DISCUSSION PAPER SERIES

DP15511

TOXIC TYPES AND INFECTIOUS COMMUNICATION BREAKDOWN

Kfir Eliaz and Alexander Frug

INDUSTRIAL ORGANIZATION

ISSN 0265-8003

TOXIC TYPES AND INFECTIOUS COMMUNICATION BREAKDOWN

Kfir Eliaz and Alexander Frug

Discussion Paper DP15511
Published 03 December 2020
Submitted 29 November 2020

Centre for Economic Policy Research
33 Great Sutton Street, London EC1V 0DX, UK
Tel: +44 (0)20 7183 8801
www.cepr.org

TOXIC TYPES AND INFECTIOUS COMMUNICATION BREAKDOWN

Abstract

We introduce a new reason for why an informed sender may not be able to communicate(via cheap talk) his private information to an uninformed receiver. Our framework has two novel features: (i) conditional on interacting, both parties agree on the optimal action to take given the sender's information, and (ii) there are some sender types (the "toxic" types) with which the receiver prefers not to interact. Our main result establishes that for a broad class of preferences, any interval equilibrium (where each message is associated with an interval of types) induces only finitely many actions in the support of the receiver's strategy. For a canonical class of preferences with uniformly distributed types, we characterize the Pareto efficient(interval) equilibria and illustrate how communication is adversely affected even with a small set of toxic types. In addition, we show that introducing a second stage in which the receiver gets a noisy signal on the sender type can have a dramatic effect on the first-stage communication.

JEL Classification: D83

Keywords: cheap talk, contagion

Kfir Eliaz - kfire@post.tau.ac.il
Tel-Aviv University and CEPR

Alexander Frug - alexander.frug@upf.edu
Universitat Pompeu Fabra

Acknowledgements
We thank Fabrizio Germano, 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.

Toxic Types and Infectious Communication Breakdown*

Kfir Eliaz and Alexander Frug†

First version: November 2020

Abstract

We introduce a new reason for why an informed sender may not be able to communicate (via cheap talk) his private information to an uninformed receiver. Our framework has two novel features: (i) conditional on interacting, both parties agree on the optimal action to take given the sender's information, and (ii) there are some sender types (the "toxic" types) with which the receiver prefers not to interact. Our main result establishes that for a broad class of preferences, any interval equilibrium (where each message is associated with an interval of types) induces only finitely many actions in the support of the receiver's strategy. For a canonical class of preferences with uniformly distributed types, we characterize the Pareto efficient (interval) equilibria and illustrate how communication is adversely affected even with a small set of toxic types. In addition, we show that introducing a second stage in which the receiver gets a noisy signal on the sender type can have a dramatic effect on the first-stage communication.

*We thank Fabrizio Germano, Navin Kartik, Elliot Lipnowski, and Joel Sobel for helpful comments.

†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 the informed party ("the sender") and the uninformed party ("the receiver") disagree on what is the optimal action in each state.[1] However, oftentimes, the receiver's interests are perfectly aligned with the sender's when the two interact, but there are some circumstances or states in which the receiver is better off not interacting with the sender.

For example, in mentoring, 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 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 that best fits his skills and experience, he may prefer not to work with a person whose skill or experience is too low, or who can only work on unprofitable tasks. Similarly, potential investors are interested only in viable projects or ideas, and in qualified entrepreneurs. In both examples, once a manager hires or assigns a candidate to a task, and an investor invests in an entrepreneur, the two sides will share the same objective, and in many situations will agree on the optimal course of action.

To capture situations like the above examples, we consider a model with the following features. Conditional on interacting with the sender, the receiver's preferences are *perfectly* aligned with those of the sender — both agree on the optimal action in each state. In addition, the players' gains from choosing the optimal action in a state can vary across states. However, while all sender types want to interact with the receiver, there is a small set of "toxic" sender types with whom

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

the receiver prefers not to interact *regardless* of the action he chooses. These types then have no incentive to reveal themselves to the receiver. Absent these types, there would be a fully revealing equilibrium. However, the presence of these toxic types—even when their measure is arbitrarily small—contaminates the ability of *all* sender types to communicate with the receiver.

Our model consists of a continuum of sender types (represented by an interval) and symmetric loss functions where conditional on interacting with the sender, both players want the receiver to match the sender's type. For expositional purposes, the toxic types in our model are placed at the bottom of the type space. However, we can easily extend our results to accommodate the toxic types residing anywhere in the type space, provided that the set of these types contains an interval. 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). In this case, toxic types are located at the top of the type space and their desire to pool with slightly lower types prevents fully informative communication with *all* types.

In our analysis, we focus on "interval equilibria" or equilibria that induce a partition of the sender type space into intervals. In these equilibria, messages can be naturally interpreted as 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 increases as the likelihood of drawing an undesirable sender type decreases, their total number remains finite. Our result has the following broader implication for environments with symmetric loss functions where the receiver and sender always agree on the optimal action in each state: 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 communication breaks down in our model

[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.

(or that the receiver's equilibrium partition has only finitely many intervals) requires different techniques from those used in the cheap-talk model of Crawford and Sobel (1982) and many of its extensions.

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 higher types.

Some of our motivating examples capture *dynamic* rather than static interaction. In such on-going relationships, sometimes the receiver observes additional signals about the sender's type as the interaction unfolds. For instance, after instructing an assistant or after assigning him to some task, a supervisor may be able to observe some measure of performance, which may provide a good indication of the assistant's skills or fitness for the job. This may allow the supervisor to re-optimize and either assign a more suitable task or end his relation with the assistant. However, since it may take some time until performance can be measured, or since such monitoring opportunities may be rare, the initial training or assignment, which is based on the assistant's self-report, can still be crucial to the overall surplus from the interaction. This raises the question of whether the initial communication phase improves or deteriorates when the receiver (the supervisor) has a future opportunity to obtain a signal about the sender's (the assistant's) type.

To address this question we consider a simple two-period example with the following features. At the beginning of the first period, an agent (the sender) reports a type to a principal (the receiver). Based on this report, the principal assigns the agent to a task or chooses not to employ him in which case the game ends. If the game continues, then at the end of the first period the principal observes a noisy signal of the agent's type. Specifically, with some probability $p < \frac{1}{2}$ the observation is pure noise, but with the complementary probability, the principal's observation coincides with the agent's type. Given his observation, the principal decides whether to keep the initial task assignment, assign the agent a different task, or fire the agent. We show that noisy monitoring severely hampers communication: in any (interval) equilibrium there is either little or no communication in the first period: the receiver's equilibrium information partition consists of at most *two* intervals. Furthermore, we illustrate that the effect on communication can be so severe that both players may be better off in the absence of monitoring.

The conflict of interest in our model, as well as its implications, is reminiscent of the effect of adverse selection in Akerlof's (1970) classic lemons market. In that model, there are always gains from trade, yet the buyer and seller may fail to trade due to the presence of seller types with whom the buyer prefers not to transact. Similarly, in our model, despite perfect interest alignment between the receiver and most sender types, the players fail to realize full gains from interaction due to the presence of some sender types with which the receiver prefers not to interact. Of course, the two frameworks differ along other dimensions.

Two distinctive features of our model are the following: (i) the receiver disagrees with some sender types over the payoff from not interacting, but conditional on interacting, there is no conflict of interest, and (ii) the gains from interaction vary with the state. Similar features also appear in Che, Dessein, and Kartik (2013), who consider a model of multidimensional, comparative cheap talk, where a sender communicates the value of a number of projects, among which the receiver can choose at most one. Both players agree on which project is preferred, but they disagree on the value from choosing neither. The authors show that the equilibrium depends on the value of not choosing any project: when it is low, there is a truthful equilibrium, when it is intermediate, the sender recommends an inferior project, which is chosen with positive probability, and when it is high, no project is chosen. By contrast, we analyze unidimensional cheap talk where coarse communication always exists, and is generated by an "infection-like" argument: the incentive of the toxic types to pool with the remaining types ruins the ability of all types to fully reveal their information.[3]

The loss of informative communication in our model is reminiscent of the loss that occurs in the literature on cheap talk with uncertainty over the sender's bias (notable examples include Morris (2001) and Morgan and Stocken (2003)). The reason coarse communication arises in these

[3]The second feature—i.e., state-dependent gains from interaction—is common in the comparative cheap-talk literature, pioneered by Chakraborty and Harbaugh (2007, 2010). This literature studies how a sender can truthfully communicate to a receiver the ranking of different options, even though the two players have conflicting interests.

papers is similar to that of Crawford and Sobel (1982) in the sense that the receiver believes that there is positive probability that the sender is biased. However, in our model, all sender types are "unbiased" in the sense that they agree with the receiver on the optimal action conditional on interaction. Another crucial distinction is that in these models, whenever the receiver gets a coarse message he is uncertain about the sender's bias. 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). However, the mechanisms of contagion in both of these literatures is very different from the one in our framework.

The rest of the paper is organized as follows. Section 2 introduces the general model. Our main result is described and proven in Section 3. Section 4 analyzes a canonical case with uniformly distributed types and Section 5 studies the effect of noisy monitoring on communication. Concluding remarks appear in Section 6.

2 Model

A sender privately draws a type θ from a distribution $F[0,1]$ with a differentiable density function f, which is strictly positive everywhere. We further assume that there exist $\underline{f}', \bar{f}' \in \mathbb{R}$ such that $\underline{f}' < f'(\theta) < \bar{f}'$ for all $\theta \in [0,1]$. 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 *null action.* 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, differentiable, and there exist $\underline{r}', \bar{r}' \in \mathbb{R}$ such that $\underline{r}' < r'(\theta) < \bar{r}'$ for all $\theta \in [0,1]$; and $L(\cdot)$ is increasing, continuous, strictly convex, 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[4] $\Lambda(\cdot)$. 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. This is a distinguishing feature of our model, which sets it apart from previous studies of communication failures. We assume that σ satisfies that each sender type strictly prefers any $a \in [0,1]$ to[5] $a = N$. 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. 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.

We analyze the perfect Bayesian Nash equilibrium of this game. Our aim is to show that even if the set of *toxic types* is arbitrarily small, in equilibrium, *all* agent types will fail to communicate their private information to the employer. The severity of the failure (i.e., the coarseness of the information transmission) will increase with the proportion of toxic types.

[4] 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 $a \in [0,1]$.

[5] Note that this assumption implies that any equilibrium will satisfy NITS condition of Chen, Kartik and Sobel (2008).

Due to the generality of our specification, in addition to *interval equilibria*, in which each equilibrium message is associated with a (perhaps degenerate) interval of types, there may exist equilibria in which the mapping between types and *non-null* actions (i.e., $a \neq N$) is non-monotonic. The non-monotonicity in the type-to-action mapping may only occur on information sets where N is optimal for the receiver, i.e., information sets that do not create value for the receiver.[6] Unlike interval equilibria where messages have appealing meaning (e.g., the quality is low/high), non-monotonic equilibria lack a natural interpretation and, moreover, the existence and structure of such equilibria depend on the fine details of the functions[7] $R(\cdot), r(\cdot), S(\cdot), s(\cdot)$, and $f(\cdot)$. In this paper, we focus on interval equilibria and refer to any such equilibrium simply as an "equilibrium."

Note that if it was common knowledge that $R(\theta) \geq \rho$, there would be a *fully revealing* equilibrium. As we show below, when a positive measure of toxic types exists, then no matter how small it is, communication is severely hampered: in any equilibrium, communication is discrete.

3 The main result

This section establishes our main result that 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.

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

In contrast to standard models of cheap talk, in each state the sender and receiver agree on the optimal action in [0,1]. The only source of disagreement is whether N is the worst or best action. While all sender types agree that N is the worst action, the receiver shares this view only when facing a sender above some threshold. For types below that threshold, N is the receiver's best action. Because of this distinctive feature, the main argument in the proof of Theorem 1 (given in

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

[7] In the Appendix, we illustrate a possible construction of such an equilibrium and offer a sufficient condition under which only interval equilibria exist.

Lemma 1 below) is substantially different from the standard method of proof as in Crawford and Sobel (1982). To lay the groundwork for our method of proof, we start with several preliminary definitions and observations.

For any $0 \leq x \leq y \leq 1$, let

$$V(x,y) = \max_{a \neq N} \int_x^y \left[R(\theta) - r(\theta)L(|\theta - a|)\right] \frac{f(\theta)}{F(y) - F(x)} d\theta$$

denote the receiver's expected payoff under the optimal action $a \neq N$, given the belief that $\theta \in [x,y]$. 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.[8]

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 non-null action following some message that is sent in equilibrium, then *every* type of the sender induces a non-null action with positive probability in that equilibrium. This leads to the following observation.

Observation 2. *In any equilibrium, the receiver's information set that contains $\theta = 0$ is a non-degenerate interval.*

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. In our case, Observation 2 follows from observation 1 because either all types induce N with certainty (in which case the interval is $[0,1]$) or an action in $[0,1]$ is a best response for the receiver (and so the length of the interval is positive since[9] $V(0,0) = 0 < \rho$). Taken together, our assumptions on $L(\cdot)$, the continuity of $f(\cdot)$ and $r(\cdot)$, and Observation 2 imply that the optimal

[8] In the next section, we show that this assumption does not guarantee the existence of equilibria in which non-null actions are played.

[9] 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.

non-null action at a non-degenerate interval $[x, y]$ is strictly below[10] y. 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 non-null actions are played. Then the receiver's information partition in that equilibrium contains a* non-degenerate *interval* $[y, z]$*.*

Clearly, if the intervals do not decrease in length as 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. The challenge in proving Theorem 1 is that in general, the intervals may *decrease* in length as the state increases. This requires us to come up with an argument, which shows that the intervals do not shrink too fast, thereby allowing the effect of toxic types to spread through the whole state space.

Assume by contradiction that there exists an equilibrium in which the receiver's information partition contains an infinite sequence of intervals $\{I_j\}_j^\infty$ in which $0 \in I_1$ and $\sup(I_j) = \inf(I_{j+1})$. Let η_j denote the length of I_j. Observations 2 and 3 imply that to prove Theorem 1 it suffices to prove the following result.

Lemma 1. *The series* $\sum_{j=1}^{\infty} \eta_j$ *diverges.*

Lemma 1 is proven in three steps. Step 1 sets an upper bound on the receiver's optimal action given a belief that the sender type is in some interval I. To derive this upper bound, we modify the receiver's optimization problem when the sender type is known to be in I such that the resulting solution is necessarily *higher* than the solution to the original optimization problem. In step 2, applying the indifference condition of a sender type that is on the boundary between two adjacent intervals, we use the upper bound on the receiver's optimal action established in step 1 to derive a *lower* bound on the ratio between the length of an interval and the length of the adjacent lower interval. In the third and final step, we construct a (sub)sequence of intervals from the receiver's equilibrium information partition. Using the lower bound on the ratio of two adjacent intervals, which we derived in the previous step, we show that the sum of interval lengths in this sequence must exceed one. Hence, there cannot exist an equilibrium in which the receiver's induced information

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

partition contains infinitely many intervals.

Proof of Lemma 1. The proof proceeds in three steps.

Step 1. Deriving an upper bound on the receiver's optimal response at an information set.

Let $I = [\underline{I}, \bar{I}] \subseteq [0,1]$ be an The (unique) action $a \neq N$ that attains $V(\underline{I}, \bar{I})$ solves

$$\min_{a \in I} \int_I L(|a-\theta|) r(\theta) f(\theta) d(\theta). \tag{1}$$

Let $-\infty < \underline{r} < \bar{r} < \infty$ 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

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

Hence, (1) can be rewritten as a problem where the receiver chooses $a \in I$ to minimize

$$\begin{aligned} &\underline{r}_I \underline{f}_I \int_I L(|a-\theta|) d\theta + \int_I L(|a-\theta|) \underline{r}_I (f(\theta) - \underline{f}_I) d\theta + \\ &+ \int_I L(|a-\theta|) \underline{f}_I \left(r(\theta) - \underline{r}_I\right) d\theta + \int_I L(|a-\theta|)(r(\theta) - \underline{r}_I)(f(\theta) - \underline{f}_I) d\theta. \end{aligned}$$

Denote the solution to this problem by $a^*(I)$. Observe that if the receiver's objective were only $\underline{r}_I \underline{f}_I \int_I L(|a-\theta|) d\theta$ (the first term in the above expression), then by the convexity of $L(\cdot)$, the optimal solution would be the midpoint $(\frac{\underline{I}+\bar{I}}{2})$ of the interval I. However, the remaining summands in the objective may push the action either below or above the midpoint. We now derive an upper bound on $a^*(I)$ by modifying each of the summands in the receiver's objective in a way that incentivizes him to choose a weakly higher action if he were minimizing each expression alone. Thus, the minimum of the sum of all modified components will be attained at an action weakly higher than $a^*(I)$.

Two forces incentivize the receiver to increase the action above $\frac{\underline{I}+\bar{I}}{2}$. First, if higher types are more likely (this is captured by changes in $f(\cdot)$), and second, if higher types are relatively "more important" (this is captured by changes in $r(\cdot)$). By assumption, the rate of change in density is at most $|\bar{f}'|$, while the rate of change in "importance" is at most $|\bar{r}'|$. We therefore modify the original objective function as follows:

$$\begin{array}{ccc} \textit{Original} & & \textit{Modified} \\ \underline{r}_I \underline{f}_I \int_I L(|a-\theta|)d\theta & \Rightarrow & \underline{r}\underline{f} \int_I L(|a-\theta|)d\theta \\ \int_I L(|a-\theta|)\underline{r}_I \left(f(\theta) - \underline{f}_I\right) d\theta & \Rightarrow & \bar{r} \int_I \left|\bar{f}'\right| (\theta - \underline{I})d\theta \cdot L(\bar{I} - a) \\ \int_I L(|a-\theta|)\underline{f}_I \left(r(\theta) - \underline{r}_I\right) d\theta & \Rightarrow & \bar{f} \int_I |\bar{r}'| (\theta - \underline{I})d\theta \cdot L(\bar{I} - a) \\ \int_I L(|a-\theta|)(r(\theta) - \underline{r}_I)(f(\theta) - \underline{f}_I)d\theta & \Rightarrow & |\bar{r}'| \left|\bar{f}'\right| \int_I (\theta - \underline{I})d\theta \cdot L(\bar{I} - a) \end{array}$$

The modified expressions have the following properties. First, as in the original problem, the first summand in the modified problem is also minimized at the midpoint of the interval. Second, each modified summand is minimized at the top boundary of the interval. Third, the weight on the first modified summand is weakly lower than the weight on the first original summand, while the weight on each of the remaining summands in the modified problem is higher than the corresponding weight in the original problem (this follows from observing that for each $\theta \in I$, the modified weight in the second, third, and fourth summands is higher than the corresponding weight in each of the original summands).[11]

It follows that the solution to the modified minimization problem,

$$\min_{a \in I} \ \underline{r}\underline{f} \int_I L(|a-\theta|)d\theta + (\bar{r}\left|\bar{f}'\right| + \bar{f}\,|\bar{r}'| + |\bar{r}'|\left|\bar{f}'\right|)L(\bar{I} - a) \int_I (\theta - \underline{I})d\theta,$$

is weakly higher than $a^*(I)$. Denote this solution by $a^{**}(I)$. Since $L(\cdot)$ is differentiable and convex, $a^{**}(I)$ can be derived from the first-order condition with respect to a :

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

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

[11] This is less obvious for the fourth summand. To see that this property is still true, notice that for each θ,

$$(r(\theta) - \underline{r}_I)(f(\theta) - \underline{f}_I) \leq |\bar{r}'|\,(\theta - \underline{I})\left|\bar{f}'\right|(\theta - \underline{I}),$$

and that $(\theta - \underline{I}) \leq 1$.

$Z \in [L'(\bar{I} - a^{**}(I)), L'(a^{**}(I) - \underline{I})]$ 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, (2) can be rewritten as

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

and since $Z \geq L'(\bar{I} - a^{**}(I))$, denoting

$$\gamma(I) \equiv \frac{(\bar{r}\left|\bar{f}'\right| + \bar{f}\left|\bar{r}'\right| + \left|\bar{r}'\right|\left|\bar{f}'\right|)}{4\underline{r}\underline{f}} \cdot |I|^2, \tag{3}$$

we obtain the upper bound, $a^*(I) \leq a^{**}(I) \leq \frac{\underline{I}+\bar{I}}{2} + \gamma(I)$. $\|$

In general, it is not guaranteed that the upper bound $a^{**}(I)$ falls inside I. However, observe that $\gamma(I)$ can be made arbitrarily small relative to $|I|$ by considering a sufficiently short interval I. We refer to an interval for which $\gamma(I) < \frac{|I|}{4}$ as a *short interval.*

To proceed with the proof of Lemma 1, it is sufficient to restrict attention to situations where all members of the receiver's information partition are short intervals.[12]

Step 2. Deriving a lower bound on the ratio of two adjacent intervals in the receiver's information partition.

Let $[x, y]$ and $[y, z]$ be two adjacent intervals in the receiver's equilibrium information partition such that $y - x \geq z - y$. The indifference condition of the threshold type y implies

$$y - a^*([x, y]) = a^*([y, z]) - y.$$

Since $a^*([x, y]) \leq \frac{x+y}{2} + \gamma([x, y])$, we have $a^*([y, z]) \geq y + \frac{y-x}{2} - \gamma([x, y])$. Combining this inequality

[12]The lemma holds trivially if the intervals do not become arbitrarily small eventually.

with $a^*([y,z]) \leq \frac{y+z}{2} + \gamma([y,z])$, we obtain $\frac{y+z}{2} + \gamma([y,z]) \geq y + \frac{y-x}{2} - \gamma([x,y])$, which gives

$$z - y \geq y - x - 2\gamma([x,y]) - 2\gamma([y,z]).$$

Dividing both sides of this inequality by $(y-x)$ yields

$$\frac{z-y}{y-x} \geq 1 - \frac{2\gamma([x,y])}{y-x} - \frac{2\gamma([y,z])}{y-x} \geq 1 - \frac{2\gamma([x,y])}{y-x} - \frac{2\gamma([y,z])}{z-y} \geq \; 1 - \frac{4\gamma([x,y])}{y-x}, \tag{4}$$

where the middle inequality follows from $y - x \geq z - y$ and the last inequality holds because $\frac{\gamma(I)}{|I|}$ increases in $|I|$. ‖

Step 3. Constructing a subsequence of equilibrium intervals with a divergent series of lengths.

Assume by contradiction that $\sum_{j=1}^{\infty} \eta_j \leq 1$. From the sequence $\{\eta_j\}_{j=1}^{\infty}$ we construct a *decreasing* subsequence of short intervals $\{\zeta_k\}_{k=1}^{\infty}$ as follows. Let $\zeta_1 = \eta_{j_1}$, where j_1 is the smallest integer such that I_j is a short interval and $\eta_j < \eta_{j_1}$ for all $j > j_1$. Such a j_1 exists since otherwise $\sum_{j=1}^{\infty} \eta_j$ diverges. Similarly, let $\zeta_2 = \eta_{j_2}$, where $j_2 > j_1$ is the smallest integer such that $\eta_j < \eta_{j_2}$ for all $j > j_2$. Continuing inductively in the same manner we define $\zeta_n = \eta_{j_n}$, where $j_n > j_{n-1}$ is the smallest integer such that $\eta_j < \eta_{j_n}$ for all $j > j_n$. Note that $\sum_{i=1}^{\infty} \zeta_i \leq 1$ (as a subseries of a convergent series of positive numbers). By construction, $\{\zeta_n\}_{n=1}^{\infty}$ is a monotonically decreasing sequence where $\zeta_{i+1} \geq \eta_{j_i+1}$. Thus, (4) implies $\frac{\zeta_{i+1}}{\zeta_i} \geq 1 - \frac{4\gamma(I_{j_i})}{\zeta_i}$ or, equivalently,[13]

$$\zeta_{i+1} \geq \left[1 - \frac{4\gamma(I_{j_i})}{\zeta_i}\right] \cdot \zeta_i.$$

As $\frac{\gamma(I)}{|I|}$ is increasing in $|I|$, the ratio $\delta_i = \left[1 - \frac{4\gamma(I_{j_i})}{\zeta_i}\right]$ increases when i increases because ζ_i decreases. Hence, by comparing $\sum_{i=k}^{\infty} \zeta_i$ to a geometric series whose first element is ζ_k and the *common* ratio is δ_k, it follows that, for any $k \in \mathbb{N}$,

$$\sum_{i=1}^{\infty} \zeta_i = \sum_{i=1}^{k-1} \zeta_i + \sum_{i=k}^{\infty} \zeta_i \geq \sum_{i=1}^{k-1} \zeta_i + \frac{\zeta_k}{1-\delta_k} = \sum_{i=1}^{k-1} \zeta_i + \frac{\zeta_k^2}{4\gamma(I_{j_k})}. \tag{5}$$

[13]To see why, recall that by definition, ζ_i is the length of the interval I_{j_i}, and that either ζ_{i+1} and ζ_{i+1} are the lengths of two adjacent intervals, or ζ_{i+1} is higher than the length of the interval adjacent to I_{j_i}.

By definition, $\lim_{k\to\infty}\sum_{i=1}^{k-1}\zeta_i = \sum_{i=1}^{\infty}\zeta_i$. However, by (3), $\frac{\zeta_k^2}{4\gamma(I_{j_k})}$ is positive and does not depend on k, a contradiction. $\square$

We have therefore established that in equilibrium, the receiver's information partition cannot consist of infinitely many intervals. This completes the proof of Theorem 1.

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

Example 1. *Suppose that when the receiver chooses $a \neq N$, 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 $a(x,y) = \frac{1}{3}x + \frac{2}{3}y$. 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[14] $\frac{1}{2}(y-x)$.

For example, if $\rho = \frac{25}{216}$, there exists 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 $\frac{1}{2}$.

[14]Denote the right-adjacent interval by $[y,z]$. Thus, $a(y,z) = \frac{1}{3}y + \frac{2}{3}z$. 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 $\frac{z-y}{y-x} = \frac{1}{2}$.

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.*

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 will focus on the case in which the distribution of types is uniform and the sender's type θ represents the value he can generate 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 action N 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.

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[15] $L(\cdot)$. 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.

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

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. In equilibrium, any message is associated with some interval of types that send this message. 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

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

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

$$V(x,y) = \frac{1}{y-x}\int_x^y \theta d\theta - \frac{1}{y-x}\int_x^y L(|\frac{x+y}{2}-\theta|)d\theta = \frac{x+y}{2} - \bar{\lambda}_{y-x}. \tag{6}$$

Since $L(\cdot)$ is increasing and strictly convex, the function $\bar{\lambda}_\delta$ 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, as a sum of two concave functions. 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 $E_{noN}=\{Q_K \ : \frac{1}{K}\geq\theta_\rho\}$.*

Proof. By (6), 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 2. *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 $\bar{\lambda}_\delta$ is an increasing function of δ,

$$V(y,z) = \frac{y+z}{2} - \bar{\lambda}_{z-y} > \frac{x+y}{2} - \bar{\lambda}_{z-y} \geq \frac{x+y}{2} - \bar{\lambda}_{y-x} = V(x,y). \tag{7}$$

Assume first that the receiver mixes on $[y, z]$. By (7), $\rho = V(y, z) > V(x, y)$. 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, (7) 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 (2) 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. $\square$

Provided that N is not selected, the players' interests coincide. The expected payoff from Q_K equals $\frac{1}{K}\sum_{k=1}^{K} V(\frac{k-1}{K}, \frac{k}{K}) = \frac{1}{2} - \bar{\lambda}_{\frac{1}{K}}$. Since $\bar{\lambda}_{\frac{1}{K}}$ 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 $V(0, \frac{1}{M^*}) \geq \rho$. Under the Pareto efficient equilibrium, the receiver never plays N.*

The presence of undesirable sender types—even when their measure is arbitrarily small—prevents *all* sender types from disclosing their type to the receiver, despite the fact that such disclosure would be mutually beneficial for all types above ρ. However, as the measure of undesirable types falls, the precision of the communication rises. The following example illustrates the strong effect of a small proportion of toxic types.

Example 2. *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 $y \leq \frac{2}{10}$, $V(0, y) = \frac{y}{2} - \bar{\lambda}_y < \frac{y}{2} \leq \frac{1}{10} = \rho$. 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—tha 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 3. *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*[16]

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

and $z > 0$ is a constant. It is easy to verify that $V(0,y)$ is increasing in y, for all $y < \frac{3}{4}$. Since $z > 0$, for all $y > \frac{3}{4}$, we have $4e^{z(\theta - \frac{3}{4})^2} > 4$. Thus,

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

with strict inequality if and only if $y > \frac{3}{4}$. The expression on the RHS, and therefore also $V(0,y)$, is maximized at $\theta = \frac{3}{4}$, where its value is $\frac{3}{16}$. Let $V(\frac{3}{4},1)(z)$ denote the value of $V(\frac{3}{4},1)$ as a function of z. $V(\frac{3}{4},1)(z)$ 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 $V(\frac{3}{4},1)(z^*) = V(0,\frac{3}{4})$. 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

[16]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, while the exponential form is convenient to guarantee that $r(\cdot)$ is differentiable.

the intervals implies that $q_l < q_r$.

Clearly, the receiver's payoff from any equilibrium that partitions the unit interval around $\frac{3}{4}$ *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 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 Communication breakdown with noisy monitoring

In this section we explore how communication is impacted when the receiver observes noisy and infrequent signals on the sender's type. On the one hand, the ability to obtain some verifiable information on the receiver's type may allow the sender to end his relations with the toxic types. On the other hand, the noise in the sender's signal may give further incentives for the toxic types (who are under risk of being excluded from interacting with the receiver) to lie. This may deteriorate even more the communication between the receiver and the desirable types. For this disrupted communication to have a significant effect, the opportunities to observe the sender's types should be relatively infrequent.

To analyze the effect of noisy monitoring on communication, we consider the following two-period extension of our model. First, when the sender's type is θ, and the receiver chooses $a \in [0,1]$, both players' payoff is $\theta - (\theta - a)^2$. When the receiver chooses $a = N$, the sender's and receiver's payoffs are, respectively, $\sigma < -1$ and $\rho \in (0, V(0,1))$.

Second, the players play the following two-period game. At the beginning of the first period, the sender privately observes his type θ, which is drawn from a uniform distribution on $[0,1]$, after which he sends a message $m(\theta)$. The receiver then chooses $a_1(m) \in [0,1] \cup \{N\}$. If $a_1 = N$, the game ends and the players receive payoffs. Otherwise, the game proceeds to the second period, in which the receiver observes a signal t about the sender's type. With probability $p \leq \frac{1}{2}$, the signal is a random draw, uncorrelated with the sender's type, from a uniform distribution on $[0,1]$. With the remaining probability, $1-p$, the signal realization is equal to the sender's type. After observing the signal realization, the receiver chooses $a_2(m,t)$, the game ends, and the players receive payoffs.

Our objective is to derive an upper bound on the number of intervals in any (interval) equilibrium (or equivalently, an upper bound on the number of distinct actions that are chosen with positive probability). We shall show that any such equilibrium is either babbling or induces only two distinct actions.

We start by characterizing the receiver's optimal behavior following the report that $\theta \in [x,y]$, under the restriction that the receiver does not choose N. Prior to observing a signal, the receiver chooses $a_1 = \frac{x+y}{2}$. To characterize a_2, let t denote the signal realization. With probability $p(1-(y-x))$, the signal realization satisfies $t \notin [x,y]$, in which case the receiver infers that he observed noise and chooses $a_2 = a_1 = \frac{x+y}{2}$. With probability $(1-p)+p(y-x)$, the signal realization falls within $[x,y]$, in which case the receiver assigns probability

$$q = \frac{p(y-x)}{(1-p)+p(y-x)}$$

to the event that he observed noise. Conditional on $t \in [x,y]$, the receiver maximizes

$$q\int_x^y (\theta - (\theta-a)^2)\frac{1}{y-x}d\theta + (1-q)(t-(t-a)^2), \tag{8}$$

which is equivalent to minimizing the expected loss, $q\int_x^y (\theta-a)^2 \frac{1}{y-x}d\theta + (1-q)(t-a)^2$. The minimum of the expected quadratic loss function is attained at the expected value. Hence,

$$a_2 = \mathbb{E}[\theta|\theta \in [x,y], t] = q\frac{x+y}{2} + (1-q)t. \tag{9}$$

The next lemma establishes that N may be selected in equilibrium only on the leftmost interval. The argument has two parts. First, in order to make $a_1 \neq N$ optimal on that interval, it must contain some viable types. Therefore, every other interval consists of only viable types. The second part of the argument shows that under the assumed loss function, the action N is suboptimal in both periods on any interval that contains only viable types.

Lemma 3. *In any interval equilibrium, N is never selected outside of the leftmost interval.*

Proof. The action N is uniquely optimal for all $\theta < \rho$ and hence, if $y < \rho$, N is uniquely optimal on $[0, y]$ given any signal realization. A monotonic information partition (i.e., a partition of $[0, 1]$ into intervals) where the length of the leftmost interval is below ρ cannot be consistent with an equilibrium because sender types below ρ would have a profitable deviation to a report that induces $a_1 \neq N$, and such a report exists because $\rho < V(0, 1) < V(x, 1)$ for all $x > 0$.

$V(0, y)$ is strictly increasing for all $y \in (0, 1]$. Therefore, $V(x, y) > x$ for all $y > x$, which guarantees that N is suboptimal whenever $\theta \sim U[x, y]$ for $\rho \leq x < y \leq 1$. Therefore, the receiver chooses $a_1 \neq N$ and $a_2 \neq N$ if the signal realization $t \notin [x, y]$ (in which case he infers that it is noise). If $t \in [x, y]$, we have

$$\max_a \; q \int_x^y (\theta - (\theta - a)^2) \frac{1}{y - x} d\theta + (1 - q)(t - (t - a)^2) >$$

$$q \int_x^y (\theta - (\theta - x)^2) \frac{1}{y - x} d\theta + (1 - q)(t - (t - x)^2) > x \geq \rho,$$

and therefore $a_2 \neq N$ on any interval above ρ. □

An important step in characterizing the class of interval equilibria is understanding the relation between the lengths of the intervals, which capture the "precision" of the messages sent by types in those intervals. The previous section established that in a *one-shot interaction* with linear values and uniform types, any pair of intervals on which the action N is never played must have equal lengths. The question is whether this feature continues to hold when we add a second period of interaction in which the receiver observes a signal on the sender's type. The added difficulty is that now the sender's report induces a *distribution* of actions in period 2. This distribution is affected by the receiver's interpretation of signals, which changes with the length of the interval reported by the sender.[17] Hence, in order to understand whether a sender type, that lies on the boundary between two adjacent intervals of different lengths prefers the longer interval, we need to compare two distributions overs the receiver's actions.

To do this, we introduce the following notation. Suppose that before observing a signal realization the receiver believes $\theta \sim U[y,z]$. Let $f_{[y,z]}$ denote the distribution over the receiver's actions (in period 2) conditional on $t \neq \theta$, i.e., conditional on the signal realization being noise. Note that $f_{[y,z]}$ is symmetric around $\frac{y+z}{2}$ where the distribution has an atom.

Lemma 4. *If $y < z_1 < z_2$, then $f_{[y,z_2]}$ FOSD $f_{[y,z_1]}$.*

Proof. For notational convenience, we set $y = 0$ ($a = N$ is ruled out and so this is just a normalization that simplifies expressions). By (9), the support of $f_{[0,z]}$ is $\left[q \cdot \frac{z}{2}\ ,\ q \cdot \frac{z}{2} + (1-q)z\right]$, where $q = \frac{pz}{(1-p)+pz}$. The probability at the atom $\frac{z}{2}$ is $1 - z$, and the density elsewhere is uniform. If we increase the value of z, the location of the atom, $\frac{z}{2}$, shifts to the right, and the probability assigned to the atom, $1 - z$, decreases. Hence, $F_{[0,z_1]}(\frac{z_1}{2}) = 1 - \frac{z_1}{2} > 1 - \frac{z_2}{2} = F_{[0,z_2]}(\frac{z_2}{2})$, for $0 < z_1 < z_2$. In addition, since $\frac{\partial q}{\partial z} > 0$, the lower bound of the support of $f_{[0,z]}$ increases with z. Therefore, for any $\eta < \frac{z_2}{2}$, $F_{[0,z_1]}(\eta) \geq F_{[0,z_2]}(\eta)$. Finally, $\frac{\partial[q \cdot \frac{z}{2} + (1-q)z]}{\partial z} = \frac{(p-1)^2}{2(pz-p+1)^2} + \frac{1}{2} > 0$, and so the upper bound of the support of $f_{[0,z]}$ increases with z. Therefore, for any $\eta > \frac{z_2}{2}$, $F_{[0,z_1]}(\eta) \geq F_{[0,z_2]}(\eta)$. □

Using this lemma we can now establish that types in the lowest interval send the most precise message, while the messages of all higher types are less precise.

Proposition 4. *In any equilibrium, the leftmost interval is the shortest and the remaining intervals are of equal length.*

Proof. Let $x < y < z$ be the thresholds of adjacent intervals of an equilibrium partition. Consider first the case where $x > 0$. By Lemma (3), the receiver chooses $a \neq N$ in both periods (before and after observing a signal realization t). Consider the expected payoff of type $\theta = y$ as a function of z. Clearly, the sender's payoff in period 1 (i.e., prior to signal realization) is higher when the interval

[17]Signals contain "less information" on larger intervals because detecting noise on such intervals is more difficult. This might be beneficial to the sender since it protects him from extreme actions in case of noise.

is shorter. To prove that his payoff monotonically decreases when z increases we now show this for the second period payoff.

With probability p, the signal generates uninformative noise, and by Lemma 4, the sender's payoff decreases when z increases in this case. With probability $1-p$, the signal realization is $t=y$, in which case the receiver chooses $a_2 = q\frac{y+z}{2} + (1-q)y$, where $q = \frac{p(z-y)}{(1-p)+p(z-y)}$. Since $\frac{\partial q}{\partial z} > 0$, the induced receiver's action increases with z and thus the sender's payoff decreases in this case as well. Hence, the payoff of type y from reporting that $\theta \in [y,z]$ is strictly decreasing in z.

When type y reports that $\theta \in [y,z]$, he induces a distribution of actions that are all above y. The sender does not care about the distribution of actions per se, but rather about the distribution of *distances* between his type and the selected action. Exactly the same distribution of distances between y and the receiver's action is obtained when the sender's report induces the receiver's belief that $\theta \in [x,y]$ when $y-x=z-y$. Since the payoff of type $\theta = y$ from reporting that $\theta \in [y,z]$ is monotonically decreasing in z, all intervals on which the receiver never selects N must be of equal length.

Finally, consider the case where $x=0$. Assume that the sender reports that $\theta \in [0,y]$. Let $a(t) \in [0,y] \cup \{N\}$ denote the receiver's optimal action when the signal realization is t, and let $\tilde{a}(t) \in [0,y]$ denote the receiver's optimal action when he is restricted from playing N. Observe that whenever $a(t) \neq N$, $a(t) = \tilde{a}(t)$. Therefore, since $\sigma < -1$, from the perspective of the sender, $a(t)$ is identical to or strictly worse than $\tilde{a}(t)$, for any realization t. By Lemma (3), the receiver never chooses N when $[y,z]$ is reported. Therefore, from the argument given earlier in the proof, it follows that *if* the receiver played according to $\tilde{a}(t)$, then type $\theta = y$ would have been indifferent between reporting $[0,y]$ and $[y,z]$ if and only if the lengths of these intervals were equal. Hence, if, conditional on reporting $[0,y]$, type $\theta = y$ assigns positive probability to signal realizations after which the receiver chooses N, the sender can be indifferent between reporting $[0,y]$ and $[y,z]$ only if $y < z-y$. □

We are now ready to provide the main result of this section, which shows that noisy monitoring has a dramatic effect on the senders' ability to communicate their type to the receiver.

Proposition 5. *Any interval equilibrium in which action N is played partitions the unit interval* *into at most two intervals.*

Proof. Assume by contradiction that there are at least three intervals. Let $\underline{a}$ denote the receiver's action on the lowest interval prior to observing a signal (period 1). Consider type 0. By reporting that θ belongs to the lowest interval, the sender receives a payoff of at most $-\underline{a}^2$ in period 1. In period 2, with probability $1-p$ the signal realization is $t = \theta = 0$, in which case the receiver chooses[18] N. With probability p there is noise that induces a distribution over actions, which may assign positive probability to action N. Denote this distribution by g. Let g' be the distribution over actions that is obtained when the signal realization is noise, but the sender is restricted to choosing an action in [0,1]. This distribution is symmetric around $\underline{a}$ and dominates the action distribution g. Due to his risk aversion, the sender prefers to induce $\underline{a}$ with certainty to the symmetric distribution g', and therefore the sender's payoff at state $\theta = 0$ from "truth-telling" can be bounded from above by

$$-(1+p)\underline{a}^2 - (1-p). \tag{10}$$

Let a denote the second-lowest action in the support of the receiver's strategy in period 1 (i.e., a is the mid-point of the second-lowest interval of the equilibrium partition). Observe that if $a > \frac{1}{2}$, then more than half of the second-lowest interval lies above $\frac{1}{2}$, and therefore the *third*-lowest interval is shorter than the lowest interval, in contradiction to Proposition (4). Thus, a necessary condition for having at least three intervals in the equilibrium partition is $a \leq \frac{1}{2}$. Clearly, type $\theta = 0$ prefers to report the second-lowest interval to any other interval on which the receiver plays higher actions. Next we show that the payoff of type 0 from deviating is bounded from below by

$$-a^2 + \big[-(1-p)a^2 - p(1 - \frac{1-2\underline{a}}{2})a^2 - p \cdot \frac{1-2\underline{a}}{2} \int_{2\underline{a}}^{2\underline{a}+\frac{1}{2}(1-2\underline{a})} \alpha^2 \cdot \frac{2}{1-2\underline{a}} d\alpha\big]. \tag{11}$$

First, since the sender's payoff from deviating increases when the number of (equal) intervals above $[0, 2\underline{a}]$ increases (because this brings the induced actions closer together), (11) is written as if there are exactly three intervals in the putative equilibrium partition. The first term, $-a^2$, is the

[18]Since N is played in equilibrium, it must be played when the receiver has the lowest expectation of the payoff from interacting with the sender. This occurs when the sender reports the lowest interval and the period-2 signal is 0.

sender's payoff in period 1, and the expression in the square brackets is a lower bound on his payoff in period 2: with probability $(1-p)$, the signal realization is informative but interpreted as noise outside of the reported interval (because of the deviation), and with probability $p(1-\frac{1-2\underline{a}}{2})$, there is noise, and the signal realization falls outside of the reported interval. In both of these cases, which correspond to the first two terms in square brackets, the receiver chooses a. Finally, with probability $p\frac{1-2\underline{a}}{2}$, the signal realization is noise that is consistent with the reported interval. Since in this case t is uniform on the reported interval, and by (9) the action is linear in t, a uniform distribution of actions is induced. Moreover, (9) also implies that the support of this distribution is contained within the second-lowest interval and is symmetric around a. Due to the convexity of the loss function, the worst distribution in this family from the perspective of type 0 is one whose support is the whole interval.

In equilibrium, (10) must be greater than (11), which, after some rearrangement, becomes

$$-(1+p)\underline{a}^2-(1-p)+(2-p)a^2+p(\underline{a}+\frac{1}{2})a^2+p\cdot\frac{1}{3}((\underline{a}+\frac{1}{2})^3-8\underline{a}^3)\geq 0. \tag{12}$$

Denote the expression on the LHS of (12) by $Z(\underline{a},a,p)$. Direct inspection reveals that $\frac{\partial Z}{\partial p}>0$, for all $0<\underline{a}<a<\frac{1}{2}$. Hence, a necessary condition for equilibrium is

$$Z(\underline{a},a,p=\frac{1}{2})\geq 0.$$

However, this inequality can hold only if $a>\frac{1}{2}$, a contradiction. □

Next we show that the possibility of free additional information during a dynamic interaction may harm the principal due to the endogenous effect on voluntary communication in the initial stage.

Example 4. *Let $\rho=\frac{1}{10}$. To have a benchmark with no monitoring, consider a two-period interaction that begins with the worker reporting his type, after which the principal assigns a task to the agent twice (it is suboptimal to assign the worker two distinct tasks in different periods since the receiver does not learn anything between the two periods). As in Example 2, there exists an equilibrium in which the receiver's information partition consists of four equal intervals (recall that we denote this partition by Q_4). The receiver's expected payoff under this no-monitoring benchmark* *is given by*

$$W_{NM} := 2\Big[\mathbb{E}[\theta] - \mathbb{E}[Var[\theta|Q_4]]\Big] = 1 - \frac{1}{96}.$$

Now consider the case where, between the periods, the receiver observes a signal as we specified at the beginning of the present section. By Proposition 5, the receiver's information partition in period 1 consists of at most two intervals. Thus, the expected payoff in period 1 is bounded from above by the (one-period) expected payoff from Q_2*. Clearly, the receiver's expected payoff in period 2 is bounded from above by the expected payoff under full information. Hence, denoting by* W_M *the receiver's expected payoff with monitoring, we obtain*

$$W_M < \underbrace{\mathbb{E}[\theta] - \mathbb{E}[Var[\theta|Q_2]]}_{\textit{period-1 upper bound}} + \underbrace{\int_0^{\rho} max\{\theta, \rho\} d\theta}_{\textit{period-2 upper bound}} = [\frac{1}{2} - \frac{1}{48}] + [\frac{1}{2} + \frac{1}{200}] < W_{NM}.$$

Hence, the benefits from monitoring are more than offset by the reduced quality of communication and thus, ex ante, the receiver is better off without monitoring. This holds for any p whenever, under the no-monitoring benchmark, the most informative equilibrium has at least 4 intervals.

6 Concluding remarks

This paper analyzed a fairly common scenario 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 bilateral 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. Moreover, we demonstrated that this communication can deteriorate even further when the uninformed party that chooses the action gets noisy observations on the type of the other party.

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] Akerlof, George (1970). "The market for "lemons": Quality uncertainty and the market mechanism." *Quarterly Journal of Economics* 84(3): 488–500.

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

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

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

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

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

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

[14] 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 non-null 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-null 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 (13) 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,

$$-s(\theta - a_l)^2 + [(1 - q)s(\theta - a_r)^2 + q\sigma], \tag{13}$$

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-null 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.

A2. Sufficient condition for existence of only interval equilibria

We now show that if $S(\theta)$ is a constant, 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.