Toxic Types and Infectious Communication Breakdown*

Kfir Eliaz and Alexander Frug†

May 5, 2023

Abstract

We study an environment where an informed sender has conflicting interests with an uninformed receiver only in some states. Using an "infection-like" argument, we show that with symmetric loss functions, the presence of such disagreement states - even if they are very rare - leads to coarse communication in all states, even those where, following communication, it is commonly known that the players' interests are perfectly aligned. However, with asymmetric loss functions, one can construct examples with truthful communication in some set of states.

Keywords: Cheap talk, contagion.
JEL Classification: D83.

*We thank Navin Kartik, Elliot Lipnowski, and Joel Sobel for helpful comments. Eliaz acknowledges financial support from the Sapir Center for Economic Development and from ISF grant 374/16. Frug acknowledges the financial support of the Spanish Ministry of Science and Innovation through the grants: AEI/FEDER, UE - PGC2018-098949-B-I00, SEV-2015-0563, and CEX2019-000915-S.

†Eliaz: School of Economics, Tel Aviv University and David Eccles School of Business, University of Utah. E-mail: kfire@tauex.tau.ac.il. Frug: Department of Economics and Business, Universitat Pompeu Fabra and Barcelona GSE. E-mail: alexander.frug@upf.edu.

1 Introduction

Traditionally, models of strategic information transmission have focused on environments where, in each and every state, there is some probability (typically one) that the informed party disagrees with the uninformed party on what is the optimal action. As is well known, such form of interest divergence leads to inefficient coarse communication in equilibrium.[1] However, oftentimes, the two parties do not always have conflicting interests – they disagree only in some states, but agree in all others. For instance, an employer would agree with an employee on what is the right position for him when the employee is indeed qualified for that position; elected officials would tend to agree on the action proposed by some lobbying group if the case made by the group is indeed correct; a judge would most likely agree on the sentence recommendation of a prosecutor if all the arguments made by the latter were indeed true. In such environments, will the mere presence of disagreement states contaminate the communication in other states? We show that the answer is yes: in a broad set of environments, even an arbitrarily small set of disagreement states restricts the players' ability to communicate about states where it is commonly known that they agree on the optimal action.

We establish our result in a setting where the receiver may opt-out and not interact with the sender but, conditional on interacting, the players' preferences are aligned. While the sender wants to interact with the receiver in all states, there is a small set of states in which the receiver prefers to opt-out. Absent these states, there would be a fully revealing equilibrium. However, the possibility of such states—the sender's "toxic types"—contaminates the ability of *all* sender types to communicate with the receiver.

Some examples that fit this setting include mentoring, where a senior colleague, a supervisor, or an advisor wishes to teach some material to a junior colleague, an advisee, or a research assistant. Both the mentor and the apprentice would benefit from the latter's understanding of the material. Yet, for the interaction to be effective, the mentor must learn what the apprentice already knows, or what his skill level is, so that he can give

[1] For a comprehensive survey of this literature, see Sobel (2013).

the appropriate explanation or guidance. Oftentimes, the mentor prefers not to take on an apprentice with low skills, which would require excessive guidance or training. Consequently, a candidate for apprenticeship who has low skills would not want to disclose that information even at the expense of getting suboptimal training that is not suited to his level. Sometimes, we may be able to directly verify the candidate's skills with a test or a task, but oftentimes that is not feasible, or it is too costly, and one must rely on the candidate's word.

Likewise, when a manager needs to assign a person to a task, he may prefer not to work with a person whose skill or experience is too low. Similarly, a potential investor is interested only in viable projects and qualified entrepreneurs. When the manager knows the worker is skilled or experienced, both sides share the same objective of succeeding in the task. Similarly, when an investor knows the entrepreneur he invested in is qualified, a conflict of interest need not arise.

Our model consists of a continuum of sender types (represented by an interval) and symmetric loss functions, such that conditional on interacting with the sender, both players want the receiver to match the sender's type. In analyzing this model, we focus on "interval equilibria" or equilibria that induce a partition of the sender type space into (possibly degenerate) intervals. In these equilibria, messages can be naturally interpreted as possibly coarse statements of the sender's characteristics (e.g., the sender's skill level or the task level most suited for him). Our main result establishes that in every equilibrium in this class, the sender chooses only *finitely* many actions. While the number of actions that are chosen in equilibrium can increase as the likelihood of disagreement states decreases, their total number remains finite. Still, as the set of toxic types goes to zero, the precision of communication goes to infinity. Specifically, for any $\Delta > 0$, there exists some ε, such that the length of every interval of sender types is below Δ if the lowest interval (the one containing the toxic types) is below ε.

Our result has the following broader implication: if for some reason there is an interval of sender types who must pool, then, in equilibrium, communication will be coarse over the entire type space. In this paper, we focus on an incentive to pool that stems from the concern that the receiver will choose to opt out, but in general, there may be other reasons for pooling.

The channel through which communication breaks down in our model is significantly different from the standard channel introduced by Crawford and Sobel (1982) that operates through a bias in the sender's preferences.[2] Consequently, proving that the receiver's equilibrium partition has only finitely many intervals requires a different approach than that used in the cheap-talk model of Crawford and Sobel (1982) and many of its extensions.

Our proof technique can be summarized as follows. Recall that the sender always prefers interacting with the receiver but the receiver prefers not to interact with sufficiently low types of the sender. Consequently, the lowest sender type will not reveal his type in any equilibrium where there is positive probability that the receiver does not opt out from the interaction. Consequently, in any interval equilibrium, the lowest-type sender can only communicate to the receiver that his type belongs to a non-degenerate interval of all types below some positive cutoff. The sender type at this cutoff must then be indifferent between pooling with the types below him and pooling with an interval of types above him. In general, the interval above the cutoff type can be longer or shorter than its left neighboring interval. For the infection argument to work throughout the whole state space, the interval above the cutoff type, as well as the next intervals that arise from the boundary types' indifference condition, cannot decrease in lengths sufficiently fast. Proving that this is indeed the case is the crux of our proof since we allow for a general distribution of types (in particular, the density function need not be decreasing in types).

Thus, the heart of our proof is showing that Lipschiz continuity of the density function and the symmetry in payoff functions guarantee that the lengths of adjacent intervals cannot shrink too fast. Indeed, symmetry of the player's loss function plays a key role in our analysis: with an asymmetric loss function we show an example with full communication by

[2] A recent paper by Dilmé (2019) also studies a model in which the sender and receiver agree on the optimal action. In his model, however, communication is imprecise only when the sender does not perfectly observe the state.

sender types above some threshold, and an example where the size of the rightmost interval of sender types is above half, no matter how small the lowest interval (which contains all the toxic types) is.

To illustrate our main result and the working of our model, we analyze a "canonical" case in which types are uniformly distributed on [0,1]. It is shown that, under a mild restriction on the players' payoffs, there exists a unique Pareto efficient (interval) equilibrium in which (i) the receiver interacts with *all* sender types, and (ii) the sender types are pooled into *equal-length* intervals. To further illustrate the adverse effect of the toxic types, we also show an example where all non-babbling equilibria have the property that the receiver mixes between interacting with the sender and ending the game such that the probability of ending the game is *higher* for more desirable types.

The remainder of the paper is organized as follows. Section 2 introduces the general model. Our main results are described and proven in Section 3. Section 4 analyzes a canonical case with uniformly distributed types. Section 5 reviews the related literature and Section 6 concludes.

2 Model

A sender privately draws a type θ from a distribution $F[0,1]$ with a strictly positive, Lipschitz continuous, and differentiable density f.The sender sends a message $m \in [0,1]$ to a receiver, after which the receiver chooses an action $a \in [0,1] \cup \{N\}$, where N is interpreted as a decision not to interact with the sender. We refer to N as the *non-participation* action and to the actions in [0,1] as the *participation* actions. Both players' payoffs depend on the action taken and on the sender's private information. The receiver's payoff is defined as follows:

$$u^R(a,\theta) = \begin{cases} R(\theta) - r(\theta)L(|a-\theta|) & , \quad a \in [0,1] \\ \rho & , \quad a = N, \end{cases}$$

where $R(\cdot)$ is non-negative, continuous, and weakly increasing; $r(\cdot)$ is positive, Lipschitz continuous and differentiable; and $L(\cdot)$ is increasing, strictly convex, differentiable, and satisfies $L(0) = 0$. The sender's payoff is similarly defined as follows:

$$u^S(a, \theta) = \begin{cases} S(\theta) - s(\theta)\Lambda(|a - \theta|) & , \quad a \in [0, 1] \\ \sigma & , \quad a = N, \end{cases}$$

with the analogous assumptions made on $S(\cdot)$, $s(\cdot)$, and $\Lambda(\cdot)$.[3] Note that these payoffs have the following property: conditional on $a \neq N$, both the sender and receiver agree on the optimal action, regardless of the sender's type. We assume that σ satisfies that each sender type strictly prefers any $a \in [0, 1]$ to $a = N$. In addition, we assume that $R(0) < \rho < R(1)$. That is, the receiver prefers $a = N$ whenever $R(\theta) < \rho$, even if he is perfectly informed about the sender's type.[4] We refer to types θ for which $R(\theta) < \rho$ as *toxic* and our objective is to study the effect of toxic types on communication with other types with which the players' interests are perfectly aligned.

We interpret the above game as capturing a situation in which an employer needs to decide whether to hire an agent, and if so, what task to assign him. The agent's success depends both on his ability θ, and on how closely the assigned task matches this ability. Both the employer and the agent benefit from success, and have aligned interests in the sense that, conditional on employing the agent, they agree on the best-fitting task. The only conflict of interest arises from the presence of toxic types: every agent type prefers to be employed and work on any task rather than be unemployed, while the employer prefers not to hire an agent if his ability is too low ($R(\theta) < \rho$), and has no conflict of interest with other types.

In our analysis, we focus on a class of perfect Bayesian Nash equilibria (PBE) in which the sender's strategy induces an information partition for the receiver where each element,

[3] An equivalent formulation is to assume that the payoff to both players is zero when $a = N$, but the receiver pays some cost to choose $a \in [0, 1]$.

[4] Note that this implies that any equilibrium satisfies the NITS condition of Chen, Kartik, and Sobel (2008).

which we refer to as an *information set* of the receiver, is a (possibly degenerate) interval. For convenience we assume that each interval is communicated by exactly one message. We refer to such equilibria as *interval equilibria*. Given the receiver's preferences, if the receiver believes that the state is an element of an interval, then his optimal action is an element of that set. Thus, in particular, under any interval equilibrium, the mapping from states to actions is monotonically increasing.

Due to the generality of our specification, in addition to interval equilibria, there may exist equilibria in which the mapping between types and participation actions is non-monotonic. However, this non-monotonicity in the type-to-action mapping can only occur on information sets where N is optimal for the receiver, i.e., information sets that do not create value for the receiver.[5] In Appendix A1, we illustrate a non-monotonic equilibrium and show that when $S(\theta)$ and $s(\theta)$ are constants, only interval equilibria exist. Henceforth, we will refer to an interval equilibrium simply as an "equilibrium."

3 Equilibrium communication

This section establishes our two main results: (i) even an arbitrarily small fraction of undesirable types prevents *all* sender types—including those whose preferences are aligned with the receiver's—from communicating mutually beneficial information, and (ii) when the size of the interval containing the toxic types goes to zero, communication gets arbitrarily close to being fully revealing.

We begin by establishing a number of preliminary observations and explaining in the main text why each of them holds. To do this, it will be useful to introduce the following notation. For any $0 \leq x \leq y \leq 1$, denote by $V(x, y)$ the receiver's expected payoff from

[5] A variant of the standard sorting argument can be used to show that on all other information sets, types are mapped into actions monotonically.

the optimal participation action $a \neq N$, given the belief that $\theta \in [x, y]$, i.e.,

$$V(x,y) = \begin{cases} \max_{a \neq N} \int_x^y \left[R(\theta) - r(\theta)L(|\theta - a|)\right] \frac{f(\theta)}{F(y)-F(x)} & , \quad x < y \\ R(x) & , \quad x = y. \end{cases}$$

If $V(0, y) < \rho$ for all $y \geq 0$, then in equilibrium the receiver chooses N with certainty and Theorem 1 holds trivially. A necessary condition for informative communication in equilibrium is $V(0, y) \geq \rho$, for some y. We maintain this assumption throughout the rest of the paper

Observation 1. *Fix an equilibrium. There exists an equilibrium message after which N is chosen with probability 1 if and only if N is played with certainty regardless of the sender's message.*

This observation follows from our assumption that every sender type prefers some action in [0,1] to the action *N*. Thus, if the receiver chooses with positive probability a participation action following some message that is sent in equilibrium, then *every* sender type would induce a participation action with positive probability in that equilibrium. This leads to the following observation.

Observation 2. *In any equilibrium, there is a threshold type $0 < \theta^* \leq 1$ such that all types in $[0, \theta^*]$ send the same message.*

This observation essentially plants the seed that leads to our main result: if for some reason an interval of sender types must pool, this creates a ripple effect that coarsens the communication with all sender types. To create this effect, the interval of pooling types can reside *anywhere* in the type space. While in our model the incentive to pool arises because the receiver prefers not to interact with sufficiently low types, more generally, some sender types may want to pool because they disagree with the receiver over the optimal action. For example, think of situations where the most profitable (from an employer's point of view) tasks to assign workers with sufficiently high ability are tasks that workers would like to avoid (e.g., suppose high-ability workers are considered more motivated and trustworthy and hence are assigned more administrative duties). Here, "toxic types" are located at the top of the type space and they prefer to pool with slightly lower types.

Observation 2 follows from Observation 1 because either all types induce N with certainty (in which case the interval is $[0, 1]$) or a participation action is a best response for the receiver (and so the length of the interval is positive since $V(0, 0) = R(0) < \rho$).[6] Taken together, our assumptions on $L(\cdot)$, the continuity of $f(\cdot)$ and $r(\cdot)$, and Observation 2 imply that the optimal *participation* action at a non-degenerate interval $[x, y]$ is strictly below y.[7] The indifference condition of the sender's boundary type between two adjacent intervals then implies the following.

Observation 3. *Suppose that $[x, y]$, where $y < 1$, is a receiver's information set in an equilibrium in which participation actions are played. Then the receiver's information partition in that equilibrium contains a* non-degenerate *interval $[y, z]$.*

We are now ready to state our first main result.

Theorem 1. *In any equilibrium, the support of the receiver's strategy consists of only finitely many actions.*

Clearly, if the intervals do not decrease in length when the state increases (as is the case in the example of uniform types that we analyze in the next section), then it is easy to see why Theorem 1 holds.[8] The challenge in proving Theorem 1 is that, in general, the intervals may *decrease* in length as the state increases. The main step in the proof is, therefore, to show that the intervals do not shrink too fast.

[6] We ignore equilibria in which the sender transmits *redundant* information since for any such equilibrium, there exists an equilibrium that does not contain redundant information transmission. For example, we identify an equilibrium in which the sender reveals whether his type is 0 or not and the receiver chooses N with certainty, with the completely uninformative equilibrium where the receiver has only one information set.

[7] We denote any interval information set whose infimum and supremum are, respectively, x and y by $[x, y]$.

[8] One condition that guarantees this is that the density function is non-increasing. An analogous condition is used in Di Pei (2017), who studies the effect of higher-order uncertainty on the sender's bias (in his environment, the contagion argument goes from high to low types, hence the condition he requires is that the density is non-decreasing).

The proof of Theorem 1 appears in Appendix A3. It relies on three key Lemmas (which also appear in Appendix A3) that are also useful for proving Theorem 2 below. To briefly describe the Lemmas, suppose that the receiver believes that $\theta \in I$ for some interval $I \subseteq [0,1]$. Lemma 3 establishes a bound on the distance between the mid-point of I and the receiver's optimal action given $\theta \in I$, that is linear in the squared length of I. Using this bound, Lemma 4 establishes a lower bound on the ratio between the length of I' - an adjacent interval to I, and the length of I. Specifically, it shows that there exists a constant $\psi > 0$, such that $\frac{|I'|}{|I|} > 1 - \psi|I|$. Intuitively, this shows that as the length of I shrinks, the intervals I and I' become arbitrarily similar in lengths. Finally, Lemma 5 shows that the sum of any sequence of positive real numbers that decreases at a pace that satisfies the bound of Lemma 4, diverges. Given these Lemmas, Theorem 1 follows by observing that the lengths of intervals in any equilibrium partition (from left to right) are bounded from below by a divergent sequence.

The above result relies on our assumption that the players' payoff functions are symmetric. With asymmetric payoff functions there would still be some infection, and desirable sender types would suffer from reduced ability to communicate their information to the receiver, but this infection might vanish and leave sufficiently high sender types unaffected.

Example 1. *Suppose that when $a \neq N$, the sender's payoff is $u^S(a,\theta) = -(\theta - a)^2$, while the receiver's payoff is,*

$$u^R(a,\theta) = \begin{cases} \theta - 4(\theta - a)^2 & , \quad \theta > a \\ \theta - (\theta - a)^2 & , \quad \theta \leq a. \end{cases}$$

Also, assume that the sender's and receiver's payoffs from the receiver's action $a = N$ are $\sigma < -1$ and $\rho > 0$, respectively. Note that, conditional on $a \neq N$, the players' interests are aligned in the sense that both would choose $a = \theta$ in every state θ. However, since $\rho > 0$, in any equilibrium, sufficiently low types must be pooled together.

Assume that the state is uniformly distributed on $[0,1]$. Suppose that when the receiver's information set is $[x,y]$, he does not choose $a = N$. In this case, it is easy to *verify that he selects the action* $\frac{1}{3}x + \frac{2}{3}y$. *If* $y < 1$, *from the indifference of the boundary sender type* $\theta = y$, *we can conclude that the length of the interval right-adjacent to* $[x, y]$ *is*[9] $\frac{1}{2}(y - x)$.

For example, if $\rho = \frac{25}{216}$, *there is an equilibrium where the receiver's information partition consists of infinitely many non-degenerate intervals whose union is* $[0, \frac{1}{2}]$ *and singletons above* $\frac{1}{2}$. *The leftmost interval in this partition is* $[0, \frac{1}{4}]$ *and the length of every other interval in* $[0, \frac{1}{2}]$ *is half the length of its left-adjacent neighbor. In this case, the adverse effect that arises from the existence of undesirable types (i.e., sender types below* ρ*) vanishes as* θ *increases and does not impact at all sufficiently high (here* $\theta > \frac{1}{2}$*) sender types.*

According to Theorem 1 the presence of undesirable sender types—even when their measure is arbitrarily small—prevents *all* sender types from disclosing their information to the receiver, despite the fact that such disclosure would be beneficial for the receiver and for all types above ρ. Our next result shows that when the size of the equilibrium interval containing the undesirable types goes to zero (which in particular means that the set of undesirable types goes to zero), communication becomes arbitrarily precise. The proof of this result appears in Appendix A3 and uses Lemmas 3 - 5 in a similar manner to the Proof of Theorem 1.

Theorem 2. *For every* $\Delta > 0$ *there exists an* $\varepsilon > 0$ *such that if the leftmost interval (that contains all the toxic types) of the equilibrium partition is of length below* ε, *then all elements of the information partition of that equilibrium are intervals of length below* Δ.

Symmetry of the receiver's loss function also plays a key role in Theorem 2. This is demonstrated in the next example, which uses the asymmetric loss function of Example 1 to show that no matter how small the leftmost equilibrium interval is, the length of the rightmost interval is greater than $\frac{1}{2}$.

[9] Denote the right-adjacent interval by $[y, z]$. Then $a^*(y, z) = \frac{1}{3}y + \frac{2}{3}z$. On the other hand, from the indifference of the sender type $\theta = y$, this action can be expressed as $a(y, z) = \frac{4}{3}y - \frac{1}{3}x$. Thus, $z = \frac{3}{2}y - \frac{1}{2}x$. Consequently, the ratio between the lengths of an interval and its *left-adjacent* neighbor is $\frac{z-y}{y-x} = \frac{1}{2}$.

Example 2. *Assume that the state is uniformly distributed on* $[0,1]$. *Suppose that when* $a \neq N$, *the sender's payoff is* $u^S(a,\theta) = -(\theta - a)^2$, *while the receiver's payoff is,*

$$u^R(a,\theta) = \begin{cases} \theta - (\theta - a)^2 & , \quad \theta > a \\ \theta - 4(\theta - a)^2 & , \quad \theta \leq a. \end{cases}$$

Suppose that when the receiver's information set is $[x,y]$, *he does not choose* $a = N$. *In this case, it is easy to verify that he selects the action* $a^*(x,y) = \frac{2}{3}x + \frac{1}{3}y$. *Denote the right-adjacent interval to* $[x,y]$ *by* $[y,z]$. *Then the receiver's optimal action under the belief* $\theta \sim U[y,z]$ *is*

$$a^*(y,z) = \frac{2}{3}y + \frac{1}{3}z.$$

On the other hand, from the indifference of the sender type $\theta = y$, *this action can be expressed as* $a^*(y,z) - y = y - a^*(x,y)$, *i.e.,* $a^*(y,z) - y = y - (\frac{2}{3}x + \frac{1}{3}y)$, *or,*

$$a^*(y,z) = \frac{5}{3}y - \frac{2}{3}x.$$

Therefore, $\frac{2}{3}y + \frac{1}{3}z = \frac{5}{3}y - \frac{2}{3}x$, *which gives* $z - y = 2(y - x)$.

Fix an equilibrium and let $2\epsilon > 0$ *denote the length of the leftmost interval in the equilibrium partition. The above relation between the lengths of adjacent intervals implies that the equilibrium partition consists of finitely many intervals, say* k, *and thus we can write* $1 = 2\epsilon + 2^2\epsilon + ... + 2^k\epsilon$, *where the* j*-th summand on the RHS corresponds to the length of the* j*-th leftmost interval. We can write*

$$1 = \epsilon \sum_{j=1}^{k} 2^j = \epsilon \Big(\sum_{j=1}^{k-1} 2^j + 2^k \Big),$$

Note that the sum of lengths of all but the rightmost interval is below the length of the rightmost interval since $\sum_{j=1}^{k-1} 2^j = 2^k - 2 < 2^k$. *Thus, in any equilibrium, the length of the rightmost interval is greater than* $\frac{1}{2}$.

4 Linear values and uniform types

This section serves two purposes. First, it illustrates Theorem 1 by characterizing the unique Pareto efficient (interval) equilibrium for a simple specification of our model. Second, it demonstrates how a small set of toxic types can have a dramatic effect on the ability of all sender types to communicate with the receiver.

We focus on the case in which the distribution of types is uniform and the sender's type θ represents the value he generates when $a = \theta$. The main part of the section is devoted to characterizing the unique Pareto efficient equilibrium under the following additional assumptions. To highlight the role of the non-participation action in generating a conflict of interest between the sender and the toxic sender types, both players will have the *same exact* payoffs from any action in [0,1]. In addition, the loss function will be invariant to the sender's type, and the receiver's payoff from the action N will be sufficiently low so that $a \neq N$ is optimal for a completely uninformed receiver. We relax these assumptions at the end of the section to illustrate an extreme effect of the toxic types' externality on all the other types. Specifically, we present an example where $a = N$ is played with positive probability *only* if the receiver is certain that he interacts with desirable sender types.

More formally, assume that $\theta \sim U[0,1]$, and that $R(\theta) = S(\theta) = \theta$ and $r(\theta) = s(\theta) = 1$. That is, the players' payoff at state θ when the receiver chooses $a \neq N$ is

$$\theta - L(|a - \theta|),$$

for some strictly convex loss function $L(\cdot)$.[10] Also, assume that the sender's and receiver's payoffs from $a = N$ are $\sigma < -L(1)$ and $\rho \in \big(0, V(0,1)\big]$, respectively.[11]

[10] Our analysis in this section would remain unchanged if we were to assume instead that $S(\theta)$ is a constant, in which case, by the proof in Appendix A2, only interval equilibria exist.

[11] Gordon (2010) studies cheap talk environments with states in which the players agree on the optimal actions. He characterizes a *necessary* condition for the existence of equilibria with infinitely many actions. We can use the current specification to illustrate why our Theorem 1 does not follow from his characterization. In the notation of Gordon (2010), when $L(\cdot)$ is quadratic, for any interior agreement type x^*, $\alpha = \frac{\partial R}{\partial s}(x^*, x^*) = \frac{1}{2}$, $\beta = \frac{\partial R}{\partial t}(x^*, x^*) = \frac{1}{2}$, and $d = \frac{\frac{\partial^2 U^S}{\partial a \partial t}(a^*, x^*)}{\frac{\partial^2 U^S}{\partial a^2}(a^*, x^*)} = 1$. Since $\alpha + \beta = d$, Gordon's necessary

First, recall that, by the convexity of $L(\cdot)$, for any belief on θ, there is a unique action in $[0,1]$ that maximizes the receiver's expected payoff. Second, since $r(\theta)=1$ (in particular, the fact that $r(\theta)$ does not vary with θ), when the receiver believes that θ is uniformly distributed on some interval, the mid-point of that interval is the uniquely optimal action for the receiver, out of all actions different from N. The next observation follows from the indifference condition of the sender's threshold type between the two intervals.

Observation 4. *Any pair of adjacent intervals in the receiver's equilibrium information partition on which the receiver never plays action N must be of equal length.*

We now offer several useful observations on the function $V(x,y)$. Denote by

$$\bar{L}_\delta = \frac{1}{\delta}\int_0^\delta L(|\frac{\delta}{2}-\theta|)d\theta$$

the average loss given a belief that the state is uniformly distributed on an interval of length δ. The function $V(x,y)$ then takes the form

$$V(x,y)=\frac{1}{y-x}\int_x^y \theta d\theta - \frac{1}{y-x}\int_x^y L(|\frac{x+y}{2}-\theta|)d\theta = \frac{x+y}{2} - \bar{L}_{y-x}. \tag{1}$$

Since $L(\cdot)$ is increasing and strictly convex, the function $\bar{L}_\delta$ is also increasing and strictly convex (as a function of δ). This, in turn, implies that given any value of x, the function $V(x,y)$ is a concave function of y. Let $\theta_\rho \in (0,1]$ be the lowest sender type for which $V(0,\theta_\rho)=\rho$. The existence of such a type follows from the fact that $V(0,y)$ is a continuous function of y that satisfies $V(0,0)<\rho\leq V(0,1)$. Moreover, from the concavity of $V(0,y)$, there exists at most one such value below 1.

Denote by Q_K the partition of $[0,1]$ into $K\in\mathbb{N}$ equal-length intervals. We now characterize equilibria in which N is never played.

Proposition 1. *The set of the receiver's information partitions that are consistent with an equilibrium in which N is never played is given by $E_{noN}=\{Q_K \ : \frac{1}{K}\geq\theta_\rho\}$.*

condition is satisfied. Indeed, Gordon (2010) shows that his condition is not sufficient.

Proof. By (1), it is immediate that, given $\delta > 0$, $V(x, x+\delta)$ is increasing in x. Thus, if the receiver weakly prefers not to play N on a given interval, he would also (even strictly) prefer not to play N on any equal-length interval whose lower bound is shifted to the right. By Observation 4, only partitions into equal intervals are consistent with equilibria in which N is not played. □

We now turn to equilibria in which the action N is played with positive probability. The next lemma is the key for the main result of this section.

Lemma 1. *Let $[x, y]$ and $[y, z]$ be two adjacent intervals in an equilibrium partition. If the receiver mixes (between N and some action $a \neq N$) on either of these intervals, then the interval $[y, z]$ is strictly longer than $[x, y]$.*

Proof. Assume by contradiction that $y - x \geq z - y$. Since $\bar{L}_\delta$ is an increasing function of δ,

$$V(y,z) = \frac{y+z}{2} - \bar{L}_{z-y} > \frac{x+y}{2} - \bar{L}_{z-y} \geq \frac{x+y}{2} - \bar{L}_{y-x} = V(x,y). \tag{2}$$

Assume first that the receiver mixes on $[y, z]$. By (2), $\rho = V(y, z) > V(x, y)$. But then N is uniquely optimal for the receiver on $[x, y]$ which, by Observation (1), cannot be consistent with the receiver's mixing on $[y, z]$. Next, suppose that the receiver plays N with positive probability on $[x, y]$. In this case, (2) implies that $V(y, z) > V(x, y) = \rho$ and, therefore, $a = \frac{y+z}{2}$ is uniquely optimal for the receiver on $[y, z]$. This also leads to a contradiction: since $y - x \geq z - y$, there exist sender types $\theta \in (x, y)$ that strictly prefer the mid-point of $[y, z]$ to a lottery between the mid-point of $[x, y]$ and N. □

Proposition 2. *For any equilibrium in which N is played, there exists a Pareto-dominating equilibrium in which the receiver never chooses N.*

Proof. Since $V(0, 1) \geq \rho$, a completely uninformative communication followed by the receiver's action $a = \frac{1}{2}$ constitutes an equilibrium. Moreover, the outcome obtained in this equilibrium Pareto dominates any (unconditional) mixing between N and $\frac{1}{2}$. Hence, by

Observation 1, it is left to consider informative equilibria (i.e., equilibria in which the receiver's information partition consists of at least two intervals), with the property that, following some sender's message, the receiver mixes between N and some action $a \neq N$.

Let e be an equilibrium and denote by $M \geq 2$ the number of intervals in the receiver's information partition under e. Since playing $a \neq N$ is optimal for the receiver on the leftmost interval, this interval must be weakly longer than θ_ρ. By Lemma 1 and Observation 4, the receiver's information partition in e consists of unequal intervals and the leftmost interval is the shortest interval in this partition. Therefore, $\frac{1}{M} > \theta_\rho$, which, in turn, implies that $Q_M \in E_{noN}$.

Since $L(\cdot)$ is increasing and convex, the partition Q_M attains the lowest expected loss among all partitions of the unit interval into M intervals. Since N is never selected under the equilibrium that corresponds to Q_M, both players strictly prefer that equilibrium to any equilibrium that partitions the unit interval into M intervals. □

Provided that N is not selected, the players' interests coincide. The expected payoff from Q_K equals $\frac{1}{K}\sum_{k=1}^{K} V(\frac{k-1}{K}, \frac{k}{K}) = \frac{1}{2} - \bar{L}_{\frac{1}{K}}$. Since $\bar{L}_{\frac{1}{K}}$ decreases when K increases, from the ex-ante perspective, the Pareto dominant partition in E_{noN} is the one with the maximal number of intervals. This leads to the following characterization.

Proposition 3. *The unique Pareto efficient equilibrium partitions the unit interval into M^* equal intervals, where M^* is the largest integer that satisfies $V(0, \frac{1}{M^*}) \geq \rho$. In the Pareto efficient equilibrium, the receiver never plays N.*

The following example illustrates the strong effect of a small proportion of toxic types.

Example 3. *Assume that $\rho = \frac{1}{10}$. The Pareto efficient equilibrium partitions the unit interval into at most 4 intervals. To see this, note that for any $y \leq \frac{2}{10}$, $V(0, y) = \frac{y}{2} - \bar{L}_y < \frac{y}{2} \leq \frac{1}{10} = \rho$. Thus, by Observation 4, every interval is strictly longer than $\frac{2}{10}$.*

We conclude this section with an illustration of another possible manifestation of the infection from toxic types in our model. The only aspect that is different from the specification considered earlier in this section is that now we allow the coefficient of the receiver's loss function, $r(\cdot)$, to vary with the sender's type. In the next example, all informative equilibria have the following structure: the state space is partitioned into two intervals and the probability of N on the left interval—the one that contains all of the toxic types—is *strictly lower* than the probability of N on the right interval. In the Pareto efficient equilibrium, N is played with positive probability *only* on the right interval (which contains only viable sender types with whom the players' interests are perfectly aligned).

Example 4. *Suppose that the sender's payoff from the receiver's action* $a \in [0,1]$ *at state* θ *is* $\theta - \Lambda(|\theta - a|)$*, while his payoff from* $a = N$ *is* $\sigma < -\Lambda(1)$*; the receiver's payoff at state* θ *from* $a \in [0,1]$ *is given by* $\theta - r(\theta)(\theta - a)^2$*, where*

$$r(\theta) = \begin{cases} 4 & , \quad \theta < \frac{3}{4} \\ 4e^{z(\theta - \frac{3}{4})^2} & , \quad \theta \geq \frac{3}{4} \end{cases}$$

and $z > 0$ *is a constant.*[12]

By Lemma 2 in Appendix A1, there exists z^* *for which* $V(\frac{3}{4},1)(z^*) = V(0,\frac{3}{4})$*. In what follows, we assume that* $z = z^*$ *and set* $\rho = \frac{3}{16}$ *(so that* $\rho = V(0,\frac{3}{4}) = V(\frac{3}{4},1)$*).*

We show below that all informative interval equilibria have the following structure: the unit interval is partitioned at the threshold $\frac{3}{4}$*; upon learning that the state belongs to the left (right) interval the receiver chooses* N *with probability* q_l (q_r)*. Conditional on not playing* N*, the receiver chooses* $a_l = \frac{3}{8}$ *on the left interval and some* $a_r \in (\frac{3}{4}, 1)$ *on the right interval. Since* a_r *is necessarily closer to the sender's threshold type* $\theta = \frac{3}{4}$*, the indifference of that type between joining either of the intervals implies that* $q_l < q_r$*.*

Clearly, the receiver's payoff from any equilibrium that partitions the unit interval into two at $\frac{3}{4}$ *is exactly* ρ*. As for the sender, note that if we start from a pair of probabilities* (q_l, q_r) *that are consistent with an equilibrium, and we decrease* q_l*, then to restore the*

[12] The objective is to increase, in a convenient parametric way, the "importance" coefficient $r(\cdot)$ for types above $\frac{3}{4}$. The particular form is inessential. We chose the exponential form because it is convenient to guarantee that $r(\cdot)$ is differentiable.

indifference of the sender's threshold type, we need to decrease q_r *as well. Since the sender benefits from both decreases, the Pareto efficient equilibrium in this family satisfies* $0 = q_l < q_r$.

In the babbling equilibrium, the receiver chooses N *with certainty* ($V(0,1) < V(0, \frac{3}{4}) = \rho$)*. Hence, the sender is strictly worse off under the babbling equilibrium relative to any equilibrium in the aforementioned family. To see that other information partitions cannot be part of an equilibrium, recall that* $V(0, y) < V(0, \frac{3}{4}) = \rho$ *for all* $y \neq \frac{3}{4}$*. Thus, on any interval* $[0, y]$ *such that* $y \neq \frac{3}{4}$*, the receiver would choose* N *with certainty. Since* $\sigma < -\Lambda(1)$*, this can be consistent with an equilibrium only if the receiver plays* N *with certainty regardless of the sender's message, which is equivalent to the babbling equilibrium outcome.*

5 Related literature

A key feature of our model is that the receiver disagrees with some sender types over the payoff from not interacting, but conditional on interacting, there is no conflict of interest. A similar feature also appear in Che, Dessein, and Kartik (2013). In their baseline model, a receiver faces three options, A, B or C. The value of C is fixed and commonly known while the values of A and B are drawn from two distinct distributions and are only observed by the sender. The payoff from A and B is the same for both players, but their payoffs from C differ: it is zero for the sender but positive for the receiver. In the game analyzed by the authors, the sender recommends either A or B, and the receiver either follows the recommendation or opts out. The main result is that for an intermediate range of C payoffs for the receiver, the sender is biased towards one of the projects in the sense that he recommends it also in states where it is the inferior project.

There are several substantial differences between this paper and ours. First, their result does not stem from an infection argument. Rather, it follows from pooling the states in which the "biased" project is the best with some states in which it is the worst such that conditional on recommending the unbiased project, the receiver is indifferent between following the recommendation and opting out. Second, in contrast to our result, there is never a state where it is common knowledge that there is no conflict of interests and yet the sender is not being truthful. Third, unlike us, for some distributions, the sender will be truthful in all the states where there is no conflict of interest (in fact, this is true in their leading example).

A recent paper by Dilmé (2019) also studies a model in which the sender and receiver agree on the optimal action. Thus, in contrast to our setting, there exists a fully separating equilibrium in which the sender reveals the state. The paper departs from the standard Crawford-Sobel setting by assuming that the two players have different asymmetric loss functions. The paper pursues a different objective than us, which is to show that introducing small noise to the sender's observation of the state coarsens the communication.

The loss of informative communication in our model is reminiscent of that in the literature on cheap talk with uncertainty over the sender's bias (notable examples include Morris (2001) and Morgan and Stocken (2003)). A key distinction between our work and these papers is the following. In these papers there is a positive probability that the sender is biased *in every state*, and hence the reason coarse communication arises is similar to that of Crawford and Sobel (1982). By contrast, in our model, even though all messages are coarse, whenever the receiver gets a message from types above the lowest equilibrium interval, he is certain that his interests are perfectly aligned with the sender.

The idea that even an extremely rare event or set of types can "infect" all the other states/types and have a dramatic effect on the equilibrium has been previously demonstrated in the literature on reputation in repeated games (pioneering works include Kreps and Wilson (1982) and Fudenberg and Levine (1989)) and in the global games literature (Rubinstein (1989), Carlsson and van Damme (1993) and the subsequent papers). More recently, Di Pei (2017) and Blume (2018) apply these contagion arguments to show that in cheap-talk games with higher-order uncertainty (either on the sender's preferences, or the set of available messages), coarse communication may arise even if interests are aligned

(Di Pei (2017)) or when there are sufficiently many messages (Blume (2018)).

Relative to these papers, our contribution is twofold: (i) we identify a new way to start the contagion, namely, through the presence of toxic types, and (ii) we characterize new sufficient conditions under which contagion can spread throughout the type space, that is, there is no accumulation point in the iteration process for finding end points of intervals. These conditions accommodate contagion in a broader set of environments where previous sufficient conditions (for example, see footnote 8).

6 Concluding remarks

This paper analyzed a setting in which two parties agree on the action that maximizes the gain from joint interaction, but one of the parties wants to enter this interaction only if the other side is sufficiently "able." Examples of such scenarios include assigning tasks that best fit a worker's skills as long as these skills are above some level, or trying to match an individual with the object most valuable to him, provided this value is above some threshold.

Our analysis focused on interactions where one agent always gains from the interaction, whereas the other agent, the one who controls the action, gains only if it interacts with types above some threshold. We showed that when types are unobserved, then even when the threshold is arbitrarily small—so that the interaction is profitable with almost all types—the two parties will fail to realize the full potential of their interaction. In particular, the incentive of the unprofitable types to hide their identity "infects" all types and prevents communication of mutually beneficial information. However, as the lowest set of types that pool in equilibrium becomes smaller, communication improves in the sense that every set of types who pool in equilibrium becomes smaller.

Our results suggest that, more generally, when an uninformed decision-maker has conflicting interests even with an arbitrarily small set of types of an informed agent, the two may be unable to communicate mutually beneficial information. While our analysis has focused on a particular form of conflicting interests, it may extend to a wider range of applications. Some potential examples include situations where the expected returns from projects are higher the more ambitious and riskier they are, but investors and entrepreneurs have different risk thresholds for taking on the projects. Similarly, the gains for a lobbyist and a politician from enacting (or canceling) a new law or regulation increases with the potential harm it prevents, but contrary to the lobbyist, the politician may be willing to push for the reform only if this harm is sufficiently high. We hope that our work will spur future research on a broader class of environments that includes these and related applications.

References

[1] Blume, Andreas (2018). "Failure of common knowledge of language in common-interest communication games." *Games and Economic Behavior* 109 : 132-155.

[2] Carlsson, Hans, and Eric Van Damme (1993). "Global games and equilibrium selection." *Econometrica* 6(5): 989–1018.

[3] Chakraborty, Archishman, and Rick Harbaugh (2007). "Comparative cheap talk." *Journal of Economic Theory* 132(1): 70–94.

[4] Chakraborty, Archishman, and Rick Harbaugh (2010). "Persuasion by cheap talk." *American Economic Review* 100(5): 2361–2382.

[5] Che, Yeon-Ke, Wouter Dessein and Navin Kartik (2013). "Pandering to persuade." *American Economic Review* 103(1): 47–79.

[6] Chen, Ying, Navin Kartik, and Joel Sobel (2008). "Selecting cheap-talk equilibria." *Econometrica* 76(1): 117–136.

[7] Crawford, Vincent P., and Joel Sobel (1982). "Strategic information transmission." *Econometrica* 50(6): 1431–1451.

[8] Di Pei, Harry (2017). "Uncertainty about Uncertainty in Communication." working paper.

[9] Francesc Dilmé (2019). "Skewed information transmission," Working Paper, University of Bonn.

[10] Fudenberg, Drew, and David Levine (1989). "Reputation and equilibrium selection in games with a patient player." *Econometrica* 57(4): 759–78.

[11] Gordon, Sidartha (2010). "On infinite cheap talk equilibria." working paper.

[12] Kreps, David M., and Robert Wilson (1982). "Reputation and imperfect information." *Journal of economic theory* 27(2): 253–279.

[13] Morgan, John, and Phillip C. Stocken (2003). "An analysis of stock recommendations." *RAND Journal of economics* 34(1): 183–203.

[14] Morris, Stephen (2001). "Political correctness." *Journal of Political Economy* 109(2): 231–265.

[15] Ottaviani, Marco, and Peter N. Sørensen (2006a). "Professional Advice." *Journal of Economic Theory* 126(1): 120–142.

[16] Ottaviani, Marco, and Peter N. Sørensen (2006b). "Reputational Cheap Talk." *Rand Journal of Economics* 37(1): 155–175.

[17] Rubinstein, Ariel (1989). "The electronic mail game: Strategic behavior under 'almost common knowledge'." *American Economic Review* 79(3): 385–391.

[18] Sobel, Joel (2013). "Giving and receiving advice." Advances in Economics and Econometrics 1: 305–341.

Appendix

A1. Constructing an example with a non-monotonic partition equilibrium

Let $\theta \sim U[0,1]$; the sender's payoff from the receiver's action $a \in [0,1]$ at state θ is $-(\theta - a)^2$ and his payoff from $a = N$ is $\sigma < -1$; the receiver's payoff at state θ from $a \in [0,1]$ is given by $\theta - r(\theta)(\theta - a)^2$ where

$$r(\theta) = \begin{cases} 4 & , \quad \theta < \frac{3}{4} \\ 4e^{z^*(\theta - \frac{3}{4})^2} & , \quad \theta \geq \frac{3}{4} \end{cases}$$

and $z^* > 0$ satisfies $\rho = \frac{3}{16} = V(0, \frac{3}{4}) = V(\frac{3}{4}, 1)$. As shown in the last example in Section 4, in the Pareto efficient equilibrium, the receiver learns whether the state is below or above $\frac{3}{4}$; in the former case he chooses $a_l = \frac{3}{8}$ and in the latter case he chooses N with probability q and some action $a_r \in (\frac{3}{4}, 1)$ with probability $1 - q$. Denote this lottery by α_r. We now modify the sender's preferences to obtain a specification where the receiver's induced information partition does not consist of intervals, namely, one where the mapping between types and participation actions will not be monotonic. For computational convenience, we will modify the above specification such that the receiver's induced information partition will be identical to the one above up to a singleton.

Let $\hat{\theta} \in (\frac{3}{4}, a_r)$ and let the payoff of type $\hat{\theta}$ from non-participation actions be $-\hat{s}(\hat{\theta} - a)^2$ such that he is indifferent between a_l and α_r. Such an $\hat{s} \in (0,1)$ is unique: considering the preferences $-s(\hat{\theta} - a)^2$, it is easy to see that when $s = 0$ type $\hat{\theta}$ strictly prefers a_l to α_r; from the equilibrium in the original specification it is obvious that when $\hat{s} = 1$ the strict preference is reversed; and the sender's gain from inducing a_l instead of α_r is monotonically decreasing in s (see (3) for $\theta = \hat{\theta}$).

We now modify the sender's preferences for types near $\hat{\theta}$ to obtain a specification that is consistent with our modeling assumptions. Since the gain from inducing a_l instead of

α_r,

$$-s(\theta - a_l)^2 + [(1-q)s(\theta - a_r)^2 + q\sigma], \tag{3}$$

is differentiable in s and θ, there exists $\varepsilon > 0$ such that $[\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon] \subset (\frac{3}{4}, a_r)$ and a differentiable function $\hat{s} : [\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon] \to (0, 1]$ such that (i) $\hat{s}(\hat{\theta}) = \hat{s}$, (ii) $\hat{s}(\hat{\theta} - \varepsilon) = \hat{s}(\hat{\theta} + \varepsilon) = 1$, (iii) $\hat{s}'(\hat{\theta} - \varepsilon) = \hat{s}'(\hat{\theta} + \varepsilon) = 0$, and (iv) a_r is strictly better than α_l for all $\theta \in [\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon] - \{\hat{\theta}\}$.

By the choice of $\hat{s}$, type $\hat{\theta}$ is indifferent between a_r and α_l. Under the sender's modified preferences from non-participation actions, $-s(\theta)(\theta - a)^2$, where

$$s(\theta) = \begin{cases} 1 & , \quad \theta \notin [\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon] \\ \hat{s}(\theta) & , \quad \theta \in [\hat{\theta} - \varepsilon, \hat{\theta} + \varepsilon], \end{cases}$$

and the receiver's original preferences, there exists an equilibrium where a_l is induced by sender types $[0, \frac{3}{4}] \cup \hat{\theta}$, and all other types induce α_r.

Lemma 2. *Let $\{R_z\}_{z \in \mathbb{R}_+}$ denote a family of receiver's preferences such that, for each $z \in \mathbb{R}_+$, the payoff at state θ from $a \in [0, 1]$ is $\theta - r_z(\theta)(\theta - a)^2$, where*

$$r_z(\theta) = \begin{cases} 4 & , \quad \theta < \frac{3}{4} \\ 4e^{z(\theta - \frac{3}{4})^2} & , \quad \theta \geq \frac{3}{4}. \end{cases}$$

Let $V(x, y)(z)$ denote the value of $V(x, y)$ under R_z. There exists $z^* > 0$ such that $V(\frac{3}{4}, 1)(z^*) = V(0, \frac{3}{4})(z^*)$.

Proof. Note that for all z, the function $V(0, y)(z)$ is increasing in y for all $y < \frac{3}{4}$. Since $z \geq 0$, for all $y > \frac{3}{4}$, we have $4e^{z(\theta - \frac{3}{4})^2} \geq 4$, with strict inequality for $z > 0$. Thus, for all z,

$$V(0, y)(z) \leq \frac{1}{y} \int_0^y \theta - 4(\theta - \frac{y}{2})^2 d\theta$$

with equality for all $y \leq \frac{3}{4}$ and a strict inequality for $y > \frac{3}{4}$ whenever $z > 0$. The expression on the R.H.S. is maximized at $y = \frac{3}{4}$, where its value is $\frac{3}{16}$. $V(\frac{3}{4}, 1)(z)$ is continuous, strictly decreasing, satisfies $V(\frac{3}{4}, 1)(0) > \frac{3}{16}$, and can be made arbitrarily small by choosing large values for z. Thus, there exists a z^* for which $V(\frac{3}{4}, 1)(z^*) = V(0, \frac{3}{4})$. $\square$

A2. Sufficient condition for existence of only interval equilibria

We now show that if $S(\theta)$ and $s(\theta)$ are constants, then all equilibria induce a monotonic partition on the set of sender types.

Assume, by contradiction, that there exists an equilibrium with the following properties. There exist three types, $\theta_1 < \theta_2 < \theta_3$, such that the support of θ_1's and θ_3's strategies include the same message m_1, while the support of θ_2's strategy includes a message m_2, which is not in the support of θ_1's and θ_3's strategies. The receiver responds to m_1 by choosing N with probability q_1 and an action a_1 with probability $1 - q_1$. He responds to m_2 by choosing N with probability q_2 and an action a_2 with probability $1 - q_2$.

Suppose that $a_1 < a_2$. If $q_1 \leq q_2$, then since type θ_3 weakly prefers the message m_1 to m_2, the lower type θ_2 must strictly prefer m_1 to m_2, a contradiction. Hence, q_1 must be greater than q_2. But then the fact that type θ_2 weakly prefers m_2 to m_1 implies that the higher type θ_3 strictly prefers m_2 to m_1, a contradiction.

Suppose next that $a_1 > a_2$. If $q_1 \leq q_2$, then the fact that type θ_1 weakly prefers m_1 to m_2 implies that the higher type θ_2 strictly prefers m_1 to m_2, a contradiction. It follows that q_1 must be greater than q_2. But then the fact that type θ_2 weakly prefers m_2 to m_1 implies that the lower type θ_1 strictly prefers m_2 to m_1, a contradiction.

Finally, if $a_1 = a_2$, then it must be that $q_1 = q_2$. But in this case, the two messages m_1 and m_2 can be merged into one message.

A3. Omitted Proofs

Lemma 3. *There exists* $c > 0$ *(that depends on the specification of the model) such that, for any non-degenerate interval* $I = [\underline{I}, \bar{I}] \subseteq [0, 1]$*, the receiver's (unique) optimal participation action* $a^*(I) \neq N$ *conditional on* $\theta \in I$*, satisfies*

$$\frac{\underline{I} + \bar{I}}{2} - c \cdot |I|^2 < a^*(I) < \frac{\underline{I} + \bar{I}}{2} + c \cdot |I|^2.$$

Proof. The (unique) action $a^*(I) \neq N$ that attains $V(\underline{I}, \bar{I})$ solves

$$\min_{a \in I} \int_I L(|a - \theta|) r(\theta) f(\theta) d(\theta). \tag{4}$$

Let $0 < \underline{r} < \bar{r} < \infty$ and $0 < \underline{f} < \bar{f} < \infty$ satisfy $\underline{f} < f(\theta) < \bar{f}$ and $\underline{r} < r(\theta) < \bar{r}$ for all $\theta \in [0, 1]$. Such values exist because $r(\cdot)$ and $f(\cdot)$ are continuous on $[0, 1]$. In addition, denote by $\underline{r}_I$ and $\underline{f}_I$ the minimal values of the functions $r(\cdot), f(\cdot)$ on I. It follows that

$$L(|a - \theta|) r(\theta) f(\theta) = L(|a - \theta|) \Big[\underline{r}_I + (r(\theta) - \underline{r}_I) \Big] \Big[\underline{f}_I + (f(\theta) - \underline{f}_I) \Big].$$

Hence, (4) can be rewritten as

$$\begin{aligned} \min_{a \in I} \ & \Big\{ \underline{r}_I \underline{f}_I \int_I L(|a - \theta|) d\theta + \int_I L(|a - \theta|) \underline{r}_I (f(\theta) - \underline{f}_I) d\theta + \\ & + \int_I L(|a - \theta|) \underline{f}_I \, (r(\theta) - \underline{r}_I) \, d\theta + \int_I L(|a - \theta|)(r(\theta) - \underline{r}_I)(f(\theta) - \underline{f}_I) d\theta \Big\}. \end{aligned} \tag{5}$$

Let $A \in \mathbb{R}$ be greater than the Lipschitz constants of both f and r. The following bounds hold for all $\theta \in I$:

$$f(\theta) - \underline{f}_I \leq A[\theta - \underline{I}]; \quad r(\theta) - \underline{r}_I \leq A[\theta - \underline{I}]; \quad (r(\theta) - \underline{r}_I)(f(\theta) - \underline{f}_I) \leq A^2 [\theta - \underline{I}], \tag{6}$$

where the first two relations follow directly from the definition of A and the last one also relies on $\theta - \underline{I} < 1$.

We now define the following modified problem.

$$\min_{a\in I}\ \underline{rf}\int_I L(|a-\theta|)d\theta+\int_I (A(\bar{r}+\bar{f})+A^2)(\theta-\underline{I})\cdot L(\bar{I}-a)d\theta. \tag{7}$$

The first summands of the two problems, (5) and (7), differ only in that the weight on $\int_I L(|a-\theta|)d\theta$ in the modified problem is (weakly) lower, $\underline{rf}\leq \underline{r}_I\underline{f}_I$. The remaining summands in (5) are replaced by an expression that is minimized at the right boundary of the interval, $\bar{I}$. Moreover, by (6), for each θ,

$$(A(\bar{r}+\bar{f})+A^2)(\theta-\underline{I})\geq \underline{r}_I(f(\theta)-\underline{f}_I)+\underline{f}_I(r(\theta)-\underline{r}_I)+(r(\theta)-\underline{r}_I)(f(\theta)-\underline{f}_I).$$

Hence, for each θ, the coefficient of $L(\bar{I}-a)$ in the modified problem (7) is strictly greater than the sum of coefficients of $L(\cdot)$ in the second, third, and fourth summands in the original problem (5).

It follows that the solution to (7), which we denote by $a^{**}(I)$, is weakly greater than $a^*(I)$. Note that (7) can be rewritten as

$$\min_{a\in I}\ \underline{rf}\int_I L(|a-\theta|)d\theta+(A(\bar{r}+\bar{f})+A^2)\cdot\frac{|I|^2}{2}\cdot L(\bar{I}-a). \tag{8}$$

$L(\cdot)$ is differentiable and convex, so $a^{**}(I)$ can be derived from the first order condition:

$$\underline{rf}\left[L(a^{**}(I)-\underline{I})-L(\bar{I}-a^{**}(I))\right]-(A(\bar{r}+\bar{f})+A^2)L'(\bar{I}-a^{**}(I))\frac{|I|^2}{2}=0. \tag{9}$$

Since $a^{**}(I)\geq\frac{\underline{I}+\bar{I}}{2}$, by the mean value theorem and the strict convexity of L, there exists a unique $Z\in[L'(\bar{I}-a^{**}(I)),L'(a^{**}(I)-\underline{I})]$ such that

$$L(a^{**}(I)-\underline{I})-L(\bar{I}-a^{**}(I))=Z\left[(a^{**}(I)-\underline{I})-(\bar{I}-a^{**}(I))\right]=Z[2a^{**}(I)-(\underline{I}+\bar{I})].$$

Thus, (9) can be rewritten as

$$a^{**}(I)=\frac{\underline{I}+\bar{I}}{2}+\frac{(A(\bar{r}+\bar{f})+A^2)}{\underline{rf}}\cdot\frac{L'(\bar{I}-a^{**}(I))}{Z}\cdot\frac{|I|^2}{4},$$

and since $Z \geq L'(\bar{I} - a^{**}(I))$, letting $c = \frac{(A(\bar{r}+\bar{f})+A^2)}{4\underline{r}\underline{f}}$, we obtain the upper bound,

$$a^*(I) \leq a^{**}(I) \leq \frac{\underline{I} + \bar{I}}{2} + c \cdot |I|^2. \tag{10}$$

The lower bound $\frac{\underline{I}+\bar{I}}{2} - c \cdot |I|^2 \leq a^*(I)$ is obtained in an analogous way by considering instead of the modified problem (7), one where $L(\bar{I} - a)$ is replaced with $L(a - \underline{I})$. $\square$

Lemma 4. *Let c be the constant from Lemma 3 and assume that the non-degenerate interval $I = [x, y]$ is an element of the receiver's information partition in a given equilibrium. Also, assume that $\frac{x+y}{2} + c \cdot (y - x)^2 \in I$. Let I' be an element adjacent to I in that information partition.*[13] *Then* $\frac{|I'|}{|I|} \geq 1 - 4c|I|$.

Proof. First we consider I's right-adjacent interval which we denote $[y, z]$. Since our objective is to determine a lower bound on $|I'|$ that satisfies $|I'| < |I|$, we can assume that $z - y \leq y - x$ (otherwise the claim holds trivially). By Lemma 3,

$$a^*([x, y]) \leq \frac{x + y}{2} + c \cdot (y - x)^2 = y - \frac{y - x}{2} + c \cdot (y - x)^2.$$

We can hence obtain the following, where the first inequality follows from the indifference of the sender's threshold type $\theta = y$ and the second from Lemma 3,

$$y + \frac{y - x}{2} - c \cdot (y - x)^2 \leq a^*([y, z]) \leq \frac{y + z}{2} + c \cdot (z - y)^2.$$

This yields $y - x - 2c \cdot (z - y)^2 - 2c \cdot (y - x)^2 \leq z - y$. The assumption $z - y \leq y - x$ implies $y - x - 4c \cdot (y - x)^2 \leq z - y$. Dividing by $y - x$ gives

$$1 - 4c \cdot (y - x) \leq \frac{z - y}{y - x}.$$

Similarly, we now consider I's left-adjacent interval which we denote $[w, x]$. Like before,

[13] The interval I' is right-adjacent to I if $\sup(I) = \inf(I')$ and it is left-adjacent to I if $\sup(I') = \inf(I)$.

we can assume $x - w \leq y - x$. By Lemma 3,

$$a^*([x, y]) \geq \frac{x+y}{2} - c \cdot (y-x)^2 = x + \frac{y-x}{2} - c \cdot (y-x)^2.$$

Hence, we can write, $x + \frac{y-x}{2} - c \cdot (y-x)^2 \geq a^*([w, x]) \geq \frac{w+x}{2} - c \cdot (x-w)^2$, which gives $x - w \geq y - x - 2c \cdot (x-w)^2 - 2c \cdot (y-x)^2$. Using the assumption that $x - w \leq y - x$, we obtain $x - w \geq y - x - 4c \cdot (y-x)^2$. Dividing by $y - x$ gives

$$\frac{x-w}{y-x} \geq 1 - 4c \cdot (y-x).$$

□

Lemma 5. *Let $(x_i)_{i=1}^{\infty}$ be a sequence of real numbers such that $x_1 > 0$ and for each $i \in \mathbb{N}$, $x_{i+1} = (1 - \psi x_i)x_i$ where the constant $\psi > 0$ satisfies $1 - \psi x_1 > 0$. Then $\sum_{i=1}^{\infty} x_i = \infty$.*

Proof. Let $\delta_i = 1 - \psi x_i$. For each $k \in \mathbb{N}$, by comparing $\sum_{i=k}^{\infty} x_i$ to the convergent geometric sequence whose first element is x_k and the common ratio is $\delta_k < 1$, we obtain

$$\sum_{i=1}^{\infty} x_i = \sum_{i=1}^{k-1} x_i + \sum_{i=k}^{\infty} x_i \geq \sum_{i=1}^{k-1} x_i + x_k \sum_{j=0}^{\infty} \delta_k^j = \sum_{i=1}^{k-1} x_i + \frac{x_k}{1-\delta_k} = \sum_{i=1}^{k-1} x_i + \frac{1}{\psi}$$

Since $\frac{1}{\psi} > 0$, assuming that $\sum_{i=1}^{\infty} x_i$ is finite leads to a contradiction, by taking the limit as $k \to \infty$ on both sides

$$\sum_{i=1}^{\infty} x_i \geq \sum_{i=1}^{\infty} x_i + \frac{1}{\psi} > \sum_{i=1}^{\infty} x_i$$

as, by definition, $\lim_{k\to\infty} \sum_{i=1}^{k-1} x_i = \sum_{i=1}^{\infty} x_i$. □

Using the above lemmas, we can now prove Theorem 1:

Proof of Theorem 1. Let $[0, \theta_1]$, where $\theta_1 > 0$, be the leftmost interval in the equilibrium information partition (Observation 2). By Observation 3, we can define a sequence of interval thresholds $(\theta_k)_{k=0}^{\mathcal{K}}$, where $\theta_0 = 0$ and $\mathcal{K}$ is either an element of $\mathbb{N}$ or ∞, such that for any $k < \mathcal{K}$, the interval $[\theta_k, \theta_{k+1}]$ (up to its boundaries) is an element of the equilibrium information partition. To prove the Theorem, we need to show that $\mathcal{K}$ must be finite.

Assume by contradiction that $\mathcal{K} = \infty$. Let $y_k = \theta_k - \theta_{k-1}$ for each $k \in \mathcal{K}$. Our assumption that $\mathcal{K} = \infty$ implies that as k increases, the elements y_k become arbitrarily small. To simplify notation, we can assume WLOG that y_1 (i.e., θ_1) is sufficiently small such that $1 - 4c \cdot y_1 > 0$ where $c = \frac{A(\bar{r}+\bar{f})+A^2}{4\underline{r}\underline{f}}$ is the constant defined in Lemma 3 (otherwise, we can begin the construction starting from an element y_k for which this holds).

We now define the sequence $(x_k)_{k=1}^{\infty}$ as follows. Let $x_1 = \theta_1$, and $x_{k+1} = (1 - 4c \cdot x_k)x_k$. By iteratively applying Lemma 4, it follows that $x_k \leq y_k$ for all $k \in \mathcal{K}$. However, by Lemma 5, the series $\sum_{k=1}^{\infty} x_k$ diverges. Hence, there can be only finitely many $\theta_K = \sum_{k=1}^{K} y_k$ for which $\theta_K \leq 1$, a contradiction to our assumption that $\mathcal{K} = \infty$. $\square$

Proof. Let $\Delta > 0$ such that $1 - 4c \cdot \Delta > 0$, where $c = \frac{A(\bar{r}+\bar{f})+A^2}{4\underline{r}\underline{f}}$ is the constant defined in Lemma 3. Consider the monotonically decreasing sequence $(x_k)_{k=1}^{\infty}$ where $x_1 = \Delta$ and $x_{k+1} = (1 - 4c \cdot x_k)x_k$ for all $k \in \mathbb{N}$. By Lemma 5, the series $\sum_{k=1}^{\infty} x_k$ diverges. Hence, there exists $K \in \mathbb{N}$ such that $\sum_{k=1}^{K-1} x_k > 1$.

Consider an equilibrium information partition and assume that its longest interval I satisfies $|I| \geq \Delta$. By applying iteratively Lemma 4 for left-adjacent intervals, we can put a lower bound on the lengths of all intervals to the left of I, using the sequence $(x_k)_{k=1}^{\infty}$. In particular, the length of every interval to the left of I must be at least x_K. In other words, if the leftmost interval of the equilibrium partition is shorter than x_K, all the intervals in the equilibrium partition must be below Δ. Hence, setting $\varepsilon = x_K$ completes the proof. $\square$