal but for a distribution with higher tail risk delegation is optimal. In other words, an increase in tail risk – i.e., in the likelihood of extreme situations – may indeed necessitate a change in the mode of decision-making from communication to delegation.

7 Tail risk and communication equilibria

Communication suffers when extreme situations become more likely. To understand the impact of tail risk on communication, we study how communication equilibria change in the shape parameter δ.

Recall that two distributions f, g with different tail risks $\delta_f < \delta_g$ and the same variance have different supports $\mathcal{S}_f \subset \mathcal{S}_g$, since $-\frac{s_f}{\delta_f} < -\frac{s_g}{\delta_g}$. Thus a direct comparison of equilibra is akin to comparing apples and oranges. To obtain a meaningful comparison, we must consider the scaled distribution $\hat{f}$ on the support $\mathcal{S}_g$. This can be done for distributions with finite support. Normalized to the same support, equilibria can be ordered by the level of tail risk.

Proposition 3 *Suppose $\delta < 0$. For any n, the equilibrium critical types and the induced actions satisfy $\frac{t^n_{i,f}}{\bar{\mathcal{S}}_f} > \frac{t^n_{i,g}}{\bar{\mathcal{S}}_g}$ and $\frac{\mu^n_{i,f}}{\bar{\mathcal{S}}_f} > \frac{\mu^n_{i,g}}{\bar{\mathcal{S}}_g}$ for all i.*

Equilibrium thresholds and actions are relatively more spread out under the less risky distribution f. Technically, the result relies on a nice property of the generalized Pareto environment: on the same support, distributions with different values of δ satisfy a monotone likelihood ratio property on each half of the support.[16] This implies on the positive half that higher realizations of Θ are more likely under distribution $\hat{f}_+$ than under distribution g_+ and thus that the conditional expectation

[15] See, for example, Alonso et al. (2008) and Rantakari (2008) who study a uniform distribution, i.e., $\delta = -1$. See also Dessein (2002).

[16] Over the entire support, the likelihood ratio is unimodal. See Section 9 for a definition. This complements Chen and Gordon (2015) which assume mlrp on the entire support.

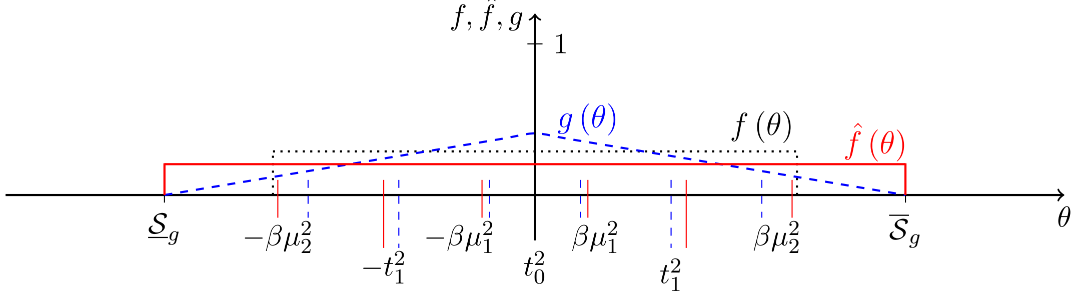

Figure 4: Scaled low-risk distribution $\hat{f}$ and corresponding equilibrium (solid red), in comparison to an equilibrium under a more risky distribution g (dashed blue).

for any arbitrary truncation is higher under $\hat{f}_+$ than under g_+.[17] See Figure 4. As a consequence, all equilibrium sender marginal types t_i^n and receiver responses μ_i^n on the positive half are – relative to the length of the support – higher under f than under g.

The exogenous variation of the tails of the distribution impacts the endogenous tails of the distribution that arise from equilibrium truncations under communication. If we make the tails heavier, the intervals at the extremes get longer relative to the support of the distribution. As a result, equilibrium partitions on the entire support are relatively more evenly spread out under distribution f than under g. Hence, in a low risk environment the receiver manages to tailor the actions better to the extreme states.

8 Tail risk and decision making in a worst case

We now switch to a worst-case perspective. We naturally define the *worst case* as the state in which the highest loss for the decision-maker arises, conditional on the

[17]Too see this, note that multiplying the likelihood ratio by an arbitrary constant does not change its monotonicity properties. If we take this constant to be the ratio of the probability masses over an interval, then it is straightforward to see that the distributions conditional on the truncation to this interval are ordered in the *likelihood ratio order* (i.e., they satisfy mlrp).

chosen mode of decision-making. The following lemma confirms the intuition that the worst cases arises in the most extreme realizations of the state.

Lemma 3 *The worst cases arise in states* $\theta \in \left\{\underline{\mathcal{S}}, \overline{\mathcal{S}}\right\}$, *giving rise to a worst case delegation loss of* $(1-\beta)^2 \overline{\mathcal{S}}^2$ *and a worst case communication loss of* $\beta^2 \left(\mu_{n+1}^n - \overline{\mathcal{S}}\right)^2$.

Next, consider which mode of decision making results in a larger worst case loss. Intuitively, for a linear bias the disagreement between sender and receiver is most extreme in the worst case, making delegation appear particularly unattractive. However, communication suffers from bad decisions that arise due to the last interval being large. It is not clear which effect dominates. To resolve the comparison, we need to better understand the choices under communication conditional on the worst case. An obstacle is that the highest critical type, t_n^n, cannot be computed in closed form – except for the special case of the uniform distribution. It turns out that we can derive an upper bound on the highest critical type. This will enable us to derive a lower bound on the loss that is made under communication in extreme situations and to identify environments in which communication performs strictly worse than delegation.

Lemma 4 *For any equilibrium, the highest equilibrium critical type is bounded from above,* $t_n^n < T := \frac{\beta s}{2-2\beta-2\delta+\beta\delta}$.

To understand the bound, note that equilibrium requires that for the highest critical type t_n^n the distance to the induced action below is equal to the distance to the induced action above. The distance to the action below is at least $(1-\beta)\, t_n^n$. Due to logconcavity and $\beta < 1$, the distance to the action above is decreasing in the value of t_n^n and gets shorter than $(1-\beta)\, t_n^n$ for t_n^n above T. The upper bound T is illustrated in Figure 5 for different values of δ. Notably, for a variation from the uniform to the Laplace distribution, for which the support increases from $\overline{\mathcal{S}}_{-1}$ to ∞, the upper bound increases only slightly from T_{-1} to T_0.

As a consequence of a large last interval, the communication loss conditional on the worst case is high. Indeed, for very risky environments, it even exceeds the delegation loss.

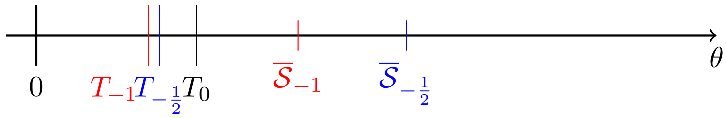

Figure 5: Upper bound of the last interval for tail risk $\delta = -1, -0.5, 0$, and $\beta = 0.75$.

Proposition 4 *The ratio of the communication relative to the delegation worst case losses is at least* $\frac{\beta^2}{(1-\beta)^2}\frac{1}{(1-\delta)^2}\left(\frac{\overline{\mathcal{S}}-T}{\overline{\mathcal{S}}}\right)^2$*, implying that for* $\delta \geq \frac{2-4\beta}{2-\beta}$*, delegation is better than communication from a worst-case perspective.*

The ratio of the losses depends ultimately on the relative length of the last communication interval proportional to the length of the support $\overline{\mathcal{S}}$. The proportion of the interval above the upper bound relative to the support $\frac{\overline{\mathcal{S}}-T}{\overline{\mathcal{S}}} = 1 + \frac{\delta\beta}{2-2\beta-2\delta+\beta\delta}$ is increasing in δ, which implies the statement. Higher risk makes it more difficult to communicate extreme states. Therefore, the receiver undershoots by a large extent. In very risky environments, the communication induced undershooting is more pronounced than the delegation induced overshooting.

8.1 A discrepancy: expected versus worst-case losses

We can now state our main insight from the generalized Pareto model.

Theorem 3 *For* $\beta \in \left[\frac{2-2\delta}{4-\delta}, \frac{2-2\delta}{3-2\delta}\right]$*, the ex ante optimal mode of decision-making is to communicate and the worst-case loss under communication exceeds the worst case loss under delegation.*

The proof is a straightforward combination of Propositions 2 and 4. For every distribution that is strictly more risky than the uniform ($\delta = -1$), there is a non-empty set of conflict parameters β such that the ex ante optimal mode of decision-making results in higher losses in the worst case. For an illustration, see Figure 6. This means that in our model, the decision maker willingly accepts higher losses than necessary in the worst case to maximize expected profits.

The consequences can be substantial. To illustrate, suppose the receiver finds it optimal to communicate despite the presence of large risks. How much worse can communication be relative to delegation under this hypothesis? Since relative

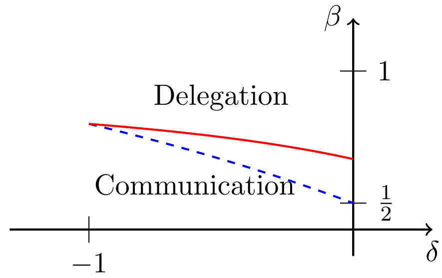

Figure 6: Worst-case (dashed blue) versus ex ante (solid red) bounds with delegation optimal above.

losses are increasing in β, we obtain the highest ratio, conditional on communication being optimal ex ante, for the borderline case where the receiver is just indifferent between communication and delegation, that is for $\beta = \frac{2-2\delta}{3-2\delta}$. In this case, the ratio of losses reduces to $\frac{4}{(\delta-1)^2}$, which ranges from 1 for the uniform to 4 for the Laplace distribution. This means that, conditional on the worst case occurring under communication under extreme risk, delegation would have performed 4 times better.[18] If the stakes are large – for example, when lives can be saved – reducing errors by 75% is a tremendous achievement.

9 Robustness: beyond the generalized Pareto

Simple delegation becomes relatively more attractive compared to communication when extreme events become more likely, both from an ex ante as well as from a worst case perspective. So far, we have used the specific functional form of the generalized Pareto framework to demonstrate these results. We now argue that the insights are robust beyond the parametrized approach. While the generalization of the worst case analysis requires no additional effort, the ex ante analysis requires some work. We address the worst case in the following paragraph and the ex ante

[18]Note that in this comparison, we constrain ourselves to keep communication ex ante optimal. The relative losses can get arbitrarily large if the ex ante optimal choice would have been to delegate. For example, for the Laplace ($\delta = 0$), the lower bound on the ratio of losses takes value $\frac{\beta^2}{(1-\beta)^2}$, which tends to infinity as $\beta \to 1$.

analysis in the following subsections.

Consider any two symmetric, logconcave densities f and g with finite supports $\mathcal{S}_f \subset \mathcal{S}_g$. To unify the supports, we can stretch the distribution f such that $\hat{f}$ is the rescaled version of f with support $\mathcal{S}_g$. This can be done by a linear transformation of the state space without altering the shape of the distribution. On the same support, conditional on the positive half, suppose that the densities satisfy the monotone likelihood ratio order, $\frac{\hat{f}_+(\theta)}{g_+(\theta)}$ increasing in θ. Then Proposition 3 states that the most extreme actions satisfy $\frac{\mu^n_{n+1,f}}{\bar{\mathcal{S}}_f} > \frac{\mu^n_{n+1,g}}{\bar{\mathcal{S}}_g}$. This implies a relatively higher worst case communication loss under the more risky distribution g. We can conclude that in the worst case, delegation becomes *relatively* more attractive in more risky environments.

From now on, we focus on the ex ante perspective. Since expected utilities depend on scale, we need to compare the distributions on their original supports (which can be infinite). We aim at comparing the value of communication for distributions with different tail risks. To this end, we introduce a stochastic order in which distributions with different likelihoods of extreme situations are ordered while keeping the variance fixed. In the following subsections, we show that communication works worse in more risky environments.

9.1 Uniform conditional variability

The following partial order helps us comparing distributions in terms of tail risk. We say that the half-distributions $f_+(\theta)$ and $g_+(\theta)$ are ordered in the *uniform conditional variability order* (see Whitt (1985), Shaked and Shanthikumar (2007)), if the following applies.

Definition 1 *Let Θ and $\tilde{\Theta}$ be two random variables with densities f_+ and g_+, respectively. The random variable Θ is* uniformly less variable *than $\tilde{\Theta}$ if the supports satisfy supp$(\Theta) \subseteq$ supp$(\tilde{\Theta})$ and the ratio $\frac{f_+(\theta)}{g_+(\theta)}$ is unimodal over the supp$(\tilde{\Theta})$, where the mode is a supremum, but Θ and $\tilde{\Theta}$ are not ordered by the usual stochastic order.*

This variability order entails differences in means and in variances over the halves; that is, $f_+(\theta)$ is *higher on average and less variable* than $g_+(\theta)$.[19]

[19]The generalized Pareto class with a constant variance satisfies this order. For this example, the

For an illustration, consider the following example in Figure 7. The top panel depicts the densities f and g of two members of the generalized Pareto family, whereby g features a relatively higher likelihood of extreme outcomes. The bottom panel depicts the likelihood ratio, $\frac{f(\theta)}{g(\theta)}$. On the positive half of the support, $\frac{f_+(\theta)}{g_+(\theta)}$ is unimodal with interior mode m.

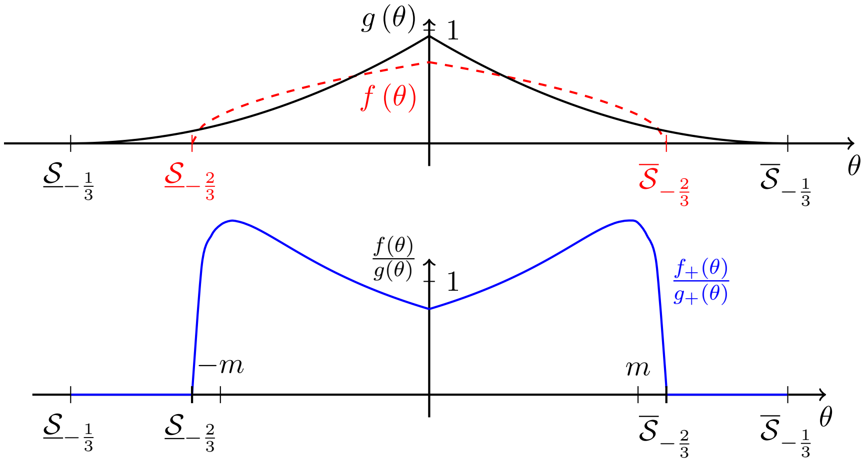

Figure 7: Top: distributions f with $\delta = -\frac{2}{3}$ (dashed red) and g with $\delta = -\frac{1}{3}$ (solid black). Bottom: the ratio $\frac{f}{g}$.

The next lemma links uniform conditional variability to another well-known concept, which ranks large classes of distributions with the same variance that differ with respect to the thickness of their tails:

Lemma 5 *Consider two symmetric distributions with the same variance and with densities f, g on $\mathbb{R}$ such that $\frac{f_+}{g_+}$ is logconcave. Then, $g_+(\theta)$ is uniformly more variable than $f_+(\theta)$.*[20]

moments conditional on the positive half satisfy $\mu_{f_+} > \mu_{g_+}$ and and $\sigma^2_{f_+} < \sigma^2_{g_+}$. Lemma A.7 in the appendix generalizes this property.

[20]From Shaked and Shanthikumar (2007) Theorem 3.A.54, it is known that relative logconcavity plus the densities crossing twice implies the uniform variability order. In contrast, we show that the uniform variability order arises from relative logconcavity plus the distributions having the same variance.

If $\frac{f_+}{g_+}$ is logconcave then f_+ is said to be *logconcave relative to* g_+ (Whitt (1985)). For example, note that $\frac{f_+}{g_+}$ is logconcave if f_+ is logconcave and g_+ is logconvex. Since logconcave (logconvex) densities on $\mathbb{R}_+$ feature a thin (thick) tail, the uniform conditional variability order applies if we compare distributions with thinner and thicker tails. The Laplace distribution features loglinear tails and divides the two classes – any logconcave density f_+ is logconcave relative to the Laplace; likewise, the Laplace is logconcave relative to any distribution with a logconvex density g_+. Another case of interest is the comparison relative to the Normal distribution.[21] Relative logconcavity is a transitive concept. Therefore, comparisons with these focal cases have implications for entire classes of distributions.

From now on we assume that the distributions have the same variances overall and that their halves satisfy the uniform conditional variability order.[22] We explore the implications of these assumptions with a primary focus on expected utilities.

9.2 Tail risk and the quality of communication

There is more information transmission and the receiver's expected utility is higher under distribution f than under g if and only if the receiver's choices are more variable under distribution f than under g. The following proposition states our generalization result.

Proposition 5 *Suppose that the densities f and g are logconcave and induce the same variance σ^2. Let g_+ be uniformly more variable than f_+. If β is sufficiently small, then there is more information transmission and the expected utilities are higher under f than under g. Formally*

$$var_f(\mu_f^n) > var_g(\mu_g^n). \tag{5}$$

[21] Excess kurtosis is typically measured relative to the Normal distribution, capturing distributions with heavier tails than the Normal. Distributions that are logconcave relative to the Normal are called *strongly logconcave*. See Wellner (2013) for a definition of strong logconcavity.

[22] We emphasize that the assumption of a constant variance – which is focal by Theorem 1 – rules out that the halves satisfy the standard stochastic order (and the distributions overall satisfy the convex order).

The left and the right side of (5) differ in three respects: (i) holding the partition of the state space induced by the sender's strategy fixed, the receiver's actions are different; (ii) still holding the partition induced by the sender's strategy fixed, the overall distributions over the receiver's actions differ; and (iii) the equilibria, i.e., the partition induced by the sender's strategy and the corresponding receiver responses, differ. We address points one to three below. We focus on finite equilibria. By taking limits to countably infinitely many actions all conclusions hold for a limit equilibrium.

An intuitive sketch of the proof is as follows: as Figure 7 reveals, close to the mean, distribution f_+ (f_-) is stochastically higher (lower) than g_+ (g_-) (in the likelihood ratio order). Imagine that all the critical types under distribution g are sufficiently close to the mean (this is the case for sufficiently small β). In this case, for the same sender partition, the distribution of receiver actions under f forms a mean preserving spread of the distribution of receiver actions under g.[23] This implies that the combination of effects (i) and (ii) increase the expected utility. Effect (iii) reinforces this: adjusting the critical types to the equilibrium partition under distribution f pushes the receiver's actions even farther away from the prior mean, leading to a further spread.

It should be emphasized that none of these arguments actually depend on quadratic losses. Neither does linearity of the ideal choice functions play a crucial role. We conclude that our insights are more general and robust beyond our leading case.

The remainder of this subsection discusses the arguments in detail. We prove the generalization result by establishing a sequence of Lemmas. The reader who is not interested in these details, can skip directly to Subsection 9.3.

As Figure 7 and Definition 1 of the uniform conditional variability order reveal, the local stochastic order depends on the location of the equilibrium thresholds considered. By symmetry, we focus on the positive half of the support only. For intervals below (above) the mode m, the truncated distributions under f_+ dominate (are dominated by) the truncated distributions under g_+ in the likelihood ratio or-

[23]This is remarkable, because the underlying distributions are not mean preserving spreads of each other.

der. To have some control over which order applies to which intervals – for example, to the first n intervals – it is helpful to establish monotonicity of equilibria in the conflict parameter β:

Lemma 6 *For any symmetric logconcave density and for any n, the equilibrium critical types $t_i^n(\beta)$ and induced means $\mu_i^n(\beta)$ are strictly increasing in β for all i.*

The result is formally a corollary to part ii) of the proof of Proposition 3. To get some intuition, consider two adjacent intervals and a critical type who is just indifferent between pooling downwards and upwards. Now increase the receiver's response parameter β. As a result, the action that the sender induces by pooling downwards has moved closer to the critical type and the action that the sender induces by pooling upwards has moved farther away from the critical type. Hence, the critical type must adjust and indeed increase. The proof is more involved than the simple intuition, because *all* critical types and *all* actions change. Logconcavity of the density implies that the receiver's responses move relatively slowly compared to the changes in the critical types, which gives stability to the system of equations that characterize the equilibrium.

To address the first difference of (5), consider a change of the distribution from g to f and (i) keep the partition of the state space induced by the critical types $\left(t_{i,g}^n(\beta)\right)_i$ fixed according to the equilibrium under g and (ii) allow the receiver to change her response from $\beta\cdot\mu_{i,g}^n=\beta\cdot\mu_g^n\left(t_{i-1,g}^n,t_{i,g}^n\right)$ to $\beta\cdot\mu_f\left(t_{i-1,g}^n,t_{i,g}^n\right)$. How do the receiver's responses change? Typically, the lowest actions increase and the highest actions decrease. As β is decreased, the former set expands while the latter set shrinks – to the point where it gets empty. In particular, we have:

Lemma 7 *For any two symmetric, logconcave densities f,g with the same variance and with truncated densities f_+,g_+ that satisfy Definition 1, there exists a unique $\hat{\beta}$ such that $\mathbb{E}_f\left[\Theta|\,\Theta\geq t_{n,g}^n\left(\hat{\beta}\right)\right]=\mathbb{E}_g\left[\tilde{\Theta}|\tilde{\Theta}\geq t_{n,g}^n\left(\hat{\beta}\right)\right]$. Moreover, for $\beta<\hat{\beta}$, all $n+1$ receiver responses under distribution f_+ are strictly higher than under g_+, $\beta\cdot\mu_f\left(t_{i-1,g}^n,t_{i,g}^n\right)>\beta\cdot\mu_g^n\left(t_{i-1,g}^n,t_{i,g}^n\right)$ for all i.*

If β is sufficiently low, then the receiver responds uniformly more conservatively – with actions that differ less from the prior mean – under distribution g compared to f.

For intervals below the mode m, the result is immediate by noting that the locally increasing likelihood ratio is preserved under truncations and that an increasing likelihood ratio implies higher truncated means under distribution f_+. For low enough β, this argument allows us to show that the first n receiver actions must be higher. However, the argument does not apply to the highest action. To compare the actions on the highest interval, we show that the tail conditional expectation function under f, $\mathbb{E}_f\left[\Theta|\,\Theta \geq x\right]$, crosses the tail conditional expectation function under g, $\mathbb{E}_g[\tilde{\Theta}|\tilde{\Theta} \geq x]$, exactly once and from above.[24] By symmetry, over the entire support all actions move farther away from zero, implying that the variance of the receiver actions increases. The first difference thus indicates an increase in the receiver's expected utility for a change from g to f if β is sufficiently low.

To address the second difference, we consider the variance of the receiver actions when the probability weighting over the partition elements is adjusted to reflect the new distribution f – still keeping the partition fixed at the equilibrium under g, $\left(t^n_{i,g}\left(\beta\right)\right)_i$. Similarly to the previous difference, we expect a positive effect on the variance of choices on intervals below the mode m and a negative effect on intervals above m. Again, the second region shrinks to the point where it becomes empty when β is low enough. As a result, the second adjustment reinforces the first one:

Let $\mu_{f,\mathbf{t}^n_g}$ denote the discrete random variable with realizations equal to the truncated means taken under distribution f and critical types taken under g.

Lemma 8 *Fix the partition of the state space at the equilibrium partition under distribution* g, $\left(t^n_{i,g}\left(\beta\right)\right)_i$. *For* $\beta \leq \hat{\beta}$ *(defined in Lemma 7), we have*

$$var_f(\mu_{f,\mathbf{t}^n_g}) > var_g(\mu^n_g).$$

By Lemma 7, all receiver actions are more spread out relative to the prior mean. To prove the current lemma, we show that the probability to take actions that are farther away from the prior mean is higher under f than under g. Taken together the results imply that the distribution of receiver actions under distribution f forms

[24] To the best of our knowledge this result is new to the literature. We note that the single crossing property implies that the moments of the distributions f_+ and g_+ satisfy $\mu_{f_+} > \mu_{g_+}$ and $\sigma^2_{f_+} < \sigma^2_{g_+}$, establishing the mean-variance trade-off as a general feature.

a mean preserving spread of the distribution of receiver actions under g. Clearly, this implies that the variance of receiver responses is increased.

Consider finally the third difference, the change in the equilibrium – i.e., critical types and induced receiver responses – when the distribution is changed from g to f. We find the following.

Lemma 9 *Suppose that $\beta \leq \hat{\beta}$. Then, all equilibrium critical types satisfy $t^n_{i,f} > t^n_{i,g}$ for all i. Moreover,*

$$var_f(\mu^n_f) > var_f(\mu_{f,\mathbf{t}^n_g}).$$

For β low enough, consider a change of the distribution from g to f. Focusing on one critical type who is just indifferent between pooling upwards and downwards on any two adjacent intervals and holding all other critical types constant, this type needs to increase. Moreover, this implies that, when we allow *all* critical types to adjust from g to f, that all of them increase. The increase in the critical types improves expected utility. For a quadratic loss function this is equivalent to increasing the variance of the actions. Intuitively, increasing actions is appreciated, because $\beta < 1$ implies that receiver's actions are too low relative to the first-best.

The Lemmas taken together prove Proposition 5. Some restriction to relatively pronounced conflicts is needed for the proof; however, our sufficient condition that $\beta < \hat{\beta}$ is far from necessary. Note that for generalized Pareto distributions expected utilities can be ordered for all values of β up to 1. Note also that the arguments are based on mean-preserving spreads and increasing utility. Hence, they are not confined to quadratic losses.

9.3 Gaussian versus Laplace

In the previous sections, we have shown that higher tail risk is detrimental to communication not only in the generalized Pareto but also in a more general framework with symmetric logconcave distributions. We now show by means of an example that our insights on the comparison of institutions also transfer to the general setup.

Suppose that the environment with relatively low tail risk is described by a Gaussian distribution, whereas the environment with thicker tails is described by a Laplace distribution. Let B_G denote the set of parameters β such that communication is preferred over delegation if the distribution is Gaussian, i.e., $B_G = \{\beta|\text{ communication} \succ \text{delegation}, f \text{ Gaussian}\}$. Thicker tails may necessitate a change in the mode of decision-making from communication to delegation:

Proposition 6 *Suppose the state follows a Gaussian distribution and $\beta \in B_G$. Then:*
i) the value of communication is higher than for a Laplace distribution with the same variance;
ii) there is a nonempty set $D \subset B_G$, such that for $\beta \in D$ delegation is preferred over communication for a Laplace distribution.

Proposition 6 generalizes the property of a decreasing slope in Figure 3: for intermediate levels of conflicts, a change from a Gaussian to a Laplace distribution makes it optimal to change the mode of decision-making from communication to delegation. The reason is that communication transmits so little information in the more risky environment.

The proposition is relevant beyond this point. It shows that the amount of conflict for which our generalization result Proposition 5 holds, remains in an interesting range. In particular, note that for large conflicts with $\beta \leq \frac{1}{2}$ delegation is suboptimal for any distribution. Proposition 6, however, shows that there exist a set of conflict parameters for which the value of communication is sufficiently low such that a switch to delegation becomes optimal.

10 Conclusions

In this paper, we study the impact of risk on the performance of communication. We find that higher likelihoods of extreme events – heavier tails – are detrimental to communication. We explore the consequences for the choice of an optimal mode of decision-making. Delegation becomes relatively more attractive in environments with higher risks. We expect that firms take these forces into account, because it helps them increase their expected profits.

We also compare losses arising from the different institutions in worst cases. Communication tends to produce higher losses in the worst case than delegation for plausible values of conflicts. Thus, there is a sense in which the decision-maker willingly accepts the possibility of relatively large losses. Judged from the worst-case perspective, this definitely amounts to *avoidable* losses.

Our analysis suggests that there is at least room for debate about a desirable objective in face of large risks, for example those that *BP* was facing when drilling in the Gulf of Mexico. A regulator that places a higher weight on avoiding catastrophic outcomes would mandate that experts take more responsibility for decision-making. We hope that our model and its analysis may contribute to a debate about reasoned regulation in face of catastrophic risks.

A Appendix

Definition A.1 *The forward equation is recursively defined as solutions* $t_{i+1}(t_{i-1}, t_i)$ *to the indifference conditions of types* t_i*. We denote an arbitrary initial value of* t_1 *by* τ*. In particular, for* $i = 1$ *we have* $t_2(0, \tau)$ *as solution to*

$$2\tau - \beta \mathbb{E}\left[\Theta | \, \Theta \in [0, \tau]\right] - \beta \mathbb{E}\left[\Theta | \, \Theta \in [\tau, t_2(0, \tau)]\right] = 0, \tag{6}$$

for $i > 1$ *we have* $t_{i+1}(t_{i-1}, t_i)$ *as solutions to*

$$2t_i - \beta \mathbb{E}\left[\Theta | \, \Theta \in [t_{i-1}, t_i]\right] - \beta \mathbb{E}\left[\Theta | \, \Theta \in [t_i, t_{i+1}(t_{i-1}, t_i)]\right] = 0. \tag{7}$$

Lemma A.1 (Szalay (2012)) *(Strict) Logconcavity of the distribution implies that*

$$\frac{\partial}{\partial t_{i-1}} \mathbb{E}\left[\Theta | \, \Theta \in [t_{i-1}, t_i]\right] + \frac{\partial}{\partial t_i} \mathbb{E}\left[\Theta | \, \Theta \in [t_{i-1}, t_i]\right] \leq (<) 1.$$

Lemma A.2 *Consider the forward equation. Logconcavity of the distribution and* $\beta < 1$ *implies that for all* $i = 1, \ldots, n-1$

$$\frac{dt_{i+1}}{dt_i} = \frac{\left(2 - \beta \frac{\partial}{\partial t_i} \mathbb{E}\left[\Theta | \, \Theta \in [t_{i-1}, t_i]\right] - \beta \frac{\partial}{\partial t_i} \mathbb{E}\left[\Theta | \, \Theta \in [t_i, t_{i+1}]\right]\right)}{\beta \frac{\partial}{\partial t_{i+1}} \mathbb{E}\left[\Theta | \, \Theta \in [t_i, t_{i+1}]\right]} > 1.$$

Proof of Lemma A.2. Consider the forward equation for t_2. The value $t_2(0, \tau)$ is the unique solution to (6). Totally differentiating (6) we find

$$\frac{dt_2}{d\tau} = \frac{\left(2 - \beta \frac{\partial}{\partial \tau} \mathbb{E}\left[\Theta | \, \Theta \in [0, \tau]\right] - \beta \frac{\partial}{\partial \tau} \mathbb{E}\left[\Theta | \, \Theta \in [\tau, t_2]\right]\right)}{\beta \frac{\partial}{\partial t_2} \mathbb{E}\left[\Theta | \, \Theta \in [\tau, t_2]\right]} > 1,$$

where the inequality follows from Lemma A.1:

$$2 - \beta \frac{\partial}{\partial \tau} \mathbb{E}\left[\Theta | \, \Theta \in [0, \tau]\right] > 1 > \beta \frac{\partial}{\partial \tau} \mathbb{E}\left[\Theta | \, \Theta \in [\tau, t_2]\right] + \beta \frac{\partial}{\partial t_2} \mathbb{E}\left[\Theta | \, \Theta \in [\tau, t_2]\right].$$

Next, consider arbitrary $i = 1, \ldots, n-1$. The sender's solution to the forward equation for t_i is given by (7). Totally differentiating (7) yields

$$\left(2 - \beta \frac{\partial}{\partial t_i} \mathbb{E}\left[\Theta | \, \Theta \in [t_{i-1}, t_i]\right] - \beta \frac{\partial}{\partial t_i} \mathbb{E}\left[\Theta | \, \Theta \in [t_i, t_{i+1}]\right] - \beta \frac{\partial}{\partial t_{i-1}} \mathbb{E}\left[\Theta | \, \Theta \in [t_{i-1}, t_i]\right] \frac{dt_{i-1}}{dt_i}\right) dt_i$$
$$= \beta \frac{\partial}{\partial t_{i+1}} \mathbb{E}\left[\Theta | \, \Theta \in [t_i, t_{i+1}]\right] dt_{i+1}.$$

Suppose as an inductive hypothesis that $\frac{dt_i}{dt_{i-1}} > 1$, so $\frac{dt_{i-1}}{dt_i} < 1$. Rearranging, we get

$$\frac{dt_{i+1}}{dt_i} = \frac{\left(2 - \beta\frac{\partial}{\partial t_i}\mathbb{E}\left[\Theta|\,\Theta \in [t_{i-1}, t_i]\right] - \beta\frac{\partial}{\partial t_i}\mathbb{E}\left[\Theta|\,\Theta \in [t_i, t_{i+1}]\right] - \beta\frac{\partial}{\partial t_{i-1}}\mathbb{E}\left[\Theta|\,\Theta \in [t_{i-1}, t_i]\right]\frac{dt_{i-1}}{dt_i}\right)}{\beta\frac{\partial}{\partial t_{i+1}}\mathbb{E}\left[\Theta|\,\Theta \in [t_i, t_{i+1}]\right]}$$
$$> 1,$$

which obtains by the inductive hypothesis and Lemma A.1:

$$\begin{aligned} & 2 - \beta\frac{\partial}{\partial t_i}\mathbb{E}\left[\Theta|\,\Theta \in [t_{i-1}, t_i]\right] - \beta\frac{\partial}{\partial t_{i-1}}\mathbb{E}\left[\Theta|\,\Theta \in [t_{i-1}, t_i]\right]\frac{dt_{i-1}}{dt_i} \\ > \;& 2 - \beta\frac{\partial}{\partial t_i}\mathbb{E}\left[\Theta|\,\Theta \in [t_{i-1}, t_i]\right] - \beta\frac{\partial}{\partial t_{i-1}}\mathbb{E}\left[\Theta|\,\Theta \in [t_{i-1}, t_i]\right] \\ > \;& 1 > \beta\frac{\partial}{\partial t_i}\mathbb{E}\left[\Theta|\,\Theta \in [t_i, t_{i+1}]\right] + \beta\frac{\partial}{\partial t_{i+1}}\mathbb{E}\left[\Theta|\,\Theta \in [t_i, t_{i+1}]\right]. \end{aligned}$$

□

Proof of Proposition 1. The following proof generalizes the proof of Proposition 1 in Deimen and Szalay (2019), which uses the functional form of the Laplace distribution. The steps of the proof are exactly the same, except for the fact that we do not use any functional form here, but rather assume the general class of logconcave densities.

The proof of the proposition consists of three lemmas. Lemma A.3 proves uniqueness of finite equilibria that do exist. Note that we do not assume symmetry here but take an arbitrary number of steps N. By symmetry of payoffs and the density, the model has symmetric equilibria. Together this implies that all finite equilibria must be symmetric around 0. Lemma A.4 then proves existence of symmetric equilibria for arbitrary N. Lemma A.5 proves existence of a limit equilibrium. □

Lemma A.3 *For any finite number N, if there exists an equilibrium with N distinct actions, then the equilibrium is unique.*

Proof of Lemma A.3. Fix N. By Lemma A.2, $\frac{dt_{i+1}}{dt_i} > 1$.

Fix an initial value $t_1 = \tau$ and take $t_2(\tau), \ldots, t_N(\tau)$ as determined by the forward equations up to and including $t_N(\tau)$. Consider the difference

$$2t_N(\tau) - \beta\mathbb{E}\left[\theta|\,\theta \in [t_{N-1}(\tau), t_N(\tau)]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq t_N(\tau)\right].$$

By Lemma A.2 this difference is a strictly monotonic function of τ as

$$\left(2-\beta\mathbb{E}\frac{\partial}{\partial t_N}\left[\theta|\,\theta\in[t_{N-1},t_N]\right]-\beta\frac{\partial}{\partial t_N}\mathbb{E}\left[\theta|\,\theta\geq t_N\right]\right)dt_N$$
$$-\beta\frac{\partial}{\partial t_{N-1}}\mathbb{E}\left[\theta|\,\theta\in[t_{N-1},t_N]\right]\frac{dt_{N-1}}{dt_N}dt_N>0.$$

Therefore, there is at most one value of τ, say τ_N^*, such that the vector $\left(t_1^N,\ldots,t_N^N\right)$ with $t_1^N:=\tau_N^*$ and $t_i^N:=t_i\left(\tau_N^*\right)$ solves the system of indifference conditions. Hence, the equilibrium is unique. □

Lemma A.4 *For any n, there exists an equilibrium inducing $N=2\,(n+1)$ actions and there exists an equilibrium inducing $N=(2n+1)$ actions. For N even (odd) the first equilibrium threshold t_1^n is decreasing in n.*

Proof of Lemma A.4. We, here, focus on the equilibria with an even number of induced actions. All the results extend to the equilibria with an odd number of induced actions.

Consider the truncated distribution, where the truncation is at zero and to the positive side. By symmetry, indifference for type zero is trivially satisfied. We construct an equilibrium as follows. We first consider the forward solution for arbitrary $t_1=\tau$ and show that for any n, the forward equation is guaranteed to have solutions up to t_n as long as $\tau\leq\tau^n$, for some well defined bound $\tau^n=\tau(n)$. Moreover, we show that $\tau^{n+1}<\tau^n$. We then consider an equilibrium of the communication game with n positive thresholds, which have to satisfy the forward equations and the closure condition:

$$2t_n(\tau)-\beta\mathbb{E}\left[\theta|\,\theta\in[t_{n-1}(\tau),t_n(\tau)]\right]-\beta\mathbb{E}\left[\theta|\,\theta\in\left[t_n(\tau),\overline{\mathcal{S}}\right]\right]=0 \tag{8}$$

for $\tau=t_1^n$. In particular, we show that there exists a unique initial value $\tau_n^*=\tau^{n+1}=t_1^n$ such that the forward solutions $t_2(t_1^n)<\ldots<t_n(t_1^n)<\overline{\mathcal{S}}$ exist and that $t_{n-1}^n=t_{n-1}(t_1^n),t_n^n=t_n(t_1^n)$ satisfy the closure condition, and hence we have an equilibrium.

If the forward equation for $t_2\,(\tau)$ exists, then it is the unique value of t_2 that satisfies equation (6). The limit as $t_2\to\tau$ of the left side of equation (6) is strictly positive as $2\tau - \beta\mathbb{E}[\theta|\, \theta \in [0,\tau]] - \beta\tau > 0$. Moreover, the left side is decreasing in t_2. In the limit as $t_2 \to \overline{\mathcal{S}}$, the left side is

$$2\tau - \beta\mathbb{E}[\theta|\, \theta \in [0,\tau]] - \beta\mathbb{E}[\theta|\, \theta \geq \tau].$$

It is well known that by logconcavity, $\mathbb{E}[\theta|\, \theta \in [0,\tau]]$ and $\mathbb{E}[\theta|\, \theta \geq \tau]$ increase with τ each at rate smaller than or equal to one. Hence, there exists a finite solution $t_2(\tau)$ if and only if $\tau < \tau^2$, where τ^2 is defined as the unique value of τ that solves

$$2\tau^2 - \beta\mathbb{E}\left[\theta|\, \theta \in \left[0,\tau^2\right]\right] - \beta\mathbb{E}\left[\theta|\, \theta \geq \tau^2\right] = 0. \tag{9}$$

Note that, for $\tau \to 0$ we have $t_2(\tau) \to 0$, and $t_2(\tau) - \tau$ is increasing in τ.

Consider next the forward solution for $t_3(\tau)$. If it exists, it is the value of t_3 that solves equation (7) for $i = 3$

$$2t_2(\tau) - \beta\mathbb{E}[\theta|\, \theta \in [\tau, t_2(\tau)]] - \beta\mathbb{E}[\theta|\, \theta \in [t_2(\tau), t_3]] = 0. \tag{10}$$

For $t_3 \to t_2(\tau)$, the left side of (10) takes value

$$2t_2(\tau) - \beta\mathbb{E}[\theta|\, \theta \in [\tau, t_2(\tau)]] - \beta t_2(\tau) > 0.$$

Moreover, the left side of (10) is decreasing in t_3. Hence, there exists a finite solution if and only if

$$2t_2(\tau) - \beta\mathbb{E}[\theta|\, \theta \in [\tau, t_2(\tau)]] - \beta\mathbb{E}[\theta|\, \theta \geq t_2(\tau)] < 0.$$

Differentiating the left side of (10) totally, we obtain

$$\left(\left(2 - \beta\frac{\partial}{\partial t_2}\mathbb{E}[\theta|\, \theta \in [\tau, t_2(\tau)]] - \beta\frac{\partial}{\partial t_2}\mathbb{E}[\theta|\, \theta \geq t_2(\tau)]\right)\frac{dt_2}{d\tau} - \beta\frac{\partial}{\partial \tau}\mathbb{E}[\theta|\, \theta \in [\tau, t_2(\tau)]]\right) d\tau.$$

As $\frac{d\tau}{dt_2} < 1$ by Lemma A.2, the expression is increasing in τ. Hence, there exists a unique value τ^3 such that a finite solution $t_3(\tau)$ exists for $\tau < \tau^3$. The value τ^3 satisfies

$$2t_2\left(\tau^3\right) - \beta\mathbb{E}\left[\theta|\, \theta \in \left[\tau^3, t_2\left(\tau^3\right)\right]\right] - \beta\mathbb{E}\left[\theta|\, \theta \geq t_2\left(\tau^3\right)\right] = 0. \tag{11}$$

At τ^3, the forward equation for $t_2(\tau^3)$, equation (6), implies that

$$2\tau^3 - \beta\mathbb{E}\left[\theta | \theta \in \left[0, \tau^3\right]\right] = \beta\mathbb{E}\left[\theta | \theta \in \left[\tau^3, t_2\left(\tau^3\right)\right]\right].$$

Substituting back into (11) gives

$$2t_2\left(\tau^3\right) - \beta\mathbb{E}\left[\theta | \theta \geq t_2\left(\tau^3\right)\right] = 2\tau^3 - \beta\mathbb{E}\left[\theta | \theta \in \left[0, \tau^3\right]\right].$$

Subtracting $\beta\mathbb{E}\left[\theta | \theta \geq \tau^3\right]$ from each side, we get

$$2\tau^3 - \beta\mathbb{E}\left[\theta | \theta \in \left[0, \tau^3\right]\right] - \beta\mathbb{E}\left[\theta | \theta \geq \tau^3\right] = 2t_2\left(\tau^3\right) - \beta\mathbb{E}\left[\theta | \theta \geq t_2\left(\tau^3\right)\right] - \beta\mathbb{E}\left[\theta | \theta \geq \tau^3\right].$$

Since

$$2t_2\left(\tau^3\right) - \beta\mathbb{E}\left[\theta | \theta \geq t_2\left(\tau^3\right)\right] = \beta\mathbb{E}\left[\theta | \theta \in \left[\tau^3, t_2\left(\tau^3\right)\right]\right],$$

by (11), the right side takes value

$$\beta\mathbb{E}\left[\theta | \theta \in \left[\tau^3, t_2\left(\tau^3\right)\right]\right] - \beta\mathbb{E}\left[\theta | \theta \geq \tau^3\right] < 0,$$

and hence

$$2\tau^3 - \beta\mathbb{E}\left[\theta | \theta \in \left[0, \tau^3\right]\right] - \beta\mathbb{E}\left[\theta | \theta \geq \tau^3\right] < 0.$$

Now recall equation (9): $2\tau^2 - \beta\mathbb{E}\left[\theta | \theta \in [0, \tau^2]\right] - \beta\mathbb{E}\left[\theta | \theta \geq \tau^2\right] = 0$. Since $2\tau - \beta\mathbb{E}\left[\theta | \theta \in [0, \tau]\right] - \beta\mathbb{E}\left[\theta | \theta \geq \tau\right]$ is increasing in τ by logconcavity, we have shown that $\tau^3 < \tau^2$.

Totally differentiating (10) gives

$$\frac{dt_3}{dt_2} = \frac{2 - \beta\frac{\partial}{\partial t_2}\mathbb{E}\left[\theta | \theta \in [\tau, t_2(\tau)]\right] - \beta\frac{\partial}{\partial t_2}\mathbb{E}\left[\theta | \theta \in [t_2(\tau), t_3]\right] - \beta\frac{\partial}{\partial \tau}\mathbb{E}\left[\theta | \theta \in [\tau, t_2(\tau)]\right]\frac{d\tau}{dt_2}}{\beta\frac{\partial}{\partial t_3}\mathbb{E}\left[\theta | \theta \in [t_2(\tau), t_3]\right]}.$$

Hence, $\frac{dt_3}{dt_2} > 1$ given that $\frac{dt_2}{d\tau} > 1$. It follows that $t_3(\tau) - t_2(\tau)$ is increasing in τ. Likewise, $t_3(\tau)$ goes to zero as $\tau \to 0$.

Suppose that the forward solutions exist up to $t_{n-1}(\tau)$ and all have the above properties. If the forward solution for $t_n(\tau)$ exists, it is defined as the value that satisfies

$$t_{n-1}(\tau) - \beta\mathbb{E}\left[\theta | \theta \in [t_{n-2}(\tau), t_{n-1}(\tau)]\right] = \beta\mathbb{E}\left[\theta | \theta \in [t_{n-1}(\tau), t_n]\right] - t_{n-1}(\tau). \quad (12)$$

At $t_n = t_{n-1}(\tau)$ the right side is negative, while the left side is positive. The right side is increasing in t_n, so there exists a unique finite solution if and only if

$$2t_{n-1}(\tau) - \beta\mathbb{E}[\theta|\,\theta \in [t_{n-2}(\tau), t_{n-1}(\tau)]] - \beta\mathbb{E}[\theta|\,\theta \geq t_{n-1}(\tau)] < 0.$$

Totally differentiating, we note that the difference is increasing in τ by the fact that $\frac{dt_{n-1}}{dt_{n-2}} > 1$. Hence, there is a unique value τ^n such that a forward solution $t_n(\tau)$ exists for any $\tau < \tau^n$, where τ^n is defined by the condition

$$2t_{n-1}(\tau^n) - \beta\mathbb{E}[\theta|\,\theta \in [t_{n-2}(\tau^n), t_{n-1}(\tau^n)]] - \beta\mathbb{E}[\theta|\,\theta \geq t_{n-1}(\tau^n)] = 0.$$

We now argue that $\tau^n < \tau^{n-1}$. Consider

$$2t_{n-2}\left(\tau^{n-1}\right) - \beta\mathbb{E}\left[\theta|\,\theta \in \left[t_{n-3}\left(\tau^{n-1}\right), t_{n-2}\left(\tau^{n-1}\right)\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq t_{n-2}\left(\tau^{n-1}\right)\right] = 0. \tag{13}$$

At τ^n, the forward equation for $t_{n-1}(\tau)$ implies

$$2t_{n-2}(\tau^n) - \beta\mathbb{E}[\theta|\,\theta \in [t_{n-3}(\tau^n), t_{n-2}(\tau^n)]] = \beta\mathbb{E}[\theta|\,\theta \in [t_{n-2}(\tau^n), t_{n-1}(\tau^n)]].$$

Hence, at τ^n,

$$\begin{aligned} & 2t_{n-2}(\tau^n) - \beta\mathbb{E}[\theta|\,\theta \in [t_{n-3}(\tau^n), t_{n-2}(\tau^n)]] - \beta\mathbb{E}[\theta|\,\theta \geq t_{n-2}(\tau^n)] \\ = \;& \beta\mathbb{E}[\theta|\,\theta \in [t_{n-2}(\tau^n), t_{n-1}(\tau^n)]] - \beta\mathbb{E}[\theta|\,\theta \geq t_{n-2}(\tau^n)] \\ < \;& 0. \end{aligned}$$

Since the left side of (13) is increasing in τ, it follows that $\tau^{n-1} > \tau^n$ is necessary to restore equality with zero.

Consider now the closure condition. Take $\tau \leq \tau^n$. A sequence of thresholds $\tau, t_2(\tau), \ldots, t_n(\tau)$ forms an equilibrium if and only if the thresholds $t_{n-1}(\tau)$ and $t_n(\tau)$ satisfy the closure condition (8). Define the left side of (8) as

$$\Delta_n(\tau) \equiv 2t_n(\tau) - \beta\mathbb{E}[\theta|\,\theta \in [t_{n-1}(\tau), t_n(\tau)]] - \beta\mathbb{E}[\theta|\,\theta \geq t_n(\tau)]. \tag{14}$$

By the now familiar argument, $\Delta_n(\tau)$ is strictly increasing in τ, so there is a unique value $\tau = \tau_n^* = t_1^n$, that solves the equation. We note that the value of t_1^n is exactly τ^{n+1}, the value such that the next forward solution just goes out of the support. This implies that all forward solutions are well defined. It follows that for any n, we can construct an equilibrium. Moreover, in any such equilibrium, the value of the first threshold t_1^n is a decreasing function of n. □

Lemma A.5 *There exists an infinite equilibrium.*

Proof of Lemma A.5. We prove the result in four claims.

Claim 0) *The last equilibrium threshold t_n^n is bounded above for all n and $\lim_{n\to\infty} t_n^n < \infty$.*

Proof: The statement is trivial for $\overline{\mathcal{S}} < \infty$. We know from Lemma A.4 that the value of the first threshold t_1^n is a monotone decreasing function of n. Since the sequence is bounded by zero it must converge. Likewise, t_n^n is bounded above: consider the closure condition, $\Delta_n(\tau) = 0$, for $\tau = t_1^n$ and Δ_n defined in (14). We have

$$\begin{aligned}\Delta_n(t_1^n) &= 2t_n^n - \beta\mathbb{E}\left[\theta|\,\theta \in \left[t_{n-1}^n, t_n^n\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq t_n^n\right] \\ &\geq 2\left(t_n^n - \beta\mathbb{E}\left[\theta|\,\theta \geq t_n^n\right]\right),\end{aligned}$$

which follows from $-\beta\mathbb{E}\left[\theta|\,\theta \in \left[t_{n-1}^n, t_n^n\right]\right] \geq -\beta\mathbb{E}\left[\theta|\,\theta \geq t_n^n\right]$. For a logconcave distribution, $t - \beta\mathbb{E}\left[\theta|\,\theta \geq t\right]$ is negative for $t = 0$, increasing in t, and goes to ∞ for $t \to \infty$. Therefore, $\lim_{n\to\infty} t_n^n < \infty$ and the sequence t_n^n is bounded above. □

Claim 1) *The equilibrium features increasing intervals,*

$$t_{i+1}^n - t_i^n > t_i^n - t_{i-1}^n \quad \forall n \text{ and } \forall i < n.$$

Proof: Consider the equilibrium indifference condition for t_1,

$$t_1 - \beta\mathbb{E}\left[\theta|\,\theta \in [0, t_1]\right] = \beta\mathbb{E}\left[\theta|\,\theta \in [t_1, t_2]\right] - t_1.$$

Logconcave densities are unimodal. By symmetry, the mode is at 0 and hence the density truncated at zero is non-increasing. This implies that for an interval of given length λ,

$$\mathbb{E}\left[\theta|\,\theta \in [t_1, t_1 + \lambda]\right] \leq t_1 + \frac{\lambda}{2}.$$

Consider $t_1 = \lambda$ and $t_2 = 2\lambda$. Then, $\lambda - \beta\mathbb{E}[\theta|\, \theta \in [0, \lambda]] \geq \lambda - \beta\frac{\lambda}{2}$ and $\beta\mathbb{E}[\theta|\, \theta \in [\lambda, 2\lambda]] - \lambda \leq \beta\frac{3}{2}\lambda - \lambda$, where the inequalities are strict if the density is strictly decreasing. Since $\lambda - \beta\frac{\lambda}{2} > \beta\frac{3}{2}\lambda - \lambda$, t_2 must increase to satisfy the equilibrium condition.

Likewise, consider

$$t_i - \beta\mathbb{E}[\theta|\, \theta \in [t_{i-1}, t_i]] = \beta\mathbb{E}[\theta|\, \theta \in [t_i, t_{i+1}]] - t_i$$

and suppose $t_i - t_{i-1} = \lambda = t_{i+1} - t_i$. Then $\beta\mathbb{E}[\theta|\, \theta \in [t_i - \lambda, t_i]] \leq \beta\left(t_i - \frac{\lambda}{2}\right)$ and $\beta\mathbb{E}[\theta|\, \theta \in [t_i, t_i + \lambda]] \leq \beta\left(t_i + \frac{\lambda}{2}\right)$ (with strict inequalities for a strictly decreasing density) imply that $t_i - \beta\mathbb{E}[\theta|\, \theta \in [t_{i-1}, t_i]] \geq t_i - \beta\left(t_i - \frac{\lambda}{2}\right)$ and $\beta\mathbb{E}[\theta|\, \theta \in [t_i, t_{i+1}]] - t_i \leq \beta\left(t_i + \frac{\lambda}{2}\right) - t_i$. Since $t_i - \beta\left(t_i - \frac{\lambda}{2}\right) > \beta\left(t_i + \frac{\lambda}{2}\right) - t_i$ for all t_i, we must again have that $t_{i+1} - t_i > t_i - t_{i-1} = \lambda$ to restore equilibrium. □

Claim 2) *The sequence* $(t_1^n)_n$ *is monotone decreasing, while the sequence* $(t_n^n)_n$ *is monotone increasing. Moreover, equilibrium thresholds are nested,*

$$t_1^{n+1} < t_1^n < t_2^{n+1} < \cdots t_n^{n+1} < t_n^n < t_{n+1}^{n+1} \qquad \forall n. \tag{15}$$

Proof: Recall the notation $t_1^n = \tau^{n+1}$ and $t_1^{n+1} = \tau^{n+2}$ from Lemma A.4. Since by Lemma A.4 the solution of the forward equation is monotonically increasing in the initial condition, τ, we have that $t_i^{n+1} < t_i^n$ for $i = 1, \ldots, n$. Hence, it remains to prove that $t_i^n < t_{i+1}^{n+1}$ for $i = 1, \ldots, n$.

We start with two preliminary observations. First, the "next" solution of the forward equation, $t_{i+1}^k(\tau)$ for $i = 1, \ldots, k-1$, and $k = n, n+1$ is monotonic in $t_i^k(\tau)$, and the length of the previous interval, $t_i^k(\tau) - t_{i-1}^k(\tau)$. To see this, note that the forward equations for t_2^k, t_3^k, and t_{i+1}^k, for $i = 3, \ldots, k-1$ and $k = n, n+1$ satisfy:

$$\tau - \beta\mathbb{E}[\theta|\, \theta \in [0, \tau]] = \beta\mathbb{E}\left[\theta|\, \theta \in \left[t_2^k, \tau\right]\right] - \tau,$$

$$t_2^k(\tau) - \beta\mathbb{E}\left[\theta|\, \theta \in \left[\tau, t_2^k(\tau)\right]\right] = \beta\mathbb{E}\left[\theta|\, \theta \in \left[t_2^k(\tau), t_3^k\right]\right] - t_2^k(\tau),$$

and

$$t_i^k(\tau) - \beta\mathbb{E}\left[\theta|\, \theta \in \left[t_{i-1}^k(\tau), t_i^k(\tau)\right]\right] = \beta\mathbb{E}\left[\theta|\, \theta \in \left[t_i^k(\tau), t_{i+1}^k\right]\right] - t_i^k(\tau).$$

Let $t_{i-1}^k(\tau) = t_i^k(\tau) - \lambda$ and substitute into the forward equation for t_{i+1}^k:

$$t_i^k(\tau) - \beta\mathbb{E}\left[\theta|\, \theta \in \left[t_i^k(\tau) - \lambda, t_i^k(\tau)\right]\right] = \beta\mathbb{E}\left[\theta|\, \theta \in \left[t_i^k(\tau), t_{i+1}^k\right]\right] - t_i^k(\tau).$$

Monotonicity follows from the fact that $t_i^k(\tau)$ decreases the value of the right-hand side by logconcavity of the density and increases the value of the left-hand side again by that property. Moreover, an increase in λ increases the left-hand side further, implying that $t_{i+1}^k(\tau)$ has to increase to restore the equality.

Second, it is impossible that $t_{n+1}^{n+1} < t_n^n$ and $t_{n+1}^{n+1} - t_n^{n+1} < t_n^n - t_{n-1}^n$. If these conditions would hold, then one of the closure conditions,

$$0 = 2t_n^n - \beta\mathbb{E}\left[\theta|\,\theta \in \left[t_{n-1}^n, t_n^n\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq t_n^n\right]$$

and

$$0 = 2t_{n+1}^{n+1} - \beta\mathbb{E}\left[\theta|\,\theta \in \left[t_n^{n+1}, t_{n+1}^{n+1}\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq t_{n+1}^{n+1}\right]$$

would necessarily be violated. To see this, take $\delta_1, \delta_2 > 0$ and suppose that $t_{n+1}^{n+1} = t_n^n - \delta_1$, $t_n^n - t_{n-1}^n = \lambda$, and $t_{n+1}^{n+1} - t_n^{n+1} = \lambda - \delta_2$. Now consider the closure conditions

$$0 = 2t_n^n - \beta\mathbb{E}\left[\theta|\,\theta \in \left[t_n^n - \lambda, t_n^n\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq t_n^n\right]$$

and

$$0 = 2\left(t_n^n - \delta_1\right) - \beta\mathbb{E}\left[\theta|\,\theta \in \left[t_n^n - \delta_1 - \left(\lambda - \delta_2\right), t_n^n - \delta_1\right]\right] - \beta\mathbb{E}\left[\theta|\,\theta \geq t_n^n - \delta_1\right].$$

By logconcavity, $\delta_1 > 0$ reduces the right-hand side of the second condition. Moreover, $\delta_2 > 0$ increases the lower bound $t_n^n - \delta_1 - \lambda + \delta_2$, so decreases the right-hand side further. Hence, one of the closure conditions must necessarily be violated.

We now show that $t_{j+1}^{n+1} > t_j^n$ for all $j \leq n$. Suppose for contradiction that the property is violated for the first time at $j = l$. Suppose $t_{j+1}^{n+1} > t_j^n$ for all $j = 1, \ldots, l-1$ and $t_{l+1}^{n+1} < t_l^n$. Taken together, these inequalities immediately imply that $t_{l+1}^{n+1} - t_l^{n+1} < t_l^n - t_{l-1}^n$. In turn, the monotonicity property of the next forward solution implies that $t_{l+2}^{n+1} < t_{l+1}^n$.

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

However, if $t_{l+2}^{n+1} < t_{l+1}^n$ and $t_{l+2}^{n+1} - t_{l+1}^{n+1} < t_{l+1}^n - t_l^n$, then $t_{l+3}^{n+1} < t_{l+2}^n$ and so forth. Hence, we would have $t_{j+1}^{n+1} < t_j^n$ and $t_{j+1}^{n+1} - t_j^{n+1} < t_j^n - t_{j-1}^n$ for all $j \geq l$ and in particular for $j = n$, leading to a violation of one of the closure conditions.

The same argument can be given for a Class II equilibrium. This is omitted. □

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

Proof: The limit is an equilibrium if the equilibrium indifference conditions are satisfied and they satisfy the order $\lim_{n\to\infty}\beta\mu_i^n \leq \lim_{n\to\infty} t_i^n \leq \lim_{n\to\infty}\beta\mu_{i+1}^n$. The indifference conditions are satisfied by construction, so what remains to show is the order. For the first inequality, since $\beta < 1$ it suffices to show that equilibrium thresholds remain ordered in the limit, $\lim_{n\to\infty} t_i^n \leq \lim_{n\to\infty} t_{i+1}^n$: For all finite n, thresholds are ordered in equilibrium, $t_i^n < t_{i+1}^n$, since they are ordered for any forward equation. By Claim 2) equilibrium thresholds converge. Denote the limits by $t_i^\infty = \lim_{n\to\infty} t_i^n$ for all i. By convergence, for any ε there is N such that for all $n > N$: $t_i^n \geq t_i^\infty - \frac{\varepsilon}{2}$ and $t_{i+1}^n \leq t_{i+1}^\infty + \frac{\varepsilon}{2}$. Suppose for contradiction that $t_i^\infty \geq t_{i+1}^\infty + \delta$ for some $\delta > 0$; this implies

$$t_i^n \geq t_i^\infty - \frac{\varepsilon}{2} \geq t_{i+1}^\infty + \delta - \frac{\varepsilon}{2} \geq t_{i+1}^n - \frac{\varepsilon}{2} + \delta - \frac{\varepsilon}{2} > t_{i+1}^n,$$

for all $\varepsilon < \delta$.

For the second inequality, we have that for all finite n the forward equations imply that $t_i^n < \beta\mu_{i+1}^n$. As before, we denote the limits by by $t_i^\infty = \lim_{n\to\infty} t_i^n$ and $\mu_i^\infty = \lim_{n\to\infty}\mu_i^n$ for all i. By convergence, for any ε there is N such that for all $n > N$: $t_i^n \geq t_i^\infty - \frac{\varepsilon}{2}$ and $\beta\mu_{i+1}^n \leq \beta\mu_{i+1}^\infty + \frac{\varepsilon}{2}$. Suppose for contradiction that $t_i^\infty > \beta\mu_{i+1}^\infty$ implying that $t_i^\infty > \beta\mu_{i+1}^\infty + \delta$ for some $\delta > 0$. This implies

$$t_i^n \geq t_i^\infty - \frac{\varepsilon}{2} > \beta\mu_{i+1}^\infty + \delta - \frac{\varepsilon}{2} \geq \beta\mu_{i+1}^n - \frac{\varepsilon}{2} + \delta - \frac{\varepsilon}{2} > \beta\mu_{i+1}^n,$$

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

Proof of Lemma 1. Since $\mathbb{E}[\mu^n] = \mathbb{E}[\theta] = 0$ and $\mathbb{E}[\mu^n\Theta] = \mathbb{E}\mathbb{E}[\mu^n\Theta|\,\Theta \in [\theta_i, \theta_{i+1}]] = \mathbb{E}[(\mu^n)^2] = var(\mu^n)$, we have

$$\begin{aligned}\mathbb{E}u_R^{com}(y_R, \Theta, \beta) &= -\mathbb{E}\left[(\beta\mu^n - \beta\Theta)^2\right] = -\beta^2\mathbb{E}\left[(\mu^n)^2 - 2\mu^n\Theta + \Theta^2\right] \\ &= \beta^2\left(var(\mu^n) - \sigma^2\right).\end{aligned}$$

We now show that $var(\mu^n) = \ell(\beta, n)\sigma^2$, for some function $\ell(\beta, n)$ that is independent of σ^2.

Consider a typical equilibrium indifference condition

$$t_i - \beta\mathbb{E}\left[\Theta\right|\Theta \in [t_{i-1}, t_i]] = \beta\mathbb{E}\left[\Theta\right|\Theta \in [t_i, t_{i+1}]] - t_i.$$

A change of variables to $z = \frac{\theta}{\sigma}$, and thus $dz = \frac{1}{\sigma}d\theta$, results in

$$\mathbb{E}\left[\Theta\right|\Theta \in [t_{i-1}, t_i]] = \frac{\int\limits_{t_{i-1}}^{t_i} \theta c \frac{1}{\sigma}\psi\left(\frac{\theta^2}{\sigma^2}\right) d\theta}{\Pr\left(\Theta \in [t_{i-1}, t_i]\right)} = \frac{\sigma \int\limits_{z_{i-1}}^{z_i} zc\psi\left(z^2\right) dz}{\Pr\left(Z \in [z_{i-1}, z_i]\right)} = \sigma\mathbb{E}\left[Z\right|\Theta \in [z_{i-1}, z_i]],$$

with $z_i = \frac{t_i}{\sigma}$. Hence, the indifference condition can be written as

$$z_i - \beta\mathbb{E}\left[Z\right| Z \in [z_{i-1}, z_i]] = \beta\mathbb{E}\left[Z\right| Z \in [z_i, z_{i+1}]] - z_i,$$

which is independent of the variance. As a consequence, the standardized equilibrium thresholds z_i are independent of the variance.

It follows that $var(\mu^n)$ is linear in σ^2, $var(\mu^n) = \ell(n, \beta)\sigma^2$, where $\ell(n, \beta)$ is independent of σ^2. □

Proof of Lemma 2.

$$\mathbb{E}u_R^{del}\left(y_S, \Theta, \beta\right) = \mathbb{E}\left[-\left(\Theta - \beta\Theta\right)^2\right] = -\left(1-\beta\right)^2\sigma^2.$$

If the receiver chooses the prior optimal action 0, then $\mathbb{E}u_R\left(0, \Theta\right) = -\beta^2\sigma^2$, implying the statement. □

Proof of Proposition 2. Straightforward integration gives for any $[a, b] \subseteq \left[0, -\frac{s}{\delta}\right]$,

$$\mathbb{E}\left[\Theta|\Theta \in [a, b]\right] = \frac{s+b}{1-\delta} - \frac{1}{1-\delta}\frac{(b-a)}{1-\left(\frac{1+\frac{\delta}{s}b}{1+\frac{\delta}{s}a}\right)^{-\frac{1}{\delta}}}. \tag{16}$$

For the special case of $b = -\frac{s}{\delta}$ and $a \in \left[0, -\frac{s}{\delta}\right]$, we get

$$\mathbb{E}\left[\Theta|\Theta \geq a\right] = \mathbb{E}\left[\Theta|\Theta \geq 0\right] + \frac{1}{1-\delta} \cdot a = \frac{s+a}{1-\delta}. \tag{17}$$

Hence, the generalized Pareto distribution features linear tail conditional expectations. Therefore, we can apply the value characterization of Deimen and Szalay

(2019), which derives the expected utility of a limit equilibrium given in (4) as an upper bound on the expected utilities of finite equilibria given in (3). Deimen and Szalay (2019) shows that the limit equilibrium exists for the special case of $\delta = 0$. Here, we extend the proof of existence in Proposition 1 to the class of all logconcave densities, which includes the generalized Pareto distribution with $\delta \in [-1, 0]$. $\square$

Proof of Theorem 2. One can show that the limit equilibrium yields a higher payoff than any finite equilibrium in the communication game. Compare the receiver's expected utility in a limit equilibrium under communication $\mathbb{E}u_R(\beta\mu^\infty, \Theta, \beta) = \beta^2 (var(\mu^\infty) - \sigma^2) = \beta^2 \left(\frac{2 - \frac{1}{1-\delta}}{2 - \frac{\beta}{1-\delta}} \sigma^2 - \sigma^2 \right) = -\beta^2 \sigma^2 \frac{1-\beta}{2-\beta-2\delta}$ to the receiver's expected utility under delegation $\mathbb{E}u_R(\Theta, \Theta, \beta) = -(1-\beta)^2 \sigma^2$. The receiver prefers delegation over communication if

$$-(1-\beta)^2 \sigma^2 \geq -\beta^2 \sigma^2 \frac{1-\beta}{2-\beta-2\delta} \quad \Leftrightarrow \quad \delta \geq \frac{2-3\beta}{2-2\beta}.$$

$\square$

Proof of Proposition 3. The proof consists of three parts. Part i) establishes that the densities f_+ and g_+ – when scaled to the same support – satisfy the monotone likelihood ratio property (mlrp). Part ii) shows that, on the same support, if all receiver best responses on the positive half of the support are higher for a given truncation, then the equilibrium critical types and the equilibrium receiver actions must be higher. In Part iii), we scale back to the original supports and prove that equilibrium critical types and equilibrium receiver actions are higher under f_+ than under g_+ relative to the length of the support.

Part i) We define the distributions $F(\theta)$ and $G(\theta)$ with densities $f(\theta) := f(\theta, s', \delta')$ and $g(\theta) := f(\theta, s'', \delta'')$ where $\delta'' > \delta'$ and $s'' < s'$ are such that the variances are identical $\sigma^2(s', \delta') = \sigma^2(s'', \delta'') = \sigma^2$. The distributions conditional on the positive half have pdfs $f_+(\theta), g_+(\theta)$ and cdfs $F_+(\theta), G_+(\theta)$.

Consider the rescaled density $\hat{f}_+ := f_+(\theta, \hat{s}, \delta')$ with scale $\hat{s}$ satisfying $-\frac{\hat{s}}{\delta'} = -\frac{s''}{\delta''}$. The densities $\hat{f}_+$ and g_+ have thus exactly the same supports (the auxiliary density $\hat{f}_+$ now induces a higher variance than g; in Part iii) of the proof we scale back to the same variance). The ratio of the likelihoods

$$\frac{\hat{f}_+(\theta)}{g_+(\theta)} = \frac{s''}{\hat{s}}\left(1+\frac{\delta'}{\hat{s}}\theta\right)^{\frac{1}{\delta''}-\frac{1}{\delta'}} \tag{18}$$

is monotonically increasing. The monotone likelihood ratio property is preserved under truncation to an arbitrary interval $[t_{i-1}, t_i]$,

$$\frac{\partial}{\partial\theta}\frac{\frac{\hat{f}_+(\theta)}{\hat{F}_+(t_i)-\hat{F}_+(t_{i-1})}}{\frac{g_+(\theta)}{G_+(t_i)-G_+(t_{i-1})}} = \frac{G_+(t_i)-G_+(t_{i-1})}{\hat{F}_+(t_i)-\hat{F}_+(t_{i-1})}\frac{\partial}{\partial\theta}\frac{\hat{f}_+(\theta)}{g_+(\theta)} > 0.$$

Part ii) Consider the positive half of the support and denote by n the number of positive critical types. In this proof, we denote equilibrium thresholds by $t^n_{i,h}$ and the forward solutions by $t_{i,h}$ under distribution $h \in \{\hat{f}, g\}$ for all i. The equilibrium conditions for $\left(t^n_{i,\hat{f}}\right)_i$ and $(t^n_{i,g})_i$ are given by

$$t^n_{i,\hat{f}} - \beta\mathbb{E}_{\hat{f}}\left[\Theta|\,\Theta \in \left[t^n_{i-1,\hat{f}}, t^n_{i,\hat{f}}\right]\right] = \beta\mathbb{E}_{\hat{f}}\left[\Theta|\,\Theta \in \left[t^n_{i,\hat{f}}, t^n_{i+1,\hat{f}}\right]\right] - t^n_{i,\hat{f}},$$

$$t^n_{i,g} - \beta\mathbb{E}_{g}\left[\tilde{\Theta}\middle|\,\tilde{\Theta} \in \left[t^n_{i-1,g}, t^n_{i,g}\right]\right] = \beta\mathbb{E}_{g}\left[\tilde{\Theta}\middle|\,\tilde{\Theta} \in \left[t^n_{i,g}, t^n_{i+1,g}\right]\right] - t^n_{i,g},$$

for $i = 1,\dots,n$, where for $i = 1$ by construction $t^n_{0,h} = 0$, and for $i = n$ we define $t^n_{n+1,h} = \overline{\mathcal{S}}_h$. Note that by monotonicity of the likelihood ratio given in equation (18) all means are ordered: $\mathbb{E}_{\hat{f}}\left[\Theta|\,\Theta \in [t_{i-1}, t_i]\right] > \mathbb{E}_g[\tilde{\Theta}|\tilde{\Theta} \in [t_{i-1}, t_i]]$. We want to show that all equilibrium thresholds are higher under distribution $\hat{F}$ than under distribution G, $t^n_{i,g} < t^n_{i,\hat{f}}$ for $i \le n$.

We first prove two useful claims.

Claim 1 *Forward solutions for a given initial condition $\tau = t_{1,\hat{f}} = t_{1,g}$ are higher under distribution G than under distribution $\hat{F}$, $t_{i,g} > t_{i,\hat{f}}$ for $i \le n$.*

Proof. Fix $t_1 = \tau$ and consider $t_{2,\hat{f}}(\tau)$. By mlrp we have

$$\begin{aligned} 0 &= 2\tau - \beta\mathbb{E}_{\hat{f}}\left[\Theta|\,\Theta \in [0,\tau]\right] - \beta\mathbb{E}_{\hat{f}}\left[\Theta|\,\Theta \in \left[\tau, t_{2,\hat{f}}(\tau)\right]\right] \\ &< 2\tau - \beta\mathbb{E}_g[\tilde{\Theta}|\tilde{\Theta} \in [0,\tau]] - \beta\mathbb{E}_g[\tilde{\Theta}|\tilde{\Theta} \in \left[\tau, t_{2,\hat{f}}(\tau)\right]]. \end{aligned}$$

Hence, we need $t_{2,g}(\tau) > t_{2,\hat{f}}(\tau)$ to restore equality with zero.

Next, we show that $t_{i,g}(\tau) > t_{i,\hat{f}}(\tau)$ for $i \leq 3$. The forward equation for $t_{3,\hat{f}}(\tau)$ is

$$0 = 2t_{2,\hat{f}}(\tau) - \beta\mathbb{E}_{\hat{f}}\left[\Theta|\,\Theta \in \left[\tau, t_{2,\hat{f}}(\tau)\right]\right] - \beta\mathbb{E}_{\hat{f}}\left[\Theta|\,\Theta \in \left[t_{2,\hat{f}}(\tau), t_{3,\hat{f}}(\tau)\right]\right].$$

Note that, since forward solutions are increasing in the initial condition, we can choose $\tau_2 < \tau$ such that $t_{2,g}(\tau_2) = t_{2,\hat{f}}(\tau)$. The value τ_2 is well defined since τ is well defined under $\hat{f}$. Fixing t_3 at $t_{3,\hat{f}}(\tau)$, we observe that by mlrp and $\tau_2 < \tau$

$$0 < 2t_{2,g}(\tau_2) - \beta\mathbb{E}_g\left[\tilde{\Theta}\middle|\,\tilde{\Theta} \in [\tau_2, t_{2,g}(\tau_2)]\right] - \beta\mathbb{E}_g\left[\tilde{\Theta}\middle|\,\tilde{\Theta} \in \left[t_{2,g}(\tau_2), t_{3,\hat{f}}(\tau)\right]\right].$$

Hence, to restore equality with zero, we need $t_{3,g}(\tau_2) > t_{3,\hat{f}}(\tau)$. Note, for future reference, that we have to choose $\tau_3 < \tau_2$ if we want to equalize $t_{3,g}(\tau_3) = t_{3,\hat{f}}(\tau)$. Finally, we need to increase τ_2 back to the original level τ. Since, by Lemma A.2 the solutions to the forward equations are increasing in the initial condition, we have $t_{i,g}(\tau) > t_{i,\hat{f}}(\tau)$ for $i \leq 3$.

Note that for each k we can choose τ_k such that $t_{k,g}(\tau_k) = t_{k,\hat{f}}(\tau)$. Suppose as an inductive hypothesis that $\tau_k < \tau_{k-1}$. The forward equation for $t_{k+1,\hat{f}}(\tau)$ is

$$0 = 2t_{k,\hat{f}}(\tau) - \beta\mathbb{E}_{\hat{f}}\left[\Theta|\,\Theta \in \left[t_{k-1,\hat{f}}(\tau), t_{k,\hat{f}}(\tau)\right]\right] - \beta\mathbb{E}_{\hat{f}}\left[\Theta|\,\Theta \in \left[t_{k,\hat{f}}(\tau), t_{k+1,\hat{f}}(\tau)\right]\right].$$

Adjusting τ_k so that $t_{k,g}(\tau_k) = t_{k,\hat{f}}(\tau)$, we have

$$0 < 2t_{k,g}(\tau_k) - \beta\mathbb{E}_g\left[\tilde{\Theta}\middle|\,\tilde{\Theta} \in [t_{k-1,g}(\tau_k), t_{k,g}(\tau_k)]\right] - \beta\mathbb{E}_g\left[\tilde{\Theta}\middle|\,\tilde{\Theta} \in \left[t_{k,g}(\tau_k), t_{k+1,\hat{f}}(\tau)\right]\right]$$

since $t_{k-1,g}(\tau_k) < t_{k-1,g}(\tau_{k-1}) = t_{k-1,\hat{f}}(\tau)$ by the inductive hypothesis.

It follows that $t_{k+1,g}(\tau_k) > t_{k+1,\hat{f}}(\tau)$ to restore equality with zero. Moreover, to obtain $t_{k+1,g}(\tau_{k+1}) = t_{k+1,\hat{f}}(\tau)$ we must have $\tau_{k+1} < \tau_k$.

Finally, we need to increase τ_k back to the original level τ. Since the solutions to the forward equations are increasing in the initial condition and $\tau_i < \tau_{i-1} < \ldots < \tau$ for $i \leq k$, we have $t_{i,g}(\tau) > t_{i,\hat{f}}(\tau)$ for $i \leq k$. □

Note that, similar to the forward equation, which takes the starting point as given, we can compute thresholds from a backward equation, which takes the last threshold $t_n = \overline{\tau}$ as given.

Claim 2 *The backward equation satisfies* $\frac{dt_i}{dt_{i+1}} > 1$ *for all* $i < n$. *Moreover, for a given initial condition* $\overline{\tau} = t_{n,\hat{f}} = t_{n,g}$ *the backward solutions are higher under distribution* G *than under distribution* $\hat{F}$, $t_{i,g} > t_{i,\hat{f}}$ *for* $i < n$.

Proof. Since the first part of the claim holds for any logconcave density, we skip the dependency on the distribution in this part. Consider the backward equation that determines $t_{n-1} = t_{n-1}(\overline{\tau})$ as a function of $\overline{\tau}$,

$$2\overline{\tau} - \beta\mathbb{E}\left[\Theta|\,\Theta \in [t_{n-1}, \overline{\tau}]\right] - \beta\mathbb{E}\left[\Theta|\,\Theta \in \left[\overline{\tau}, \overline{\mathcal{S}}\right]\right] = 0.$$

Totally differentiating and rearranging, we obtain

$$\frac{dt_{n-1}}{d\overline{\tau}} = \frac{\left(2 - \beta\frac{\partial}{\partial\overline{\tau}}\mathbb{E}\left[\Theta|\,\Theta \in [t_{n-1}, \overline{\tau}]\right] - \beta\frac{\partial}{\partial\overline{\tau}}\mathbb{E}\left[\Theta|\,\Theta \in \left[\overline{\tau}, \overline{\mathcal{S}}\right]\right]\right)}{\beta\frac{\partial}{\partial t_{n-1}}\mathbb{E}\left[\Theta|\,\Theta \in [t_{n-1}, \overline{\tau}]\right]}.$$

We find that $\frac{dt_{n-1}}{d\overline{\tau}} \geq 1$.

Suppose as an inductive hypothesis that $\frac{dt_i}{dt_{i+1}} \geq 1$ for $i \geq k$. Totally differentiating the backward equation for t_{k-1} and rearranging we get

$$\begin{aligned}\frac{dt_{k-1}}{dt_k} &= \frac{2 - \beta\frac{\partial}{\partial t_k}\mathbb{E}\left[\Theta|\,\Theta \in [t_{k-1}, t_k]\right] - \beta\frac{\partial}{\partial t_k}\mathbb{E}\left[\Theta|\,\Theta \in [t_k, t_{k+1}]\right] - \beta\frac{\partial}{\partial t_{k+1}}\mathbb{E}\left[\Theta|\,\Theta \in [t_k, t_{k+1}]\right]\frac{dt_{k+1}}{dt_k}}{\beta\frac{\partial}{\partial t_{k-1}}\mathbb{E}\left[\Theta|\,\Theta \in [t_{k-1}, t_k]\right]} \\ &\geq 1,\end{aligned}$$

where the inequality follows from the inductive hypothesis $\frac{dt_{k+1}}{dt_k} \leq 1$, and from log-concavity of the density and Lemma A.1.

The proof of the second statement is analogous to the proof of Claim 1. □

We are now ready to prove Part ii).

$\boldsymbol{i = 1}$. The first equilibrium threshold under distribution $\hat{F}_+$ is higher than the first equilibrium threshold under G_+, $t^n_{1,g} < t^n_{1,\hat{f}}$:

We note that the equilibrium values of the thresholds are necessarily solutions of the forward equations. Consider thus the equilibrium condition for $t^n_{n,\hat{f}}$. We can write for $\tau = t^n_{1,\hat{f}}$

$$2t_{n,\hat{f}}(\tau) - \beta\mathbb{E}_{\hat{f}}\left[\Theta|\,\Theta \in \left[t_{n-1,\hat{f}}(\tau), t_{n,\hat{f}}(\tau)\right]\right] - \beta\mathbb{E}_{\hat{f}}\left[\Theta|\,\Theta \in \left[t_{n,\hat{f}}(\tau), \overline{\mathcal{S}}\right]\right] = 0.$$

Choose $\tau_{n-1} < \tau$ such that $t_{n-1,g}(\tau_{n-1}) = t_{n-1,\hat{f}}(\tau)$, implying by Claim 1 that $t_{n,g}(\tau_{n-1}) > t_{n,\hat{f}}(\tau)$. By logconcavity, Lemma A.1 implies that

$$2t_{n,g}(\tau_{n-1}) - \beta\mathbb{E}_{\hat{f}}[\Theta|\,\Theta \in [t_{n-1,g}(\tau_{n-1}), t_{n,g}(\tau_{n-1})]] - \beta\mathbb{E}_{\hat{f}}\left[\Theta|\,\Theta \in \left[t_{n,g}(\tau_{n-1}), \overline{\mathcal{S}}\right]\right] > 0.$$

By mlrp,

$$2t_{n,g}(\tau_{n-1}) - \beta\mathbb{E}_{g}\left[\tilde{\Theta}\middle|\,\tilde{\Theta} \in [t_{n-1,g}(\tau_{n-1}), t_{n,g}(\tau_{n-1})]\right] - \beta\mathbb{E}_{g}\left[\tilde{\Theta}\middle|\,\tilde{\Theta} \in \left[t_{n,g}(\tau_{n-1}), \overline{\mathcal{S}}\right]\right] > 0.$$

Finally, increasing τ_{n-1} to τ, we obtain by Lemma A.1 and Lemma A.2

$$2t_{n,g}(\tau) - \beta\mathbb{E}_{g}\left[\tilde{\Theta}\middle|\,\tilde{\Theta} \in [t_{n-1,g}(\tau), t_{n,g}(\tau)]\right] - \beta\mathbb{E}_{g}\left[\tilde{\Theta}\middle|\,\tilde{\Theta} \in \left[t_{n,g}(\tau), \overline{\mathcal{S}}\right]\right] > 0.$$

Hence, to restore equilibrium – i.e., equality with zero – we need to have $t^n_{1,g} < t^n_{1,\hat{f}}$.

$\boldsymbol{i = n}$. The last equilibrium threshold under distribution $\hat{F}_+$ is higher than the last equilibrium threshold under G_+, $t^n_{n,g} < t^n_{n,\hat{f}}$.

This can be shown analogously to the proof that $t^n_{1,g} < t^n_{1,\hat{f}}$, using the backward equation instead of the forward equation.

$\boldsymbol{1 < i < n}$. The equilibrium thresholds under distribution $\hat{F}_+$ are higher than the equilibrium thresholds under G_+, $t^n_{i,g} < t^n_{i,\hat{f}}$, for $1 < i < n$:

We know that $t^n_{1,g} < t^n_{1,\hat{f}}$. Now, either $t^n_{i,g} < t^n_{i,\hat{f}}$ is satisfied for $i = 1, \ldots, n$, which completes the proof, or there must exist some k such that $t^n_{i,g} < t^n_{i,\hat{f}}$ for $i \leq k - 1$ and $t^n_{k,g} \geq t^n_{k,\hat{f}}$. Recall that the equilibrium thresholds are also forward equations. By the proof of Claim 1, this implies that the initial condition $t^n_{1,g}$ must satisfy $t^n_{1,g} \in (\tau_k, \tau_{k-1})$. As a consequence, $t_{i,g}(t^n_{1,g}) > t_{i,\hat{f}}(t^n_{1,\hat{f}})$ for $i > k + 1$, in particular for $i = n$. This contradicts $t^n_{n,g} < t^n_{n,\hat{f}}$.

Part iii) Recall from Part i) that $\overline{\mathcal{S}}_f = -\frac{s'}{\delta'}$, $\overline{\mathcal{S}}_g = -\frac{s''}{\delta''}$, and that the random variable Θ with distribution f is scaled up to the same support of random variable $\tilde{\Theta}$ with distribution g, $\overline{\mathcal{S}}_{\hat{f}} = \overline{\mathcal{S}}_g$. The scale $\hat{s}$ needed to equalize supports satisfies $-\frac{\hat{s}}{\delta'} = -\frac{s''}{\delta''}$. We can write $-\frac{s'}{\delta'}\frac{\hat{s}}{s'} = -\frac{s''}{\delta''}$ and define

$$\gamma := \frac{s'}{\hat{s}} = \frac{\frac{s'}{\delta'}}{\frac{s''}{\delta''}}.$$

The factor $\gamma < 1$ adjusts the level of the scale parameter back to its original level.

We can replicate the result of Lemma 1 – which was stated for the variance – for the scale parameter: Consider a typical indifference condition for type t_i, $t_i - \beta\mathbb{E}\left[\Theta|\,\Theta \in [t_{i-1}, t_i]\right] = \beta\mathbb{E}\left[\Theta|\,\Theta \in [t_i, t_{i+1}]\right] - t_i$ and substitute the specific expression for truncated means for the generalized Pareto distribution from equation (16),

$$t_i - \beta\left(\frac{s + s\frac{t_i}{s}}{1-\delta} - \frac{1}{1-\delta}\frac{s\left(\frac{t_i}{s} - \frac{t_{i-1}}{s}\right)}{1 - \left(\frac{1+\frac{\delta}{s}t_i}{1+\frac{\delta}{s}t_{i-1}}\right)^{-\frac{1}{\delta}}}\right) = \beta\left(\frac{s + s\frac{t_{i+1}}{s}}{1-\delta} - \frac{1}{1-\delta}\frac{s\left(\frac{t_{i+1}}{s} - \frac{t_i}{s}\right)}{1 - \left(\frac{1+\frac{\delta}{s}t_{i+1}}{1+\frac{\delta}{s}t_i}\right)^{-\frac{1}{\delta}}}\right) - t_i.$$

This makes it evident that we can change variables to $\zeta_i = \frac{t_i}{s}$ and solve for the equilibrium in the scale-normalized space. The equilibrium critical types ζ_i^n are scale free. Therefore, t_i^n is linear in s. We can thus scale back and write $t_{i,f}^n = \gamma \cdot t_{i,\hat{f}}^n$ and $\mu_{i,f}^n = \gamma \cdot \mu_{i,\hat{f}}^n$. By Part ii), $t_{i,\hat{f}}^n > t_{i,g}^n$ and $\mu_{i,\hat{f}}^n > \mu_{i,g}^n$ and so

$$t_{i,f}^n > \gamma \cdot t_{i,g}^n \quad \text{and} \quad \mu_{i,f}^n > \gamma \cdot \mu_{i,g}^n,$$

which is equivalent to the statement in the proposition. □

Proof of Lemma 3. By symmetry, consider the positive half of the distribution. Under delegation, the loss is $(1-\beta)^2\theta^2$ with maximum $(1-\beta)^2\overline{\mathcal{S}}^2$. Under communication, the receiver's loss conditional on $\theta \in \left[t_{i-1}^n, t_i^n\right]$ is $\beta^2\left(\mu_i^n - \theta\right)^2$. Since the density is decreasing, we have $\mu_i^n \leq \frac{t_{i-1}^n + t_i^n}{2}$, implying that the loss is maximal for $\theta = t_i^n$. Moreover, the relevant expression $t - \mu\left([t-\Delta, t]\right)$ is increasing in the length of the interval Δ and increasing in the location of the interval t, due to logconcavity of the density (Lemma A.1). With the convention that $t_{n+1}^n = \overline{\mathcal{S}}$, it follows that $\arg\max_i \beta^2\left(\mu_i^n - t_i^n\right)^2 = n+1$. Since intervals are increasing in i, the largest loss under communication is $\beta^2\left(\mu_{n+1}^n - \overline{\mathcal{S}}\right)^2$. □

Proof of Lemma 4. Using expression (16), we can write the indifference condition for type t_n as

$$t_n - \beta\left(\frac{s + t_n}{1-\delta} - \frac{1}{1-\delta}\frac{(t_n - t_{n-1})}{1 - \left(\frac{1+\frac{\delta}{s}t_n}{1+\frac{\delta}{s}t_{n-1}}\right)^{-\frac{1}{\delta}}}\right) = \beta\frac{s + t_n}{1-\delta} - t_n.$$

Note that the mean $\mathbb{E}\left[\Theta|\,\Theta\in[t_{n-1},t_n]\right]=\frac{s+t_n}{1-\delta}-\frac{1}{1-\delta}\frac{(t_n-t_{n-1})}{1-\left(\frac{1+\frac{\delta}{s}t_n}{1+\frac{\delta}{s}t_{n-1}}\right)^{-\frac{1}{\delta}}}$ is increasing in the bounds, t_{n-1} and t_n. To make the analysis transparent, we add $t_n\beta\frac{\delta}{\delta-1}$ to each side and rearrange,

$$\beta\frac{1}{1-\delta}\frac{(t_n-t_{n-1})}{1-\left(\frac{1+\frac{\delta}{s}t_n}{1+\frac{\delta}{s}t_{n-1}}\right)^{-\frac{1}{\delta}}}-\beta\frac{s}{1-\delta}+t_n\beta\frac{\delta}{\delta-1}=\beta\frac{s}{1-\delta}+2\beta\frac{t_n}{1-\delta}-2t_n+t_n\beta\frac{\delta}{\delta-1}.$$

Now the left side is non-negative for $t_{n-1}<t_n$ and converges to zero for $t_{n-1}\to t_n$. To see this, note that $\lim_{t_{n-1}\to t_n}\beta\frac{1}{1-\delta}\frac{(t_n-t_{n-1})}{1-\left(\frac{1+\frac{\delta}{s}t_n}{1+\frac{\delta}{s}t_{n-1}}\right)^{-\frac{1}{\delta}}}=\beta\frac{t_n\delta+s}{1-\delta}$, where the left side is decreasing in t_{n-1}. The right side becomes negative for any t_n higher than $T(\beta,\delta,s)$ as defined in the lemma. This implies that t_n must be bounded by T. □

Proof of Lemma 5. Since the supports are assumed to be $\mathbb{R}$, we have $supp(f)\subseteq supp(g)$. It remains to be shown that the ratio $\frac{f_+(\theta)}{g_+(\theta)}$ is unimodal with mode m an interior maximum.

Logconcavity of the ratio $\frac{f_+(\theta)}{g_+(\theta)}$ is equivalent to $\frac{\partial}{\partial\theta}\left(\frac{\frac{\partial}{\partial\theta}f_+(\theta)}{f_+(\theta)}-\frac{\frac{\partial}{\partial\theta}g_+(\theta)}{g_+(\theta)}\right)\leq 0$. That the difference is falling implies that one of three cases holds: either the difference is positive for all θ, $\frac{\frac{\partial}{\partial\theta}f_+(\theta)}{f_+(\theta)}>\frac{\frac{\partial}{\partial\theta}g_+(\theta)}{g_+(\theta)}$, negative for all θ, $\frac{\frac{\partial}{\partial\theta}f_+(\theta)}{f_+(\theta)}<\frac{\frac{\partial}{\partial\theta}g_+(\theta)}{g_+(\theta)}$, or changes sign once, i.e., there is some value m such that $\frac{\frac{\partial}{\partial\theta}f_+(\theta)|_{\theta=m}}{f_+(m)}=\frac{\frac{\partial}{\partial\theta}g_+(\theta)|_{\theta=m}}{g_+(m)}$ and $\frac{\frac{\partial}{\partial\theta}f_+(\theta)}{f_+(\theta)}>\frac{\frac{\partial}{\partial\theta}g_+(\theta)}{g_+(\theta)}$ for $\theta\in[0,m)$ and $\frac{\frac{\partial}{\partial\theta}f_+(\theta)}{f_+(\theta)}<\frac{\frac{\partial}{\partial\theta}g_+(\theta)}{g_+(\theta)}$ for $\theta\in(m,\overline{\mathcal{S}}]$.

The first two cases amount to mlrp on the positive half and can be ruled out by the following argument: Monotonicity of the likelihood ratio for all $\theta>0$ implies that $F_+(\theta)$ and $G_+(\theta)$ are ranked in the standard stochastic order (one distribution first order stochastically dominates the other one, FOSD). By symmetry, this implies that $F(\theta)$ and $G(\theta)$ are ordered in the convex order (SOSD). Finally, this implies that the distributions must have different variances, contradicting our assumption.

Hence, case three applies, implying that $\frac{f_+}{g_+}$ is unimodal with unique interior mode m. By concavity the mode is a maximum. □

Lemma A.6 *(Metzger and Rüschendorf (1991))*
Let $\frac{f_+(\theta)}{g_+(\theta)}$ be unimodal with interior mode m. The function $\frac{F_+(x)}{G_+(x)}$ inherits unimodality with mode $m_1 > m$, the function $\frac{(1-F_+(x))}{(1-G_+(x))}$ inherits unimodality with mode $m_2 < m$. Moreover, there exists a unique $\hat{x}$ such that $F_+(\theta) < G_+(\theta)$ for $\theta \in (0, \hat{x})$, $F_+(\hat{x}) = G_+(\hat{x})$, and $F_+(\theta) > G_+(\theta)$ for $\theta \in (\hat{x}, \infty)$.

Proof. Metzger and Rüschendorf (1991) Section 2. □

For the following lemma, since $\int_x^{\overline{\mathcal{S}}_h} (1 - H_+(\theta))\, d\theta = \int_x^{\infty} (1 - H_+(\theta))\, d\theta$ as $H_+(\theta) = 1$ for $\theta \geq \overline{\mathcal{S}}_h$, we unify notation and write $\int_x^\infty$ for infinite as well as for finite supports, $[0, \overline{\mathcal{S}}_h]$.

Lemma A.7 *(i) Let m denote the mode of the function $\frac{f_+(\theta)}{g_+(\theta)}$. Conditional on $\theta \in [0, m)$, the distributions f_+ and g_+ satisfy the monotone likelihood ratio property.*

(ii) The function $\frac{\int_x^\infty (1-F_+(\theta))d\theta}{\int_x^\infty (1-G_+(\theta))d\theta}$ is unimodal in $x \in [0, \overline{\mathcal{S}}_f]$ with mode $m' \in (0, m_2)$; for $0 \leq x \leq (<)\, m'$, we have $\mathbb{E}_f [\Theta | \Theta \geq x] \geq (>)\, \mathbb{E}_g \left[\tilde{\Theta} \middle| \tilde{\Theta} \geq x \right]$.

Proof of Lemma A.7. (i) Follows from the proof of Lemma 5.

(ii) We first show that $\frac{\int_x^\infty (1-F_+(\theta))d\theta}{\int_x^\infty (1-G_+(\theta))d\theta}$ is unimodal with mode m'. We then show that the mode m' is interior.

Straightforward differentiation gives

$$\frac{\partial}{\partial x} \frac{\int_x^\infty (1 - F_+(\theta))\, d\theta}{\int_x^\infty (1 - G_+(\theta))\, d\theta} = \frac{-(1 - F_+(x)) \int_x^\infty (1 - G_+(\theta))\, d\theta + (1 - G_+(x)) \int_x^\infty (1 - F_+(\theta))\, d\theta}{\left(\int_x^\infty (1 - G_+(\theta))\, d\theta \right)^2}.$$

The sign of the derivative is positive if and only if

$$(1 - F_+(x)) \int_x^\infty (1 - G_+(\theta))\, d\theta < (1 - G_+(x)) \int_x^\infty (1 - F_+(\theta))\, d\theta.$$

Note that by an integration by parts for any $x \in [0, \overline{\mathcal{S}}_h)$, we have that for $h_+ \in \{f_+, g_+\}$ and $H_+ \in \{F_+, G_+\}$

$$\mathbb{E}\left[\Theta \middle| \Theta \geq x\right] = \frac{\int_x^\infty \theta h_+(\theta)\, d\theta}{1 - H_+(x)} = x + \frac{\int_x^\infty (1 - H_+(\theta))\, d\theta}{1 - H_+(x)}.$$

Hence, $\frac{\partial}{\partial x} \frac{\int_x^\infty (1-F_+(\theta))d\theta}{\int_x^\infty (1-G_+(\theta))d\theta} \gtreqless 0$ if and only if $\mathbb{E}_f\left[\Theta \middle| \Theta \geq x\right] \gtreqless \mathbb{E}_g\left[\tilde{\Theta} \middle| \tilde{\Theta} \geq x\right]$.

Since a mode is an extremum, it is either at the boundary or satisfies the first order condition $\mathbb{E}_f\left[\Theta \middle| \Theta \geq x^*\right] = \mathbb{E}_g\left[\tilde{\Theta} \middle| \tilde{\Theta} \geq x^*\right]$. We next prove that there is at most one such value $x^* = m'$.

By Lemma A.6, the function $\frac{(1-F_+(x))}{(1-G_+(x))}$ is unimodal with mode m_2. Thus for $x \geq m_2$ the function is decreasing, equivalent to the conditional distribution of $\tilde{\Theta}$ conditional on $\tilde{\Theta} \geq x$ under distribution G_+ first order stochastically dominating the conditional distribution of Θ conditional on $\Theta \geq x$ under F_+: for $x \geq m_2$,

$$\frac{1 - F_+(x)}{1 - G_+(x)} > \frac{1 - F_+(\theta)}{1 - G_+(\theta)} \quad \Leftrightarrow \quad \frac{F_+(\theta) - F_+(x)}{1 - F_+(x)} > \frac{G_+(\theta) - G_+(x)}{1 - G_+(x)}.$$

By implication, for $x \geq m_2$ we have $\mathbb{E}_f\left[\Theta \middle| \Theta \geq x\right] < \mathbb{E}_g\left[\tilde{\Theta} \middle| \tilde{\Theta} \geq x\right]$ and $\frac{\int_x^\infty (1-F_+(\theta))d\theta}{\int_x^\infty (1-G_+(\theta))d\theta}$ is strictly decreasing.

For $x^* < m_2$, recall that by the first order condition we have

$$-\left(1 - F_+(x^*)\right) \int_{x^*}^\infty (1 - G_+(\theta))\, d\theta + \left(1 - G_+(x^*)\right) \int_{x^*}^\infty (1 - F_+(\theta))\, d\theta = 0.$$

Differentiating a second time and evaluating at x^*, we get

$$\begin{aligned} & f_+(x^*) \int_{x^*}^\infty (1 - G_+(\theta))\, d\theta - g_+(x^*) \int_{x^*}^\infty (1 - F_+(\theta))\, d\theta \\ < \;\; & g_+(x^*) \frac{1 - F_+(x)}{(1 - G_+(x))} \int_{x^*}^\infty (1 - G_+(\theta))\, d\theta - g_+(x^*) \int_{x^*}^\infty (1 - F_+(\theta))\, d\theta = 0, \end{aligned}$$

where the equality follows from the first order condition. For the inequality note that the function $\frac{(1-F_{+}(x))}{(1-G_{+}(x))}$ is increasing if and only if the hazard rates of the distributions satisfy

$$\frac{f_{+}(x)}{1-F_{+}(x)} < \frac{g_{+}(x)}{(1-G_{+}(x))},$$

thus for $x < m_2$. The second derivative being negative implies that any stationary point must be a maximum, hence there is at most one such point m'.

Finally, we prove that the mode m' of $\frac{\int_x^\infty (1-F_{+}(\theta))d\theta}{\int_x^\infty (1-G_{+}(\theta))d\theta}$ must be interior. For contradiction suppose that m' is at the boundary. From the first part of the proof, $m' \leq m_2$, so that m' cannot be at the upper end of the support. Thus suppose that $m' = 0$, so that $\frac{\partial}{\partial x}\frac{\int_x^\infty (1-F_{+}(\theta))d\theta}{\int_x^\infty (1-G_{+}(\theta))d\theta} < 0$ for all $x \in [0, \overline{\mathcal{S}}_f]$.

The variance of the distribution over the whole support (positive and negative) can by symmetry ($h_{+} = 2h$) and by integrating by parts twice be written as

$$\int_{-\infty}^{\infty} \theta^2 h(\theta)\, d\theta = \int_0^\infty \theta^2 h_{+}(\theta)\, d\theta = 2\int_0^\infty \theta (1-H_{+}(\theta))\, d\theta = 2\int_0^\infty \int_x^\infty (1-H_{+}(\theta))\, d\theta dx,$$

with $h \in \{f, g\}$, $h_{+} \in \{f_{+}, g_{+}\}$, and $H_{+} \in \{F_{+}, G_{+}\}$.

We can further rewrite and integrate by parts to obtain

$$\begin{aligned}
2\int_0^\infty \int_x^\infty (1-F_{+}(\theta))\, d\theta dx &= 2\int_0^\infty \frac{\int_x^\infty (1-F_{+}(\theta))\, d\theta}{\int_x^\infty (1-G_{+}(\theta))\, d\theta} \int_x^\infty (1-G_{+}(\theta))\, d\theta dx \\
&= -2\frac{\int_z^\infty (1-F_{+}(\theta))\, d\theta}{\int_z^\infty (1-G_{+}(\theta))\, d\theta}\int_z^\infty \int_x^\infty (1-G_{+}(\theta))\, d\theta dx\Bigg|_0^\infty + 2\int_0^\infty \frac{\partial}{\partial z}\frac{\int_z^\infty (1-F_{+}(\theta))\, d\theta}{\int_z^\infty (1-G_{+}(\theta))\, d\theta}\int_z^\infty \int_x^\infty (1-G_{+}(\theta))\, d\theta dx dz \\
&= 2\frac{\int_0^\infty (1-F_{+}(\theta))\, d\theta}{\int_0^\infty (1-G_{+}(\theta))\, d\theta}\int_0^\infty \int_x^\infty (1-G_{+}(\theta))\, d\theta dx + 2\int_0^\infty \frac{\partial}{\partial z}\frac{\int_z^\infty (1-F_{+}(\theta))\, d\theta}{\int_z^\infty (1-G_{+}(\theta))\, d\theta}\int_z^\infty \int_x^\infty (1-G_{+}(\theta))\, d\theta dx dz
\end{aligned}$$

Substituting for $\mu_{h_{+}} = \int_0^\infty (1-H(\theta))\, d\theta$ and $\sigma_h^2 = 2\int_0^\infty \int_x^\infty (1-H_{+}(\theta))\, d\theta dx$, we have that

$$\sigma_f^2 - \frac{\mu_{f_+}}{\mu_{g_+}}\sigma_g^2 = 2\int_0^\infty \frac{\partial}{\partial z}\frac{\int_z^\infty (1-F_+(\theta))\,d\theta}{\int_z^\infty (1-G_+(\theta))\,d\theta}\int_z^\infty\int_x^\infty (1-G_+(\theta))\,d\theta dx dz.$$

We have that $m' = 0$ implies $\frac{\mu_{f+}}{\mu_{g+}} \leq 1$. Moreover, by assumption $\sigma_f^2 = \sigma_g^2$. Hence the left side is non-negative. However, the right side is strictly negative due to our contradictory hypothesis that $\frac{\partial}{\partial z}\frac{\int_z^\infty (1-F_+(\theta))d\theta}{\int_z^\infty (1-G_+(\theta))d\theta} < 0$ for all $z \in [0, \overline{\mathcal{S}}_f]$. □

Proof of Lemma 7. By Lemma A.7, the tail conditional expectation functions, $\mathbb{E}_f\left[\Theta \middle| \Theta \geq x\right]$ and $\mathbb{E}_g\left[\tilde{\Theta} \middle| \tilde{\Theta} \geq x\right]$, cross exactly once in the interior of the positive half of the support. The intersection is at $x = m'$, the mode of the ratio $\frac{\int_x^\infty (1-F_+(\theta))d\theta}{\int_x^\infty (1-G_+(\theta))d\theta}$. Hence, $\mathbb{E}_f\left[\Theta \middle| \Theta \geq t_{n,g}^n(\beta)\right] \geq \mathbb{E}_g\left[\tilde{\Theta} \middle| \tilde{\Theta} \geq t_{n,g}^n(\beta)\right]$ if and only if $t_{n,g}^n(\beta) \leq m'$. By Lemma 6, $t_{n,g}^n(\beta)$ is strictly increasing in β, so by continuity there is a unique $\hat{\beta}$ such that $t_{n,g}^n\left(\hat{\beta}\right) = m'$ and moreover, $t_{n,g}^n(\beta) < m'$ for $\hat{\beta} < \beta$.

By Lemma A.7, the distributions below $t_{n,g}^n(\beta)$ satisfy that $f_+(\theta)/g_+(\theta)$ increasing in θ for all $\theta \leq m$ if $t_{n,g}^n(\beta) \leq m$. By Lemma A.7, $m' < m_2$. By Lemma A.6, $m_2 < m$. Hence, $\beta \leq \hat{\beta}$ implies that $f_+(\theta)/g_+(\theta)$ is increasing for all $\theta \leq t_{n,g}^n(\beta)$. Since the monotone likelihood ratio property is preserved under multiplication of a constant, the truncated distribution below $t_{n,g}^n(\beta)$ satisfies the monotone likelihood ratio property, $\frac{\partial}{\partial\theta}\frac{f_+(\theta)}{F_+\left(t_{n,g}^n(\beta)\right)}/\frac{g_+(\theta)}{G_+\left(t_{n,g}^n(\beta)\right)} > 0$. More generally, the conditional distributions truncated to any interval $\left[t_{i-1,g}^n(\beta), t_{i,g}^n(\beta)\right)$ satisfy $\frac{\partial}{\partial\theta}\frac{f_+(\theta)}{F_+\left(t_{i,g}^n(\beta)\right)-F_+\left(t_{i-1,g}^n(\beta)\right)}/\frac{g_+(\theta)}{G_+\left(t_{i,g}^n(\beta)\right)-G_+\left(t_{i-1,g}^n(\beta)\right)} > 0$ for $i = 1, \ldots, n$. As is well known, the monotone likelihood ratio property implies the standard stochastic order (FOSD) which in turn implies $\mu_f\left(t_{i-1,g}^n, t_{i,g}^n\right) > \mu_g^n\left(t_{i-1,g}^n, t_{i,g}^n\right)$ for $i = 1, \ldots, n$. □

Proof of Lemma 8. We prove the Lemma in three steps. In the first step, we compare the values of the cdfs F_+ and G_+ at the critical types $t_{i,g}^n$, for $i = 1, \ldots, n$. In the second step, we investigate the distribution of induced truncated means on the positive half. In the third step, we show that – on the entire support – the distribution of the truncated means under f is a mean preserving spread of the distribution of the truncated means under g. This implies that the variances of the truncated means are ordered as claimed.

i) For $\beta \leq \hat{\beta}$, by Lemma 7, $t^n_{n,g}(\beta) \leq m'$. By Lemma A.7, $m' < m_2$. Recall, from Lemma A.6 the definition of m_2 as the mode of $\frac{(1-F_+(x))}{(1-G_+(x))}$ and the definition of $\hat{x}$ as the unique point for which $F_+(\theta) = G_+(\theta)$ in the interior of the positive half of the support. Moreover, recall that $G_+(\theta) > F_+(\theta)$ for all $\theta \in (0, \hat{x})$. Since $\frac{(1-F_+(\hat{x}))}{(1-G_+(\hat{x}))} = 1$ and $\frac{(1-F_+(\theta))}{(1-G_+(\theta))} > 1$ for $\theta < \hat{x}$, it follows that $m_2 < \hat{x}$ and by implication, that $G_+(m') > F_+(m')$. Thus $G_+\left(t^n_{i,g}(\beta)\right) > F_+\left(t^n_{i,g}(\beta)\right)$ for any $i = 1, \ldots, n$, and $\beta \leq \hat{\beta}$.

ii) Let Y_{f_+} and Y_{g_+} denote random variables on $\mathbb{R}_+$ with probability distributions

$$\Pr\left(Y_{f_+} \leq z\right) = \begin{cases} F_+\left(t^n_{i-1,g}(\beta)\right) & for\ z \in \left[t^n_{i-1,g}, \mu_f\left(t^n_{i-1,g}, t^n_{i,g}\right)\right) \\ F_+\left(t^n_{i,g}(\beta)\right) & for\ z \in \left[\mu_f\left(t^n_{i-1,g}, t^n_{i,g}\right), t^n_{i,g}(\beta)\right) \end{cases}$$

and likewise

$$\Pr\left(Y_{g_+} \leq z\right) = \begin{cases} G_+\left(t^n_{i-1,g}(\beta)\right) & for\ z \in \left[t^n_{i-1,g}, \mu_g\left(t^n_{i-1,g}, t^n_{i,g}\right)\right) \\ G_+\left(t^n_{i,g}(\beta)\right) & for\ z \in \left[\mu_g\left(t^n_{i-1,g}, t^n_{i,g}\right), t^n_{i,g}(\beta)\right), \end{cases}$$

for $i = 1, \ldots, n+1$, where by convention $t^n_{0,g} = 0$ and $t^n_{n+1,g} = \infty$. The functions are right continuous and take upward jumps at the induced means over the given intervals; the size of each upward jump corresponds to the probability mass over the interval. Both the mean and the probability mass depend on the distributions, f_+ or g_+. By Part i), if the jumps occurred at the same points, $\mu_f\left(t^n_{i-1,g}, t^n_{i,g}\right) = \mu_g\left(t^n_{i-1,g}, t^n_{i,g}\right)$ for all i, then the distributions would be ranked by FOSD. It suffices to note that $\mu_g\left(t^n_{i-1,g}, t^n_{i,g}\right) < \mu_f\left(t^n_{i-1,g}, t^n_{i,g}\right)$ implies that the upward jumps of $\Pr\left(Y_{g_+} \leq z\right)$ are to the left of the upward jumps of $\Pr\left(Y_{f_+} \leq z\right)$. It follows that

$$\Pr\left(Y_{g_+} \leq z\right) \geq \Pr\left(Y_{f_+} \leq z\right) \quad \text{for all } z \geq 0.$$

The inequality is strict for $z \in \left[\mu_g\left(0, t^n_{1,g}\right), \mathbb{E}_f\left[\Theta | \Theta \geq t^n_{n,g}\right]\right)$. For $z \in \left[0, \mu_g\left(0, t^n_{1,g}\right)\right)$ both functions take value zero, for $z \geq \mathbb{E}_f\left[\Theta | \Theta \geq t^n_{n,g}\right]$ both functions take value one.

iii) Consider now the distribution of induced means over the entire support. By the law of iterated expectations, the expected values of the conditional means are equal to the prior mean, zero, for both distributions. Together with symmetry and FOSD on each half, we obtain that the distribution of induced means under f are a mean preserving spread of the distribution of induced means under distribution g.

$\square$

Proof of Lemma 9. By Lemma 7, for $\beta \leq \hat{\beta}$

$$\begin{aligned}&\mathbb{E}_f\left[\Theta \middle| \Theta \in \left[t^n_{i-1,g}, t^n_{i,g}\right]\right] + \mathbb{E}_f\left[\Theta \middle| \Theta \in \left[t^n_{i,g}, t^n_{i+1,g}\right]\right] \\ >&\mathbb{E}_g\left[\tilde{\Theta} \middle| \tilde{\Theta} \in \left[t^n_{i-1,g}, t^n_{i,g}\right]\right] + \mathbb{E}_g\left[\tilde{\Theta} \middle| \tilde{\Theta} \in \left[t^n_{i,g}, t^n_{i+1,g}\right]\right],\end{aligned} \tag{19}$$

where we take $t^n_{n+1,g} = \overline{\mathcal{S}}_g$. We now show that for all i, condition (19) implies that under distribution f the equilibrium critical types under distribution f are strictly higher and strictly better for the sender than the equilibrium critical types under distribution g.

Recall that the expected utilities under communication of the sender and the receiver are

$$\mathbb{E}u_S(\beta\mu, \Theta) = \beta(2-\beta)\,\mathbb{E}\left[(\mu^n)^2\right] - \sigma^2 \quad \text{and} \quad \mathbb{E}u_R(\beta\mu, \Theta, \beta) = \beta^2\left(\mathbb{E}\left[(\mu^n)^2\right] - \sigma^2\right).$$

Given our quadratic loss assumption, expected utility is higher if and only if the variance of the induced means is higher.

The sender's expected utility for distribution f_+ and arbitrary thresholds $\boldsymbol{t} = (t_i)^n_{i=0}$ with $t_0 = 0$ is

$$\mathbb{E}u_{S,f}(\mathbf{t}) := -\sum_{i=1}^{n} \int_{t_{i-1}}^{t_i} \left(\theta - \beta\mu_{i,f,\boldsymbol{t}}\right)^2 f_+(\theta)\, d\theta - \int_{t_n}^{\overline{\mathcal{S}}_f} \left(\theta - \beta\mu_{n+1,f,\boldsymbol{t}}\right)^2 f_+(\theta)\, d\theta. \tag{20}$$

For any fixed $t^n_{i-1,g}$, we denote by $t^{(*)}_{j,f}$ for all $j \geq i$ the "partial" equilibrium thresholds under f, where the distribution is adjusted from g to f on the entire support but the equilibrium thresholds are adjusted only above $t^n_{i-1,g}$ but not below, thus $t_j = t^n_{j,g}$ for $j < i$ and $t_j = t^{(*)}_{j,f}$ for $j \geq i$.

Consider the following iterative procedure. At iteration one, keep all thresholds $t_i = t^n_{i,g}$ for $i = 1, \ldots, n-1$ fixed at the equilibrium values under g and let t_n adjust to $t^{(*)}_{n,f} = t^{(*)}_{n,f}\left(t^n_{n-1,g}\right)$. At $t^{(*)}_{n,f}$, the sender is indifferent under f between pooling upwards or downwards given that the receiver best replies under f to the truncation above $t^n_{n-1,g}$. At iteration j, keep all thresholds $t_i = t^n_{i,g}$ for $i = 1, \ldots, n-j$ fixed, adjust threshold t_{n-j+1} to make the sender indifferent at $t^{(*)}_{n-j+1,f} = t^{(*)}_{n-j+1,f}\left(t^n_{n-j,g}\right)$, and keep the sender indifferent at all thresholds $t^{(*)}_{l,f}$ for $l \geq n-j+2$. Note that all $t^{(*)}_{l,f}$ depend recursively on the initial value $t^n_{n-j,g}$ and on their respective predecessors $t^{(*)}_{n-j+1,f}, \ldots, t^{(*)}_{l-1,f}$.

For iteration step one, differentiate the sender's utility (20) with respect to t_n

$$\begin{aligned}\frac{\partial}{\partial t_n}\mathbb{E}u_{S,f}(\mathbf{t}) = &-\left(t_n - \beta\mu_{n,f}\left(t^n_{n-1,g}, t_n\right)\right)^2 f_+(t_n) + \left(t_n - \beta\mu_{n+1,f}\left(t_n, \overline{\mathcal{S}}_f\right)\right)^2 f_+(t_n) \\ &+ 2\beta\frac{\partial\mu_{n,f}\left(t^n_{n-1,g}, t_n\right)}{\partial t_n}\int_{t^n_{n-1,g}}^{t_n}\left(\theta - \beta\mu_{n,f}\left(t^n_{n-1,g}, t_n\right)\right) f_+(\theta)\, d\theta \\ &+ 2\beta\frac{\partial\mu_{n+1,f}\left(t_n, \overline{\mathcal{S}}_f\right)}{\partial t_n}\int_{t_n}^{\overline{\mathcal{S}}_f}\left(\theta - \beta\mu_{n+1,f}\left(t_n, \overline{\mathcal{S}}_f\right)\right) f_+(\theta)\, d\theta. \qquad (21)\end{aligned}$$

We note that the integral in the second line can equivalently be written as $\left(F_+(t_n) - F_+\left(t^n_{n-1,g}\right)\right)(1-\beta)\,\mu_{n,f}\left(t^n_{n-1,g}, t_n\right)$ which is strictly positive for $t_n \geq t^n_{n,g}$. Likewise the integral in the third line is equivalent to $(1 - F_+(t_n))(1-\beta)\,\mu_{n+1,f}\left(t_n, \overline{\mathcal{S}}_f\right)$ and strictly positive. Since conditional means are increasing in the thresholds, we can conclude that the second and third lines are positive. The first line is positive if

$$t_n - \beta\mu_{n,f}\left(t^n_{n-1,g}, t_n\right) < \beta\mu_{n+1,f}\left(t_n, \overline{\mathcal{S}}_f\right) - t_n.$$

For $t_n = t^n_{n,g}$, the inequality holds by condition (19) and the equilibrium condition $t^n_{n,g} - \beta\mu^n_{n,g} = \beta\mu^n_{n+1,g} - t^n_{n,g}$. Moreover, by logconcavity of the density, $\mu_{n+1,f}\left(t_n, \overline{\mathcal{S}}_f\right)$ and $\mu_{n,f}\left(t^n_{n-1,g}, t_n\right)$ each increase in t_n less than one for one. Hence, there exists a unique $t^{(*)}_{n,f} > t^n_{n,g}$ such that

$$t^{(*)}_{n,f} - \beta\mu_{n,f}(t^n_{n-1,g}, t^{(*)}_{n,f}) = \beta\mu_{n+1,f}(t^{(*)}_{n,f}, \overline{\mathcal{S}}_f) - t^{(*)}_{n,f}. \qquad (22)$$

It follows that $\frac{\partial}{\partial t_n}\mathbb{E}u_{S,f}(\mathbf{t}) > 0$ for all $t_n \in \left[t^n_{n,g}, t^{(*)}_{n,f}\right]$.

Consider an arbitrary iteration step $l < n$. Suppose that all thresholds have been adjusted including down to $t^{(*)}_{l+1,f}$. Differentiating the sender's expected payoff with respect to t_l and readjusting the thresholds above accordingly, we find that

$$\begin{aligned}
&\frac{\partial}{\partial t_l}\mathbb{E}u_{S,f}(\mathbf{t})\big|_{t_l=t^n_{l,g}} \\
&\geq -\left(t_{l,g}-\beta\mu_{l,f}\left(t^n_{l-1,g},t_{l,g}\right)\right)^2 f_+(t_{l,g}) + \left(t_{l,g}-\beta\mu_{l+1,f}(t_{l,g},t^{(*)}_{l+1,f})\right)^2 f_+(t_{l,g}) \\
&>0.
\end{aligned}$$

The first inequality follows from two insights. First, note that when differentiating the sender's utility the effects through the boundaries of the integrals $t^{(*)}_{l+j,f}$ for $j > 0$ are zero by an envelope argument: a typical derivative is given by

$$\left(-\left(t^{(*)}_{l+j,f}-\beta\mu_{l+j,f}(t^{(*)}_{l+j-1,f},t^{(*)}_{l+j,f})\right)^2+\left(t^{(*)}_{l+j,f}-\beta\mu_{l+j+1,f}(t^{(*)}_{l+j,f},t^{(*)}_{l+j+1,f})\right)^2\right)f_+(t^{(*)}_{l+j,f})\frac{dt^{(*)}_{l+j,f}}{dt_{l+j-1}}\cdots\frac{dt^{(*)}_{l+1,f}}{dt_l}=0.$$

Second, as we have seen in (21), the effects through changes of the thresholds on the means are strictly positive, because the means are increasing in the truncation points and $t^{(*)}_{l+j,f}$ is increasing in t_{l+j-1}. The second (strict) inequality is implied by the equilibrium condition for $t^n_{l,g}$ under g, condition (19), and $t^{(*)}_{l+1,f} > t^n_{l+1,g}$.

It remains to be shown that there is a unique $t_l = t^{(*)}_{l,f}$ such that

$$\left(\beta\mu_{l+1,f}(t_l,t^{(*)}_{l+1,f})-t_l\right)-\left(t_l-\beta\mu_{l,f}\left(t^n_{l-1,g},t_l\right)\right)=0, \tag{23}$$

and moreover, that $\frac{\partial}{\partial t_l}\mathbb{E}u_{S,f}(\mathbf{t}) > 0$ for all $t_l \in \left[t^n_{l,g}, t^{(*)}_{l,f}\right]$. Differentiating the left side of (23) with respect to t_l, we get

$$-2+\beta\frac{\partial}{\partial t_l}\mu_{l,f}\left(t^n_{l-1,g},t_l\right)+\beta\frac{\partial}{\partial t_l}\mu_{l+1,f}(t_l,t^{(*)}_{l+1,f})+\beta\frac{\partial}{\partial t_{l+1}}\mu_{l+1,f}(t_l,t^{(*)}_{l+1,f})\frac{dt^{(*)}_{l+1,f}}{dt_l}.$$

By logconcavity, $\frac{dt^{(*)}_{l+1}}{dt_l} \leq 1$ implies that this expression is negative. We show that $\frac{dt^{(*)}_{l+1}}{dt_l} \leq 1$ holds by induction: Totally differentiating (22) with respect to to $t^{(*)}_n$ and t_{n-1}, we find that

$$\frac{dt^{(*)}_n}{dt_{n-1}}=\frac{\beta\frac{\partial}{\partial t_{n-1}}\mu_{n,f}\left(t_{n-1},t^{(*)}_{n,f}\right)}{2-\beta\frac{\partial}{\partial t^{(*)}_n}\mu_{n,f}\left(t_{n-1},t^{(*)}_{n,f}\right)-\beta\frac{\partial}{\partial t^{(*)}_n}\mu_{n+1,f}\left(t^{(*)}_{n,f},\overline{\mathcal{S}}_f\right)}\leq 1,$$

where the inequality is due to logconcavity of the density.

Next, suppose that $\frac{dt_{l+1}^{(*)}}{dt_l} \leq 1$. Totally differentiating (23) we get

$$\frac{dt_{l,f}^{(*)}}{dt_{l-1}} = \frac{\beta \frac{\partial}{\partial t_{l-1}} \mu_{l,f}(t_{l-1}, t_{l,f}^{(*)})}{2 - \beta \frac{\partial}{\partial t_{l,f}^{(*)}} \mu_{l,f}(t_{l-1}, t_{l,f}^{(*)}) - \beta \frac{\partial}{\partial t_{l,f}^{(*)}} \mu_{l+1,f}(t_{l,f}^{(*)}, t_{l+1,f}^{(*)}) - \beta \frac{\partial}{\partial t_{l+1,f}^{(*)}} \mu_{l+1,f}(t_{l,f}^{(*)}, t_{l+1,f}^{(*)}) \frac{dt_{l+1,f}^{(*)}}{dt_{l,f}^{(*)}}} \leq 1$$

by logconcavity of the density and the assumption that $\frac{dt_{l+1}^{(*)}}{dt_l} \leq 1$. This concludes the argument. □

Proof of Proposition 6. The value of communication for the generalized Pareto case is derived by dynamic programming. The expression $\frac{1}{1-\delta}$ in the denominator of (4) is the slope of the tail conditional expectation. It can be shown that a lower bound on the value of communication is obtained if we use the minimal slope of the tail conditional expectation. It is well known that the Normal distribution features a convex tail conditional expectation (see Sampford (1953)), so the minimal slope obtains at $\theta = 0$:

$$\frac{\partial}{\partial z} \mathbb{E}\left[\Theta \middle| \Theta \geq z\right]\Big|_{z=0} = \left(\mathbb{E}\left[\Theta \middle| \Theta \geq z\right] - z\right) \frac{f(z)}{1 - F(z)}\Big|_{z=0} = \frac{\mu_+}{\sigma} 2 \frac{1}{\sqrt{2\pi}}.$$

Moreover, for the Gaussian distribution, we have

$$\mathbb{E}\left[\Theta \middle| \Theta \geq z\right]\big|_{z=0} = \mu_+ = \sigma \left. \frac{\phi(z)}{1 - \Phi(z)} \right|_{z=0} = \sigma \frac{\sqrt{2}}{\sqrt{\pi}}.$$

Substituting in (4) for μ_+ and the minimal slope, we obtain

$$var(\mu^\infty) \geq \frac{\frac{4}{\pi}}{2 - \beta \frac{2}{\pi}} \sigma^2.$$

We can now prove the statements:

i) If the state is Gaussian, we obtain that communication is preferred over delegation (using the lower bound) for

$$\beta^2 \left(\frac{\frac{4}{\pi}}{2 - \beta \frac{2}{\pi}} \sigma^2 - \sigma^2 \right) \geq -\left(1 - \beta\right)^2 \sigma^2,$$

which holds for $\beta \lesssim 0.702$.

Comparing the values of communicating under a Gaussian and a Laplace distribution, we find that the Gaussian induces a higher value of communication than the Laplace

$$\beta^2 \left(\frac{\frac{4}{\pi}}{2 - \beta \frac{2}{\pi}} \sigma^2 - \sigma^2 \right) \geq \beta^2 \left(\frac{1}{2 - \beta} \sigma^2 - \sigma^2 \right),$$

for $\beta \lesssim 0.858$.

ii) Recall that for the Laplace distribution, delegation is preferred to communication for $-(1-\beta)^2 \sigma^2 \geq \beta^2 \left(\frac{1}{2-\beta} \sigma^2 - \sigma^2 \right)$, i.e., for $\beta \geq \frac{2}{3}$.

Hence, for $\beta \in \left(\frac{2}{3}, 0.702 \right)$ delegation is strictly optimal if the state follows a Laplace distribution while communication is strictly optimal if the state follows a Gaussian distribution. □

References

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.

Antić, N. and Persico, N. (2020). Cheap talk with endogenous conflict of interest, *Econometrica* **88**(6): 2663–2695.

Artzner, P., Delbaen, F., Eber, J.-M. and Heath, D. (1999). Coherent measures of risk, *Mathematical finance* **9**(3): 203–228.

Cambanis, S., Huang, S. and Simons, G. (1981). On the theory of elliptically contoured distributions, *Journal of Multivariate Analysis* **11**(3): 368–385.

Chen, Y. and Gordon, S. (2015). Information transmission in nested sender–receiver games, *Economic Theory* **58**(3): 543–569.

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

Deimen, I. and Szalay, D. (2019). Delegated expertise, authority, and communication, *American Economic Review* **109**(4): 1349–74.

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

Dessein, W., Lo, D. and Minami, C. (forthcoming). Coordination and organization design: Theory and micro-evidence., *American Economic Journal: Microeconomics* .

Gómez, E., Gómez-Villegas, M. A. and Marín, J. M. (2003). A survey on continuous elliptical vector distributions, *Revista matemática complutense* **16**(1): 345–361.

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

Liu, S. and Migrow, D. (2019). Designing organizations in volatile markets. Working paper.

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

Metzger, C. and Rüschendorf, L. (1991). Conditional variability ordering of distributions, *Annals of Operations Research* **32**(1): 127–140.

National Comission on the BP Deepwater Horizon oil spill and offshore drilling (2011). Deep water. the gulf oil disaster and the future of offshore drilling.

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

Rantakari, H. (2013). Organizational design and environmental volatility, *The Journal of Law, Economics, & Organization* **29**(3): 569–607.

Sampford, M. R. (1953). Some inequalities on mill's ratio and related functions, *The Annals of Mathematical Statistics* **24**(1): 130–132.

Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic orders*, Springer.

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

Wellner, J. A. (2013). Strong log-concavity is preserved by convolution, *High Dimensional Probability VI*, Springer, pp. 95–102.

Whitt, W. (1985). Uniform conditional variability ordering of probability distributions, *Journal of Applied Probability* **22**(3): 619–633.