Persuasion without Priors*

Alexei Parakhonyak† Anton Sobolev‡

September 21, 2022

Abstract

We consider an information design problem when the sender faces ambiguity regarding the probability distribution over the states of the world, the utility function and the prior of the receiver. The solution concept is minimax loss (regret) – that is, the sender minimizes the distance from the full information benchmark in the worst-case scenario. We show that in the binary states and binary actions setting the optimal strategy involves a mechanism with continuum of messages, which admits a representation as a randomization over mechanisms consisting of two messages. Small uncertainty regarding the receiver's prior makes the sender more truthful than in the full information benchmark, but as uncertainty increases at some point the sender starts lying more. If the sender either knows the probability distribution over the states of the world, or knows that the receiver knows it, then the maximal loss is bounded from above by $1/e$. This result generalizes to an infinite state model, provided that the set of admissible mechanisms is limited to cut-off strategies.

Keywords: Persuasion, Robustness, Multiple priors, Minimax regret

JEL-classification: D81, D82, D83

*We are grateful to Piotr Dworczak, Ludwig Sinander, Karl Schlag, Nicolas Schutz, Thomas Tröger, Karolina Vocke, and various seminar audiences for useful comments. The support by the Deutsche Forschungsgemeinschaft through PE 813/2-2 and CRC TR 224 (projects B03 and B04) is gratefully acknowledged.

†University of Oxford, Department of Economics and Lincoln College. E-Mail: aleksei.parakhonyak@economics.ox.ac.uk.

‡University of Mannheim, Department of Economics and MaCCI. E-Mail: anton.sobolev@uni-mannheim.de.

1 Introduction

The standard Bayesian persuasion framework, pioneered by Kamenica and Gentzkow (2011), imposes strong informational assumptions on the sender. In order to design the optimal mechanism, the sender must know the probability distribution over the states of the world, the receiver's prior and her utility function. Arguably, such assumptions are questionable in many practical settings. For example, an auto seller who designs a test-drive procedure might not know what the buyer's prior thoughts about a car are or which offers the buyer already has. A prosecutor might not be aware of the judge's prior belief about the probability of the defendant being guilty, or of the judge's preferences for acquitting a guilty person versus sentencing an innocent one.

In this paper, we study an environment in which the sender potentially faces uncertainty about all three inputs of the optimal persuasion mechanism. Moreover, the sender does not have any prior assumption over the distribution of these inputs, for example, the receiver's priors. Such an uninformed sender evaluates the policy based on *loss* or regret, i.e. the difference between the payoff from the current persuasion mechanism and what the sender could have obtained under complete information. We use the term 'complete information' in an ex ante sense, i.e. the completely informed sender knows all the parameters of the model, but not the future realisation of the state of the world. The sender is interested in minimising loss in the *worst-case scenario*, i.e. loss that could occur over all possible distribution over the states, priors and utility.

The receiver in our game is a standard Bayesian. That is, the receiver observes the choice of the mechanism by the sender, the realisation of the signal produced by this mechanism, and combines this information with her prior to choose the optimal action.

The main part of our analysis focuses on the binary state (good or bad), and binary action (adopt or reject) problem. There is a clear ranking about the states, but the sender does not know the probability of the good state, the receiver's prior about the probability of the good state and the outside option of the receiver. We show that the receiver's parameters can be combined to a single integral characteristic – the receiver's optimism. We solve for the optimal persuasion mechanism by representing the original problem as a zero-sum game between the sender and adversarial nature. The sender chooses a persuasion strategy and wants to minimise loss, while nature chooses the probability of the good state and the receiver's optimism and tries to maximise loss. As is standard for zero-sum games, our game has an equilibrium only in mixed strategies.

The zero-sum game formulation of the problem allows us to represent the optimal mechanism, which consists of continuum of messages, as randomisation over standard mechanisms, where an adoption recommendation is sent in both states, and a rejection recommendation is sent in the bad state only. The strategy of adversarial nature involves a negative correlation between the receiver's optimism and the true probability of being in the good state. That is, the worst-case scenarios are those when having a high probability of being in a good state is coupled with a receiver who is hard to persuade, either because of a low prior or a high outside option and vice versa. The intuition is that lying too much, which is more likely to happen when a facing pessimistic receiver, is particularly costly when the good state is very likely. Similarly, when the bad state is likely it is costly to be too truthful, i.e. too often revealing the good state. This is more likely to happen when facing an optimistic receiver, who could be persuaded by less truthful strategy.

The minimax loss is increasing in the uncertainty over parameters and is bounded from above by the omega constant.[1] When the sender faces no uncertainty over the receiver's optimism, the minimax loss equals to zero. If the sender knows the receiver's optimism, he can maximise the probability of persuasion in the bad state, even though the probability of this state may be unknown to the sender. When the sender faces maximal uncertainty over the receiver's optimism but knows the probability of the good state, the maximal loss is bounded above by $1/e$, but can be lower if the good state is relatively likely. These two

[1] See `https://en.wikipedia.org/wiki/Omega_constant`

results highlight the fact that information about the receiver is crucially more important than information about the probability distribution over the states.

The equilibrium mixed strategy of the sender in our zero-sum game is equivalent to playing the optimal mechanism which involves infinitely many messages, one message fully revealing that the state is bad, and each of the other messages persuading a receiver with optimism above a certain threshold. Any such message is equivalent to a choice of a single two-message mechanism, designed to persuade a specific receiver, from the support of the sender's mixed strategy.

We investigate how the probability of lying, i.e. not sending a revealing message in the bad state, depends on uncertainty about the receiver's optimism. We show that initially this probability decreases in uncertainty, i.e. when uncertainty is small, the sender employs a more truthful strategy than in the case of complete information. However, as the level of uncertainty increases, at some point the sender starts lying more, potentially becoming less truthful than in the full information case.

We extend our model by considering the 'informed receiver' setting, i.e. the situation in which the receiver's prior coincides with the true probability distribution over the states of the world. This imposes a constraint on the strategy of adversarial nature, as it no longer can introduce negative correlation between the receiver's optimism and the probability of the good state. In this case the loss of the sender is bounded by $1/e$. Therefore, the maximal loss of the sender is the same irrespective of whether he knows the probability distribution over the states, or knows that the receiver knows it.

Finally, we extend our analysis to multiple states of the world. We assume that the receiver's utility is monotone and that the only uncertainty the sender faces is about the receiver's prior and her outside option. We show that the sender's loss approaches its maximal possible value of one as the number of states approaches infinity. This negative result, however, relies on a peculiar optimal mechanism, in which the complete information sender would find it optimal to exclude specific states, randomly chosen by adversarial nature, from the support of adoption recommendation. If the sender (as well as the complete information benchmark) is restricted to cut-off strategies, i.e. when the adoption recommendation is sent above a certain threshold, then the optimal mechanism is equivalent to the one we derived in the binary states model, resulting in similar strategies and identical losses.

Our paper extends techniques of robust decision-making to the Bayesian persuasion problem. The most popular criterion for decision making without priors, or robust decision making, is maximin utility. Under this criterion, the decision-maker, sender in our case, aims for the best payoff for a specific, worst-case, prior. In our setting the result is trivial: nature can choose a non-persuadable receiver resulting in zero payoff. The way around this problem is to introduce an extra constraint, for example, restrict priors to distributions with certain probability mass above or below a given point, see Kosterina (2022). In our approach, usually referred to as the minimax regret or minimax loss, the sender is concerned about being close to the optimum irrespective of the prior. An alternative interpretation of this approach is a setting in which the sender faces a receiver drawn from an unknown distribution and tries to be close on average to the sender who observes the receiver's type under the most disadvantageous distribution of receivers. This approach was applied to monopoly pricing (see Bergemann and Schlag (2008)), search (see Bergemann and Schlag (2011), Parakhonyak and Sobolev (2015), Schlag and Zapechelnyuk (2021), Schlag and Sobolev (2022)), monopoly regulation (see Guo and Shmaya (2019b)) and statistical treatment choice (see Manski (2004)).

Bayesian persuasion literature was pioneered by Kamenica and Gentzkow (2011) and since then experienced explosive growth, see Bergemann and Morris (2019) and Kamenica (2019) for extensive surveys of existing literature. Our paper is closely related to the literature with multiple receiver types (see Kolotilin et al. (2017), Laclau and Renou (2017), Guo and Shmaya (2019a), Parakhonyak and Vikander (2019), Best and Quigley (2020)). In our setting, unlike the aforementioned papers, the distribution of the receiver types is unknown to the sender.

Another key feature of our model is that the prior of the receiver may be different from the prior of the informed sender benchmark. Standard Bayesian persuasion setting with heterogeneous priors was first studied by Alonso and Camara (2016).

Literature on robust persuasion mainly utilises the maximin utility approach. Hu and Weng (2021) consider a setting in which nature chooses private information of the receiver. Similar to our result, they find that even in the binary setting the optimal mechanism may involve infinitely many messages, with some messages persuading a receiver with certain private beliefs, but failing to persuade if her private beliefs were different. In contrast to Hu and Weng (2021), our paper does not rely on the common prior assumption (and a restriction on private beliefs which follows from it) and uses minimax loss approach, rather than maximin utility. Dworczak and Pavan (2020) also rely on the common prior assumption in maximin utility framework, and look at the setting in which nature can send an extra signal to the receiver, conditional not only on the mechanism chosen by the sender but also on the sender's signal realisation. This gives adversarial nature substantially more power than in our model, where nature and the sender are engaged in a simultaneous move game. Kosterina (2022), similar to our paper, focuses of the problem of the sender who does not know the prior of the receiver. After restricting the set of receiver's priors to those having some probability mass above a certain threshold, she, similar to aforementioned papers, utilises the maximin utility approach which, essentially, replaces our simultaneous move game with a sequential game in which nature moves last. The optimal mechanism in Kosterina (2022) does not depend on the sender's prior. In our model, the sender does not have a prior about the distribution of states. However, even if the sender had such a prior (see Corollary 2), the choice of the optimal mechanism would depend on it, as well as the minimax loss.

Finally, the closest paper to ours is by Babichenko et al. (2021). They also apply the minimax loss/regret solution concept to a Bayesian persuasion problem. There are two key differences between our models. Firstly, in our model the utility function from adoption is fixed (up to the outside option) and monotone, while in Babichenko et al. (2021) nature is free to choose the receiver's utility function. Secondly, in their paper sender and receiver share a common prior, while in our model priors can differ and are not known to the sender. It turns out, that if the sender is sure about the distribution of states (i.e. has a point prior), the unknown receiver's prior assumption produces qualitatively similar results to the unknown utility assumption in Babichenko et al. (2021), see Section 4 of our paper. In particular, as the number of states increases and there are no restrictions on the prior (on utility in Babichenko et al. (2021)), loss in the worst-case scenario approaches one. If, however, the sender is restricted to cut-off strategies (utility is monotone in Babichenko et al. (2021)), the minimax loss is bounded by $1/e$. If the sender faces ambiguity regarding the distribution of states, as in our main setting, the analysis is substantially different from these two limiting cases and Babichenko et al. (2021).

The rest of the paper is organised as follows. Section 2 introduces the model in a binary setting, derives the informed sender benchmark, and sets up the loss function. Section 3 contains all main results for binary setting, including the informed receiver version of the model. Section 4 extends our results to a framework with multiple states. All the proofs are presented in the Appendices.

2 Binary Model and Informed Sender Benchmark

In this section we introduce necessary notation, the informed sender benchmark and set up the uninformed sender problem.

We consider a Bayesian persuasion problem with binary states and binary actions. Suppose that the set of possible states of the world is $\Omega = \{0, 1\}$. The receiver chooses an action $a \in \{0, 1\}$. In what follows we refer to $a = 1$ as adopting (e.g. purchasing the product) and

$a = 0$ as rejecting. Utility functions are given by

$$u_R = r + a(\omega - r), \qquad u_S = a,$$

where r is the outside option of the receiver.[2] That is, the receiver prefers to adopt when $\omega = 1$ and the sender prefers adoption in any state of the world.

We assume that the receiver has a prior belief that probability of $\omega = 1$ equals to β. The true probability of $\omega = 1$ might be different from the receiver's belief and is denoted by α. The sender and the receiver in our model may have different priors. The sender (he) commits to a persuasion mechanism. Formally, let M be a set of messages and $\mathcal{M}$ be a Borel sigma-algebra generated by this set. The sender's strategy is a mapping $\mu_M : \Omega \times \mathcal{M} \to [0, 1]$ such that for any $\omega \in \Omega$ and $B \in \mathcal{M}$, $\mu_\omega(B) \equiv \mu(B|\omega)$ is a probability measure. After observing the *mechanism* μ_M and a *message* $m \in M$ generated by this mechanism, the receiver (she), updates the prior in a Bayesian way and chooses an action $a \in \{0, 1\}$ which maximises her expected utility.

Consider a mechanism $\mu_M = \{\mu_1, \mu_0\}$. We define a measure μ_R (corresponding the posterior of a Bayesian receiver) such that for any $B \in \mathcal{M}$

$$\mu_R(B) = \beta\mu_1(B) + (1 - \beta)\mu_0(B).$$

Now we define a conditional probability (of $\omega = 1$ conditional on message m) as a random variable $P_\beta(m)$ such that for almost all $B \in \mathcal{M}$ we have

$$\int_B P_\beta(m) d\mu_R(m) = \beta\mu_1(B). \tag{1}$$

We define the acceptance state as a set of messages that result in a posterior sufficiently high

[2]Our linear utility specification, naturally, is not restrictive in the binary setting. We keep this specification for the multiple states extension, but all our results for multiple states, presented in Section 4, can be generalised for any strictly increasing utility function.

to justify adoption:

$$A = \{m \in M : P_\beta(m) \geq r\}. \tag{2}$$

Now we consider the informed sender benchmark. That is, the sender (i) has a prior about the probability of the good state α, and knows (ii) the prior of the receiver β and the outside option r. Although the informed sender knows *the probability distribution over the states*, he does not know *the realisation of the state* when he chooses his mechanism. The objective function of such sender can be rewritten as

$$\pi(\mu; \alpha, \beta, r) = \alpha \int_A d\mu_1 + (1-\alpha) \int_A d\mu_0 = \alpha\mu_1(A) + (1-\alpha)\mu_0(A), \tag{3}$$

where A is the acceptance set for mechanism μ_M and parameters β and r. The optimal mechanism, described in Kamenica and Gentzkow (2011), contains two messages, one of which is sent with probability 1 in a good state.

Lemma 1. *The optimal persuasion mechanism (of the informed sender) contains two messages m^+ and m^- such that:*

1. *if $\beta < r$ then $\hat{\mu}_1(m^+) = 1$ and $P_\beta(m^+) = r$;*

2. *if $\beta \geq r$ then $\hat{\mu}_1(m^+) = 1$ and $P_\beta(m^+) = \beta$;*

We denote this optimal mechanism as $\hat{\mu}$ and the profit associated with this mechanism is given by

$$\pi(\hat{\mu}; \alpha, \beta, r) = \alpha + (1-\alpha) \min\left\{\frac{\beta(1-r)}{r(1-\beta)}, 1\right\}.$$

Now we proceed with the uninformed sender problem, which is the key focus of this paper. The uninformed sender knows neither the probability of the good state α, nor the receiver's prior β and her outside option r. Suppose that such a sender commits to some mechanism μ_M. We define the loss function as the difference between the payoffs of the informed and the uninformed senders:

$$L(\mu_M; \alpha, \beta, r) = \pi(\hat{\mu}; \alpha, \beta, r) - \pi(\mu_M; \alpha, \beta, r). \tag{4}$$

The objective of the uninformed sender is to minimize loss in the worst-case scenario:

$$\inf_{G \in \mathcal{G}} \sup_{F \in \mathcal{F}} \int \int L(\mu_M; \alpha, \beta, r) dF dG, \tag{5}$$

where $\mathcal{G}$ is a set of all probability measures over possible persuasion mechanisms[3] defined on $(M, \mathcal{M})$ and $\mathcal{F}$ is a set of probability distributions over the admissible parameters of the model, i.e. $[\underline{\alpha}, \overline{\alpha}] \times [\underline{\beta}, \overline{\beta}] \times [\underline{r}, \overline{r}] \subset (0,1)^3$. That is, we allow the sender to have reasonable restrictions on the probability of the good state (we rule out situations in which one state can occur with certainty), and the receiver's prior and her outside option. When $\underline{\alpha} = \overline{\alpha} = \alpha_0$ the sender has (subjectively) perfect knowledge of the experimental environment, but faces uncertainty about the type of the receiver. For example, a seller of a product knows the probability of the product being faulty, or a doctor knows the probability of treatment failing to be successful, but they do not know the buyer's subjective view on the likelihood of the product being good (treatment being efficient).

3 Analysis

3.1 Optimal Persuasion Mechanism

Solution to problem (5) is obtained by solving a zero-sum game between the sender, who chooses persuasion mechanism μ_M and malicious nature, which simultaneously chooses parameters of the environment, (α, β, r), in order to maximize the loss. Thus, in our analysis

[3]For representational convenience, we derive the sender's strategy as a randomisation over a set of mechanisms. Intuitively, such randomisation can be represented as a single persuasion mechanism, see Corollary 1.

we are going to focus on the equilibrium of this game.

Although the optimal strategy of the informed sender has a very simple two-message structure, the strategy of the uninformed one may involve many different messages. We show that it is indeed the case, see Corollary 1. We show that despite its complexity this mechanism can be represented as a randomisation over standard mechanisms (see Lemma 2), in which one message perfectly reveals the negative state.

Suppose that malicious nature plays a distribution $\mathbb{F}(\alpha, \beta, r)$. Suppose that the sender plays mechanism $\mu_M = \{\mu_0(m), \mu_1(m)\}$, $m \in M$. The following Lemma states that the set of best responses of the uninformed sender to any strategy of adversarial nature includes a mechanism consisting of two messages.

Lemma 2. *For any $\mathbb{F}$ and any mechanism μ_M the expected loss of the uninformed sender $\mathbb{E}_{\mathbb{F}} L(\mu_M; \alpha, \beta, r)$ can be achieved by a randomisation over mechanisms consisting of at most two messages. Moreover, in each of these mechanisms one of the messages is sent with probability 1 if the state is $\omega = 1$.*

Intuitively, the proof of Lemma 2 relies on the observation that any message m from the original mechanism corresponds to a specific binary mechanism chosen with the same probability $\mu_1(m)$ as message m is sent in the original mechanism. Lemma 2 immediately implies that the payoff from any randomisation over general mechanisms can be matched by an appropriate randomisation over binary mechanisms.

Lemma 2 states that it is sufficient to consider binary recommendation mechanisms with $\mu_1(m^+) = 1$. Thus, any mechanism is fully characterised by a single number, the probability of sending the adoption recommendation in the bad state. Note, that for the binary message mechanism, the sender adopts if her posterior is higher than r, i.e.

$$\frac{\beta}{\beta + (1-\beta)\mu_0(m^+)} \geq r \;\Leftrightarrow\; \lambda \geq \mu_0(m^+),$$

where

$$\lambda = \frac{\beta(1-r)}{r(1-\beta)} \tag{6}$$

is a characteristic of receiver's behaviour, to which we refer as 'receiver's optimism'. Note that the receiver's decision to adopt depends solely on the optimism parameter λ, but the actual decomposition into β and r does not play any role. The higher the λ, the more optimistic receiver is and the more often the sender can send m^+ in the bad state. In what follows, we denote this number as $\mu \equiv \mu_0(m^+)$. From Lemma 1 we get that for the informed sender $\hat{\mu}_0(m^+) = \lambda$. Using Lemma 2 we can rewrite loss function (4) as

$$L(\mu; \alpha, \lambda) = \begin{cases} (1-\alpha)(\lambda - \mu), & \mu \leq \lambda \\ \alpha + (1-\alpha)\lambda, & 1 \geq \mu > \lambda \end{cases}. \tag{7}$$

The first line of the loss function corresponds to the case when the sender is too modest in his recommendation. Recommending adoption in the bad state with higher probability would still persuade the receiver, and thus the sender incurs loss from too infrequent recommendation. The second line of the loss function corresponds to the case when the sender is too ambitious in his recommendation. Recommending adoption in the bad state too frequently leads to failure of message m^+ to persuade the receiver (in both states of the world). When $\mu = \lambda$ the sender chooses the optimal mechanism and loss equals to zero.

As the sender and malicious nature play a zero-sum game, there is no pure strategy equilibrium (the sender wants to choose $\mu = \lambda$ and nature wants to set $\lambda \neq \mu$ to create some loss). In Theorem 1 we characterise a mixed strategy equilibrium of this zero-sum game.[4]

Theorem 1. *There exists a mixed strategy equilibrium in which:*

1. *the sender and nature choose μ and λ from common support $[\underline{\mu}, \overline{\lambda}]$, where $\underline{\mu} \in (\underline{\lambda}, \overline{\lambda})$;*

2. *the sender's strategy is characterised by C.D.F. $G(\mu)$ which is continuous on $(\underline{\mu}, \overline{\lambda}]$;*

[4] We do not prove uniqueness of the equilibrium. However, in zero-sum games of two players all equilibria are payoff-equivalent.

3. *nature's strategy is characterised by C.D.F. $F(\lambda)$ which is continuous on $[\underline{\mu}, \overline{\lambda})$;*

4. *nature sets the probability of a good state according to function*

$$\alpha(\lambda) = \begin{cases} \overline{\alpha}, & \lambda < \lambda_0 \\ \underline{\alpha}, & \lambda \geq \lambda_0 \end{cases} \tag{8}$$

with $\lambda_0 \in (\underline{\mu}, \overline{\lambda})$.

Moreover, this equilibrium is unique in the class of binary message mechanism strategies.

Expressions for G and F are given by equations (27) and (28) at the end of the proof of the Theorem in the Appendix, see page 44. Now we discuss the properties of the equilibrium.

First, the sender and nature randomise over the same convex support. Moreover, if the upper bound of the support of F was below $\overline{\lambda}$, then for any strategy of the sender nature could increase loss by choosing a higher λ, see case 1 in formula (7). This logic does not work for the sender at the lower bound: there is no reason to choose μ below the lower bound of the support of nature's strategy, as this would only increase loss. Hence, it may be the case that $\underline{\mu} > \underline{\lambda}$.

Second, in the interior of the support their strategies are continuous. The intuition is that if the sender had an atom at some point, then nature would like to put some probability mass just below this point, thus increasing loss by making the receiver less persuadable. Hence, the only point which the sender can potentially play with positive probability is when it is impossible to undercut, i.e. $\underline{\lambda}$.

Third, if nature had an atom at some point, the sender would like to transfer probability mass from just below this point slightly upwards. The only case when this is not possible is the upper bound $\overline{\lambda}$.

Fourth, conditional on nature playing some λ, its choice of α is deterministic, see equation (8), and takes extreme values only. Intuitively, such choice makes it harder for the uninformed sender to target the optimal mechanism. Parameter λ_0 identifies the critical level of receiver's optimism: for $\lambda < \lambda_0$ nature sets $\alpha(\lambda) = \overline{\alpha}$ and for $\lambda \geq \lambda_0$ assigns the lowest possible probability, $\underline{\alpha}$, to the good state. That is, nature makes the probability of the good state negatively correlated with receiver's optimism: whenever the receiver is sufficiently optimistic, the probability of the good state is minimal and vice versa.[5]

The sender's strategy G described in Theorem 1 may be interpreted as a mechanism involving many messages. Indeed, such strategy chooses a binary mechanism $\{\mu_0(m^+) = \mu, \mu_1(m^+) = 1\}$ with the probability $g(\mu)$. Note that this binary mechanism persuades any receiver with $\lambda \geq \mu$. Alternatively, the sender could commit to a large mechanism which sends a message "adopt if your $\lambda \geq \mu$" with probability $g(\mu)$ in the good state and $\mu g(\mu)$ in the bad state. Since for each (α, λ) this mechanism induces the same adoption probability, the best response of nature is still F.[6] We sum up this result in the following corollary.

Corollary 1. *The sender can minimise the worst-case loss by choosing the following mechanism:*

- *send message "adopt if your $\lambda \geq \mu$" with density $g(\mu)$ if $\omega = 1$;*
- *send message "adopt if your $\lambda \geq \mu$" with density $\mu g(\mu)$ if $\omega = 0$;*
- *send message "reject" with probability $1 - \mathbb{E}_G \mu$ if $\omega = 0$;*

where g is the density function corresponding to C.D.F. G defined in Theorem 1.

Comparing the mechanism from Corollary 1 with the informed sender benchmark mechanism from Lemma 1 leads to the following observations. Firstly, in the presence of uncertainty the sender finds in optimal to use a continuum of messages instead of two. Secondly, under

[5]More precisely, if λ is high, then the sender is more likely to end up in the first case in formula (7), implying that it is optimal to have the lowest possible α. Similarly, when λ is low, then the second case is more likely, meaning that the choice of the largest possible α maximises loss.

[6]Note that this construction essentially reverse-engineers the steps in the proof of Lemma 2 where we approximated an arbitrary mechanism with a randomisation over binary mechanisms.

uncertainty the sender randomises over messages in the good state of the world, while the informed sender sticks to a pure strategy in this state. Finally, both mechanisms involve a message which perfectly reveals state $\omega = 0$.

Now we touch upon the question of uniqueness of the optimal mechanism. Theorem 1 established uniqueness of the equilibrium in the class of binary mechanisms, in which the sender randomises between the messages only when $\omega = 0$. However, as it follows from Corollary 1 this randomisation can be formulated as an alternative mechanism with many messages. It is possible to construct other mechanisms, for example, by simply relabeling the messages. However, each such mechanism, according to Lemma 2, can be replicated by an appropriate mixture of binary mechanisms. The following Proposition verifies that this replication is unique, i.e. any equilibrium mechanism is pay-off equivalent to the mixed strategy described in Theorem 1. Moreover, the strategy of nature is identical for each such mechanism.

Proposition 1. *Let G and F be the equilibrium strategies of the sender and nature described in Theorem 1. Then, the set of equilibrium mechanisms and strategies of nature (G', F') is such that $F' = F$, and G' and G result in the same loss, i.e.*

$$\int\int L(\mu_M; \alpha, \beta, r) dF dG' = \int\int L(\mu_M; \alpha, \beta, r) dF dG.$$

3.2 Properties of Minimax Loss

Now we proceed with characterisation of the minimax loss. We start with deriving global properties and then focus on the impact of uncertainty on the minimax loss.

Lemma 3. *The minimax loss of the sender is given by*

$$\overline{L} = \ln\left(\frac{\underline{\alpha} + (1-\underline{\alpha})\overline{\lambda}}{\underline{\alpha} + (1-\underline{\alpha})\lambda_0}\right). \tag{9}$$

This result is obtained directly from the equilibrium strategies derived in Theorem 1. Note that λ_0, i.e. the point where nature switches from the highest to the lowest α depends also on $\overline{\alpha}$ and $\underline{\lambda}$, meaning that the minimax loss depends on all four boundaries of the support of the nature's strategy. Now we turn our attention to the properties of the minimax loss.

Proposition 2. *The minimax loss is characterised by the following properties:*

1. *$\overline{L}$ strictly increases in $\overline{\alpha}$ and $\overline{\lambda}$, strictly decreases in $\underline{\alpha}$ and weakly decreases in $\underline{\lambda}$;*
2. *$\sup_{\underline{\alpha},\overline{\alpha},\underline{\lambda},\overline{\lambda}} \overline{L} = \mathbf{\Omega}$, where $\mathbf{\Omega}$ is the solution to $\mathbf{\Omega} e^{\mathbf{\Omega}} = 1$;*
3. *if $\underline{\lambda} = \overline{\lambda}$ then $\overline{L} = 0$;*
4. *if $\underline{\alpha} = \overline{\alpha}$ then $\overline{L} \leq \frac{1}{e}$.*

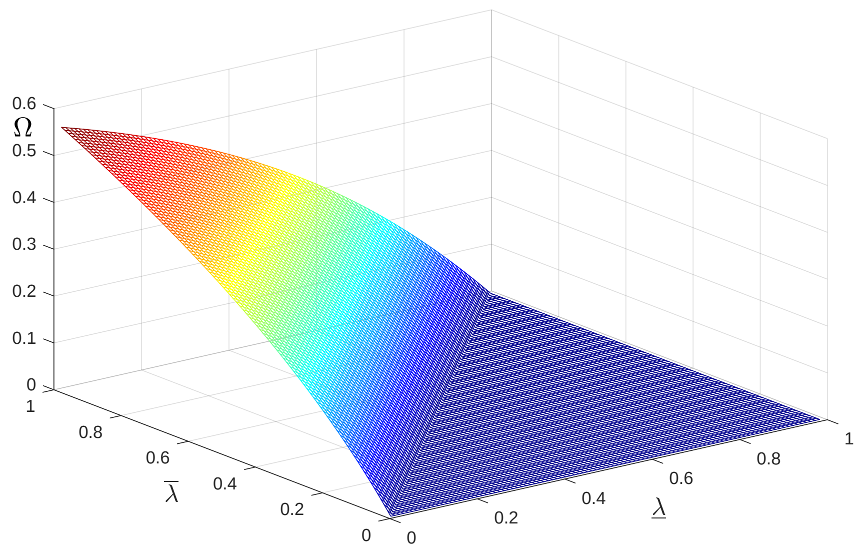
The first part of Proposition 2 is very intuitive: more uncertainty always harms the sender, as it makes it harder to target the optimal mechanism, while the informed sender benchmark is always optimal given the parameters.

The second part states that if uncertainty is maximal, then the minimax loss is approaches to what is called the 'omega constant', which equals approximately to 0.567.[7]

The third part of Proposition 2 says that when there is no uncertainty about the receiver's type then the minimax loss equals to zero. This is intuitive: the optimal mechanism employed by the informed sender does not depend on α, thus the uninformed sender can perfectly match this mechanism. The only impact of uncertainty is that the uninformed sender does not know the exact value of the equilibrium payoff, but is certain that the maximal possible payoff will be achieved, whatever its value is. Parts two and three of the Proposition are illustrated in Figure 1. In this figure we allowed for the maximal dispersion in α. In this case when $\underline{\lambda} = 0$ and $\overline{\lambda} = 1$ the minimax loss equals to $\mathbf{\Omega}$, see upper left corner. When $\underline{\lambda} = \overline{\lambda}$, which corresponds to the points on the diagonal, the minimax loss equals to zero.

[7]See `https://en.wikipedia.org/wiki/Omega_constant`

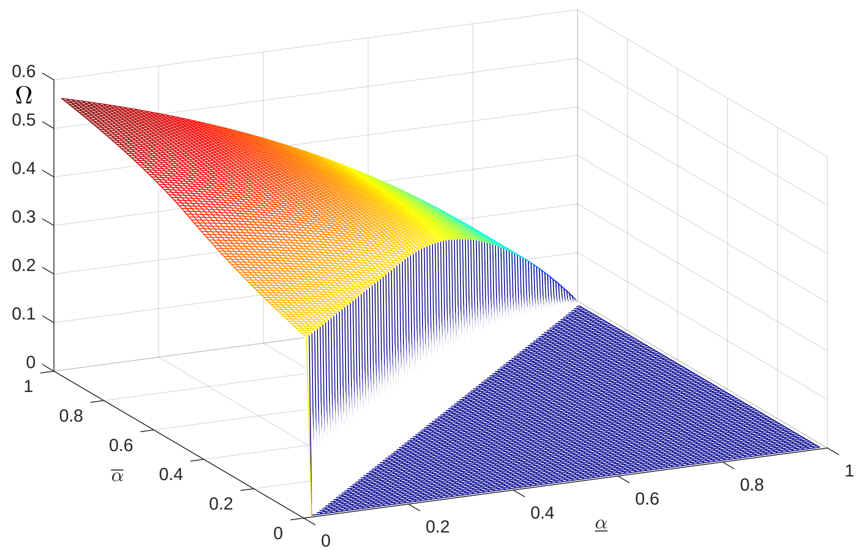
Figure 1: Equilibrium Losses in $[\underline{\lambda}, \overline{\lambda}]$ for $\underline{\alpha} \to 0$, $\overline{\alpha} \to 1$.



Finally, the fourth part of Proposition 2 illustrates the role of uncertainty in λ. Suppose that the probability of the good state is known and takes value α. Then, for generic $\underline{\lambda}$ and $\overline{\lambda}$ the minimax loss first increases and then decreases in α. Moreover, if we allow for the largest possible uncertainty in λ, i.e. $\underline{\lambda} = 0$ and $\overline{\lambda} = 1$, the minimax loss is maximal for all $\alpha < 1/e$ and then strictly decreases in α. This case corresponds to the diagonal in Figure 2. When $\alpha \to 1$ loss from using too truthful strategy approaches zero, while loss of excessive lying approaches one. In this case the sender tends to use almost truthful mechanisms, which, given that $\alpha \to 1$, generates a vanishingly small minimax loss.

This case of known α can be of particular interest for many applications. For example, a seller might know the true quality of the product, modelled as a probability to fit a particular buyer. What the seller does not know, however, is which product the buyer is currently using and what the buyer's beliefs about how likely the product will fit her are. The seller designs an advertising strategy or a testing procedure, i.e. a persuasion mechanism, to attain the maximal probability of selling the product. Note that the strategy of the sender in this case is

Figure 2: Equilibrium Losses in $[\underline{\alpha}, \overline{\alpha}]$ for $\underline{\lambda} = 0$ and $\overline{\lambda} = 1$.



obtained directly from Theorem 1 by plugging in $\underline{\alpha} = \overline{\alpha} = \alpha$.[8] We present these much simpler expressions in the following Corollary.

Corollary 2. *Suppose that $\underline{\alpha} = \overline{\alpha} \equiv \alpha$ and $\lambda \in [\underline{\lambda}, \overline{\lambda}]$. Then equilibrium strategy of the sender is*

$$G(\mu) = 1 + \ln\left(\frac{\alpha + (1-\alpha)\mu}{\alpha + (1-\alpha)\overline{\lambda}}\right), \tag{10}$$

with support $\left[\underline{\mu}, \overline{\lambda}\right]$, where

$$\underline{\mu} = \max\left\{\underline{\lambda}, \frac{1}{1-\alpha}\left(\frac{\alpha + (1-\alpha)\overline{\lambda}}{e} - \alpha\right)\right\}.$$

The associated minimax loss is given by

$$\overline{L} = [\alpha + (1-\alpha)\underline{\lambda}] \ln\left(\frac{\alpha + (1-\alpha)\overline{\lambda}}{\alpha + (1-\alpha)\underline{\lambda}}\right) \leq 1/e.$$

Interestingly, when nature is not restricted in its strategies, the value of knowing the true

[8]See formulae (27) and (28) in the proof the Theorem in the Appendix.

state of the world reduces the highest minimax loss by approximately one third. That is,

$$\Delta\overline{L} = \Omega - \frac{1}{e}.$$

This difference can be interpreted in the following way. Suppose, that nature is free in its choice of λ and α, but now is required to reveal its choice of α to the sender prior to the choice of the persuasion mechanism. In this case, nature would be indifferent between all $\alpha < 1/e$ (the flat part of the diagonal in Figure 2) and the sender would enjoy a decrease in the expected loss of $\Delta\overline{L}$. Note that the diagonal of Figure 2 is identical to Figure 2 in Babichenko et al. (2021), i.e. to the highest minimax loss in a persuasion problem with multiple states of the world, where nature chooses a monotone utility function over the states. In our setting the utility function is fixed up to the outside option, and the ability to choose a prior still leaves the problem of nature being more restricted than in Babichenko et al. (2021).

Proposition 2 considers global properties of the minimax loss. Now we are going to explore the marginal impact of small uncertainty in parameters on the minimax loss. Suppose that either α or λ is fixed, with complimentary variable taking values from a permissible range. Now we allow nature to vary corresponding parameter, for example, the probability of a good state, in a neighbourhood $[\alpha - \varepsilon/2, \alpha + \varepsilon/2]$. Part 1 of Proposition 2 establishes that the minimax loss should increase. Now we are going to establish by *how much* it increases in response to a small increase in uncertainty.

Proposition 3. *Consider small parameter uncertainty.*

1. *Suppose that* $\alpha \in [\underline{\alpha}, \overline{\alpha}]$ *and let* $\lambda \in [\hat{\lambda} - \varepsilon/2, \hat{\lambda} + \varepsilon/2]$. *Then, when* $\varepsilon \to 0$,

$$\overline{L} = (1 - \underline{\alpha})\varepsilon + o(\varepsilon).$$

2. Suppose that $\lambda \in [\underline{\lambda}, \overline{\lambda}]$ and let $\alpha \in [\hat{\alpha} - \varepsilon/2, \hat{\alpha} + \varepsilon/2]$. Then, when $\varepsilon \to 0$,

$$\overline{L} = \overline{L}_0 + k\varepsilon + o(\varepsilon), \quad k < 1/2,$$

where $\overline{L}_0 = [\hat{\alpha} + (1 - \hat{\alpha})\underline{\lambda}] \ln \left(\frac{\hat{\alpha}+(1-\hat{\alpha})\overline{\lambda}}{\hat{\alpha}+(1-\hat{\alpha})\underline{\lambda}} \right)$.

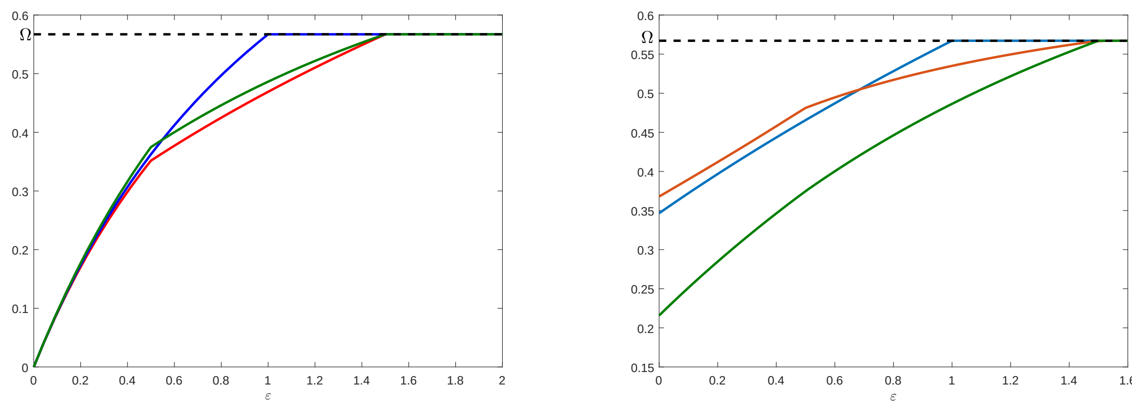
The first part of Proposition 3 says that the impact of the uncertainty in λ crucially depends on $\underline{\alpha}$, but does not depend on $\overline{\alpha}$. The intuition is as follows. Consider equation (9). We know that $\overline{\lambda}$ increases by $\varepsilon/2$, and in the worst-case λ_0 can decrease by $\varepsilon/2$ (i.e. not by more than the decrease in $\underline{\lambda}$) which, for small ε, results in total impact of $(1 - \underline{\alpha})\varepsilon$ (due to the coefficient $(1 - \underline{\alpha})$ in front of both λ's). The upper bound $\overline{\alpha}$ affects the minimax loss only indirectly via λ_0, but does not play any role for sufficiently small ε. Thus, overall impact of uncertainty in λ is generically less than one-to-one. For $\varepsilon = 0$ the minimax loss equals zero, as follows from part 3 of Proposition 2.

The second part of Proposition 3 says that uncertainty in α of size ε generates at most a $\frac{1}{2}\varepsilon$ increase in the minimax loss. The precise value of k depends on which part of diagonal in Figure 2 we are, explicit forumlae for each of the cases are given by equations (37) and (39) in the Appendix. For $\varepsilon = 0$ the minimax loss is positive, unlike in the first part of the Proposition. This is because uncertainty in λ alone, unlike just uncertainty in α, is sufficient to generate a positive loss. Moreover, from part 4 of Proposition 2 we know that $\overline{L}_0 \leq 1/e$. Intuitively, for any $\varepsilon \geq 0$ the highest minimax loss is obtained when nature is not restricted in its choice of λ.

Proposition 3 characterises the effects of uncertainty locally, for a marginal increase in the range of nature's possible choices. Figure 3.2 illustrates numerically what happens when uncertainty grows large.

In Figure 3(a) we leave nature's choice of α unrestricted, i.e. $\underline{\alpha} \to 0$ and $\overline{\alpha} \to 1$, and fix one of three starting values of $\hat{\lambda} \in \{1/4, 1/2, 3/4\}$. Then we allow $\varepsilon/2$ approach 1. Firstly, note that in the neighbourhood of zero for either case losses increase in one-to-one proportion

Figure 3: Impact of Uncertainty on Loss



(a) The minimax loss in ε for unrestricted α; $\hat{\lambda} = 0.25$ (red), $\hat{\lambda} = 0.5$ (blue), $\hat{\lambda} = 0.75$ (green)

(b) The minimax loss for unrestricted λ, $\hat{\alpha} = 0.25$ (red), $\hat{\alpha} = 0.5$ (blue), $\hat{\alpha} = 0.75$ (green)

to ε, as claimed in the first part of Proposition 3 for $\underline{\alpha} \to 0$. When ε grows larger, at some point either $\underline{\lambda}$ or $\overline{\lambda}$ hit the boundary, either 0 or 1, while vis-a-vis boundary is still interior. For example, when $\varepsilon = 1/2$, $\underline{\lambda} = 0$ and $\overline{\lambda} = 1/2$ for the red line, and $\underline{\lambda} = 1/2$ and $\overline{\lambda} = 1$ for the green line. If ε grows even further, then uncertainty feeds into the minimax loss via just one boundary of the support rather than both boundaries as before, thus making $\overline{L}$ shallower in ε. Finally, when both support boundaries hit the edges, i.e. zero and one, the minimax loss reaches its maximum, $\mathbf{\Omega}$. Note that for the initial choice of $\hat{\lambda} = 1/2$ we have that $\underline{\alpha}$ reaches zero and $\overline{\alpha}$ reaches one for the same $\varepsilon = 1$, thus only single kink is present.

Figure 3(b) leaves λ unrestricted and allows uncertainty about α grow larger. Note that even if $\varepsilon = 0$, $\overline{L} \geq 0$ (see second part of Proposition 3) as uncertainty in λ alone is sufficient to generate positive loss. The slope of all graphs, however, is less than $1/2$, as described in Proposition 3. Interestingly, the green line ($\hat{\alpha} = 3/4$) does not have a kink at $\varepsilon = 1/2$, i.e. when $\overline{\alpha} \to 1$. This is because the mimimax loss given by formula (9) depends on $\overline{\alpha}$ only via λ_0 and $\lim_{\overline{\alpha} \to 1} \frac{\partial \lambda_0}{\partial \overline{\alpha}} = 0$.[9]

[9] See Lemma C.5 in Appendix C.

3.3 Probability of Lying

We now consider the probability of the sender lying in the bad state, i.e. recommending to adopt when $\omega = 0$. This probability is given by the expected value of μ given the optimal strategy of the sender. The following Lemma characterises the average probability of lying.

Lemma 4. *The probability of lying, i.e. recommending adoption in the bad state, can be represented as*

$$\mathbb{E}_G \mu = \overline{\lambda} - \frac{\overline{L}}{1 - \underline{\alpha}}.$$

The probability of lying is inversely related to the minimax loss. That is, situations when the sender suffers high loss are associated with high degree of truth-telling, which usually happens in more 'pessimistic' environments. In the following Proposition we describe the properties of the probability of lying.

Proposition 4. *The probability of lying $\mathbb{E}_G \mu$ strictly increases in $\overline{\lambda}$, weakly increases in $\underline{\lambda}$, and strictly decreases in $\overline{\alpha}$ and $\underline{\alpha}$.*

If either $\overline{\lambda}$ or $\underline{\lambda}$ go up, then the receiver is on average easier to persuade. The sender naturally responds to a more optimistic distribution of receivers with a choice of a more ambitious strategy, i.e. he is more likely to recommending adoption in the bad state.

Whenever $\overline{\alpha}$ or $\underline{\alpha}$ increases, the probability of being in a good state goes up. This, in turn, increases the cost of not persuading the receiver, as choosing too high μ would result in the receiver rejecting regardless of the state of the world, while using a more modest strategy with lower μ harms the sender only in the bad state, which is relatively less likely than before. Thus, higher costs of failing to persuade the receiver result in the choice of a more truthful mechanism.

The impact of increasing uncertainty on the probability of lying is ambiguous. First, consider the impact of increase in uncertainty in λ, where $\overline{\lambda}$ increases and $\underline{\lambda}$ decreases by $\varepsilon/2$. In this case, if uncertainty is small, we know that the minimax loss increases proportional to

$(1-\underline{\alpha})\varepsilon$, see Proposition 3. Then, from Lemma 4 we get that

$$\Delta \mathbb{E}_G \mu \approx \frac{\varepsilon}{2} - \varepsilon = -\frac{\varepsilon}{2}.$$

That is, when uncertainty is small the probability of lying decreases in the level of uncertainty. In other words, the sender who is just a little bit unsure about the receiver's prior is more truthful than the informed one.

Second, consider larger levels of uncertainty in λ. At some point it may be the case that the lower bound $\underline{\mu}$ becomes larger than $\underline{\lambda}$. In this case $\overline{L}$ does not depend on $\underline{\lambda}$, thus we can apply Proposition 4, which states that $\mathbb{E}_G \mu$ increases in $\overline{\lambda}$ and thus, increases in uncertainty.

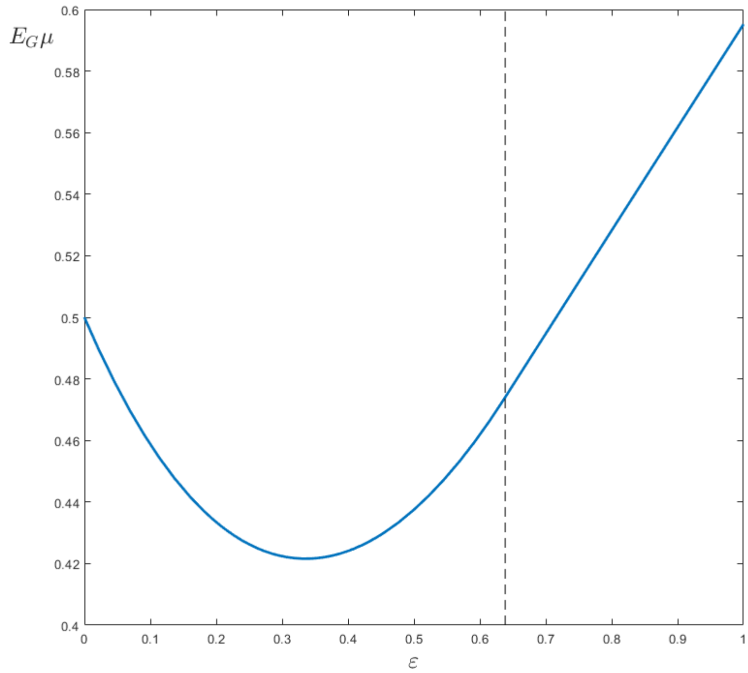
A combination of these two results is presented in Figure 4. In this figure $\lambda \in [\hat{\lambda} - \varepsilon/2, \hat{\lambda} + \varepsilon/2]$ and we vary ε between zero and one. When there is no uncertainty, i.e. $\varepsilon = 0$, the sender uses the optimal strategy, i.e. recommends adopting with probability $\mu = \hat{\lambda} = 1/2$. For small ε the probability of lying decreases in uncertainty, as explained above. For ε very large, when the lower bound of the support is larger than $\underline{\lambda}$ (all points to the right of the vertical dashed line) , the probability of lying strictly increases in ε. As one can see, there is an interior minimum of the probability of lying which is attained at such levels of uncertainty that the lower bound is still determined by $\underline{\lambda}$. Interestingly, the largest level of uncertainty corresponds to the least truthful strategy of the sender.

The impact of uncertainty in $\underline{\alpha}$ or $\overline{\alpha}$ on the sender's strategy is ambiguous and depends on the initial choice of the parameters. Our numerical analysis shows that the probability of lying can both increase or decrease in uncertainty even in the neighbourhood of zero, depending on the initial choice of parameters.

3.4 Informed Receiver

In Corollary 2 we showed that the sender can reduce the minimax loss from $\mathbf{\Omega}$ to $\frac{1}{e}$ by being informed about distribution over the states of the world. Knowing the outside option of the

Figure 4: $\mathbb{E}_G(\mu)$ as a function of ε, $\underline{\alpha} = 0.1$, $\overline{\alpha} = 0.9$, $\hat{\lambda} = 0.5$.



receiver in that setting does not help: by just varying β nature could generate any value of λ for any fixed $r \in (0, 1)$. Now we look at an alternative setting, where the sender knows the outside option r, but also knows that the receiver holds a *correct prior* about the probability distribution over the states of the world, i.e. $\beta = \alpha$. This puts a restriction of the strategy of nature: it cannot move λ and α independently.[10] As we show, in this setting the upper bound on the sender's minimax loss is the same as in the case when the sender himself is informed about the state of the world, i.e. $1/e$ as in Corollary 2.

We introduce several modifications to the main model. We focus on the situation in which receiver knows the true underlying distribution over the states, i.e. $\beta = \alpha$ and this fact is known to the sender, although the exact value of α remains unknown. Moreover, we assume

[10]Note that our assumption that the receiver knows r is crucial here. Otherwise, nature would have freedom in moving λ by just adjusting r, which would make the model identical to our baseline.

that the outside option r is known to the sender. We define

$$\lambda_r(\alpha) \equiv \frac{\alpha}{1-\alpha}\frac{1-r}{r}. \tag{11}$$

That is, $\lambda_r(\alpha)$ is the level of the receiver's optimism similar to the one defined in (6). The important difference between the two is that in (6) nature is free in its choice of λ for any given choice of α, while in (11) the value of λ is determined by nature's choice of α.

In what follows we are going to focus on the case where the strategy of nature is not restricted, i.e. $\underline{\alpha} \to 0$ and $\overline{\alpha} \to 1$. In this case for a given choice of mechanism μ and nature's choice of α the loss function is given by

$$L_r(\mu;\alpha) = \begin{cases} (1-\alpha)(\min\{\lambda_r(\alpha),1\}-\mu), & \mu \leq \lambda_r(\alpha) \leq 1 \\ \alpha + (1-\alpha)\lambda_r(\alpha), & \mu > \lambda_r(\alpha) \end{cases} \tag{12}$$

Note that $\lambda_r(\alpha) \geq 1$ if and only if $\alpha \geq r$, i.e. when the receiver's prior is sufficiently strong to justify adoption without any additional information. Also, loss in the second case can be rewritten as $\alpha + (1-\alpha)\lambda_r(\alpha) = \alpha/r$.

For the same reasons as in our baseline model, both the sender and malicious nature are playing mixed strategies. We use $\underline{\mu}$ and $\overline{\mu}$ to denote the lower and the upper bounds of the support of the sender's strategy $G(\mu)$, and $\underline{\alpha}_F \geq 0$ and $\overline{\alpha}_F \leq 1$ to denote the lower and the upper bounds of the support of the strategy of nature $F(\alpha)$. Now we are ready to characterise the optimal mechanism of the sender.

Proposition 5. *Suppose that $r \in (0,1)$. Then, there exists a mixed strategy equilibrium of the game, such that*

1. *sender chooses mechanism according to a continuous distribution function*

$$G(\mu) = \frac{1}{1-\underline{\mu}^r}\left[1-\left(\frac{\underline{\mu}}{\mu}\right)^r\right],$$

with support $[\underline{\mu}, 1]$, where $\underline{\mu}$ solves

$$\underline{\mu}^r + \frac{r^2}{1-r}\underline{\mu} - (1-r) = 0; \tag{13}$$

2. nature plays a distribution function F, which is continuous on $[\underline{\alpha}_F, \overline{\alpha}_F)$ with

$$\underline{\alpha}_F = \lambda_r^{-1}(\underline{\mu}) = \frac{r\underline{\mu}}{1-r+r\underline{\mu}}, \quad \overline{\alpha}_F = \lambda_r^{-1}(1) = r.$$

The equilibrium strategy of the sender is always a continuous function, without an atom at the lower bound, which is possible in our baseline model. The reason is that the support of nature's equilibrium strategy is wide enough. Moreover, in the equilibrium supports of λ and μ coincide. Note that unlike in the baseline model where α and λ are correlated in a negative way, now the receiver's optimism and the probability of the good state are positively related via function $\lambda_r(\alpha)$ given by equation (11). Potentially, nature could choose large α, so that $\lambda_r(\alpha) > 1$ and the receiver would have strict incentives to adopt based on the prior. However, this would come at a cost: when $\mu \leq \lambda_r(\alpha)$ in equation (12), loss is decreasing in α whenever $\lambda_r(\alpha) > 1$. Thus, playing α which pushes $\lambda_r(\alpha)$ above one is sub-optimal. However, nature still finds it optimal to push $\overline{\alpha}_F$ all the way to r, so that the largest λ played equals to one, as in this case loss given by (12) when $\mu \leq \lambda_r(\alpha)$ is increasing in α given that $\lambda_r(\alpha)$ is defined by (11).

The fact that the receiver holds a correct prior belief replaces a negative dependence between the probability of the good state and receiver's optimism, which was optimal from nature's point of view, with a positive one. Not surprisingly, this reduces the equilibrium loss of the sender. The following proposition quantifies this impact.

Proposition 6. *The minimax loss in the model with informed receiver is given by*

$$\overline{L}_r = \frac{\underline{\alpha}_F}{\overline{\alpha}_F} = \frac{\underline{\mu}}{1-r+r\underline{\mu}}. \tag{14}$$

$\overline{L}_r$ decreases in r and is bounded above by $\lim_{r\to+0}\overline{L}_r = 1/e$.

Equation (14) might seem counter-intuitive, since the higher the ratio between the lower and the upper bound, the higher the equilibrium loss. However, one should bear in mind that in equilibrium both $\underline{\alpha}_F$ and $\overline{\alpha}_F$ are functions of r. The mechanism at work is as follows. As r increases, the receiver becomes more difficult to persuade. As a result, the sender tends to use more truthful strategies, so that ending up in case $\mu \le \lambda_r(\alpha)$ in equation (12) is more likely. Also, as r increases, the distribution $F(\alpha)$ shifts up (simultaneously widening the support), resulting in lower loss in case $\mu \le \lambda_r(\alpha)$ in (12).

Proposition 6 contains another key result. The maximum value of loss equals to $1/e$, the same as in the case when the sender knows exactly the true state of the world, see Corollary 2. This implies, that if there are no exogenous restrictions on the strategy of nature, then the sender is equally well-off if he *knows* α or *knows that the receiver knows* α. Intuitively, in both cases nature's strategy is restricted to choosing a single variable, leading to the same level of flexibility and the same minimax loss.

4 Model with Large Number of States

In this section we generalise our model to a richer state space. We assume that the state space consists of $n+1$ states:

$$\Omega = \left\{0, \frac{1}{n}, \frac{2}{n}, \dots, \frac{n-1}{n}, 1\right\}.$$

We also assume that, similarly to the setting in Corollary 2, the uninformed sender knows the true probability distribution over the states of world, which is now represented as a vector $(\alpha_\omega)_{\omega\in\Omega}$ with $\sum_\Omega \alpha_\omega = 1$. In all other features the model we analyse here is identical to our baseline. That is, adversarial nature chooses receiver's prior β_ω and outside option r, while the sender chooses a persuasion mechanism which maps Ω to the message space M. Upon learning about the mechanism and a signal it generated, the receiver updates her beliefs in a

Bayesian way and decides whether adopt or reject.

We start with a negative result for large number of states, i.e. $n \to \infty$. We assume that for any two states $\omega', \omega'' \in \Omega$ we have $\alpha_{\omega'}/\alpha_{\omega''} < \overline{A}$, where $\overline{A} \in \mathbb{R}$. That is, as the number of states grows larger, none of the states becomes infinitely more likely than some other state. Note that $\alpha_\omega < \overline{A}/n$ for any $\omega \in \Omega$. This implies that the distribution of states α_ω converges to a continuous distribution function. We again denote the minimax loss as $\overline{L}$.

Theorem 2. *Suppose that the distribution of states $(\alpha_\omega)_{\omega\in\Omega}$ is known to the sender and nature is unrestricted in its choice of the receiver's prior and the outside option. Furthermore, suppose that for any n $\alpha_\omega < \overline{A}/n$ for some $\overline{A} \in \mathbb{R}_{++}$. Then,*

$$\lim_{n\to\infty} \overline{L} = 1.$$

The mechanism behind Theorem 2 is as follows. For illustration purposes suppose that all states are equally likely. Then adversarial nature can choose the following combination of the receiver's prior and the outside option. Nature assigns the prior of the receiver in the following way:

1. puts a large probability mass on state $\omega = 1$;

2. randomly selects $\lfloor \sqrt{n} \rfloor$ states among the remaining ones,[11] which we call 'mine states', and assigns the same small probability mass to each of them; we denote this mass as δ, require that the total mass of mine states approaches zero as n approaches infinity;

3. assigns probability mass ε, such that $\sqrt{n}\varepsilon/\delta \to 0$, to each of the $n - \lfloor \sqrt{n} \rfloor$ remaining states;

4. chooses $r \in \left(\frac{n-1}{n}, 1\right)$ such that pooling state $\omega = 1$ with state $\omega = (n-1)/n$ happening with probability δ leads to rejection, but if state $\omega = 1$ is pooled even with the lowest

[11] $\lfloor x \rfloor$ is the largest integer smaller or equal to x.

$n - \lfloor \sqrt{n} \rfloor$ states, each happening with probability ε, then the receiver finds it optimal to adopt.

Intuitively, nature puts $\lfloor \sqrt{n} \rfloor$ 'mines' in n states below $\omega = 1$. These states are assigned with high enough prior δ, such that pooling them with the top state drags the expected utility of the receiver down sufficiently to result in rejection. The informed sender can avoid these 'mines', and pool all the remaining states, as receiver's prior belief of being in these states, ε, is so low that the expected utility does not decrease enough to justify rejection. Thus, the informed sender can persuade the receiver in $n - \lfloor \sqrt{n} \rfloor$ states, so almost always in large state spaces since

$$\lim_{n \to \infty} \frac{n - \lfloor \sqrt{n} \rfloor}{n} = 1.$$

In order to persuade the receiver in a number of states which does not become vanishingly small as $n \to \infty$, the uninformed sender has to include $\lfloor \gamma n \rfloor$ out of n states in the receiver's posterior, $\gamma \in (0, 1)$. However, for any fixed γ the probability of having a 'mine' among any $\lfloor \gamma n \rfloor$ states approaches one as n approaches infinity. As having a 'mine' in a posterior leads to rejection, the uninformed sender can only persuade in a vanishingly small fraction of states, meaning that the minimax loss approaches one.

Theorem 2 has interesting connection to results of Babichenko et al. (2021). They consider a setting with a common prior, but unknown utility of the receiver. If the utility of the receiver can take an arbitrary form, then the minimax loss approaches one, as the number of states approaches infinity, see Theorem 3.5 in Babichenko et al. (2021). If utility is monotone, then the minimax loss is bounded above by $1/e$, see their Theorem 3.6. In our setting, we have monotone utility, but it is partially compensated by nature's freedom to choose the receiver's prior. In the limit, this delivers a similar result, but for finite n inability to vary the best state makes nature's task more difficult.

Although the results of Theorem 2 might seem rather pessimistic for prospects of the uninformed sender in large state environments, it is important to understand that the result relies on having rather peculiar optimal persuasion mechanisms. In such mechanisms the informed sender hand-picks some small number of states in which rejection is recommended. Most of real-life mechanisms are unlikely to exhibit this feature. We now look at a restricted set of mechanisms, which arguably are more likely to be implemented in reality, namely cut-off mechanisms. In a cut-off mechanism, there is a state ω_0, such that message m^+ is sent in all $\omega > \omega_0$, sends message m^- in all states $\omega < \omega_0$ and potentially randomises between both messages at ω_0. We assume that both the informed and the uninformed senders are restricted to cut-off strategies. We are going to establish the equivalence between the problem restricted to cut-off strategies with the binary problem we studied in Section 3.

We start with showing that any cut-off rule can be described by a real number in $[0, 1]$. Suppose that the sender plays a mechanism in which message m^+ is sent with probability one in all $\omega > \omega_0$, where $\omega_0 < 1$,[12] and with probability $\eta \in [0, 1]$ in state ω_0.

We define $\lambda = 1 - \omega_0 - (1 - \eta)\frac{1}{n}$. Then, since

$$\omega_0 = \frac{\lfloor (1-\lambda)n \rfloor}{n} \quad \text{and} \quad \eta = 1 - ((1-\lambda)n - \lfloor (1-\lambda)n \rfloor),$$

we obtain that there is one-to-one mapping between any cutoff mechanism and $\lambda \in [0, 1]$. Analogously to our main model, λ can be interpreted as receiver optimism: in all states $\omega > \lambda$ the informed sender can send m^+ with probability 1, and in state just below λ with probability η, and still persuade the receiver.

We define

$$y(\lambda) = \frac{1}{1-\alpha_n}\left[-\alpha_n + \sum_{\omega > \lfloor (1-\lambda)n \rfloor / n} \alpha_\omega + \alpha_{\frac{\lfloor (1-\lambda)n \rfloor}{n}} (1 - (1-\lambda)n + \lfloor (1-\lambda)n \rfloor)\right],$$

that is, the total probability of sending message m^+ conditional on $\omega \neq 1$ if a mechanism characterised by λ is being played. Note that $y(\lambda) : [0, 1] \to [0, 1]$ is a strictly increasing

[12] It is always optimal to send m^+ in $\omega = 1$.

function. The term in square brackets equals to total probability of adoption minus α_n.

Now, suppose that for some realisation of the strategy of nature the informed sender uses the optimal mechanism which persuades the receiver with optimism λ, i.e. a mechanism which recommends adoption in all states except the top with probability $y(\lambda)$ and with probability one in $\omega = 1$. Suppose that the uninformed sender uses a cut-off mechanism characterised by μ. Then, the loss function can be represented by

$$L(\mu; \alpha, \lambda) = \begin{cases} (1 - \alpha_n)(y(\lambda) - y(\mu)), & \mu \leq \lambda \\ \alpha_n + (1 - \alpha_n)y(\lambda), & 1 \geq \mu > \lambda \end{cases}. \tag{15}$$

That is, loss depends on how much the sender lies in states below $\omega = 1$ relative to full information benchmark. Note that this loss function (15) is equivalent to loss in the binary problem (7) with α replaced with α_n, λ and μ replaced with $y(\lambda)$ and $y(\mu)$ respectively. Thus, we can directly apply Corollary 2 with $\underline{\lambda} = 0$ and $\overline{\lambda} = 1$. This allows us to formulate the following result.

Proposition 7. *Suppose that the sender knows the probability distribution over the states of the world α and nature is free in its choice of the receiver's prior and the outside option. Moreover, suppose that the sender is restricted to cut-off strategies. Then, the minimax loss of the uninformed sender depends only on the probability of the best state, and does not depend on the number of states in Ω.*

Proposition 7 implies that the probability of the best state is the only important parameter for determining the minimax loss. If $\alpha_n < 1/e$ then the minimax loss equals to $1/e$, otherwise it equals to $-\alpha_n \ln \alpha_n$ and converges to 0 as α_n approaches one. This result is related to Babichenko et al. (2021), see Theorem 3.6, where the cut-off property of the optimal mechanism is delivered by a combination of monotonic utility and common prior. Proposition 7 implies that it is the cut-off property, rather than the common prior assumption is crucial for obtaining this result.

5 Conclusion

In this paper we characterised the optimal persuasion mechanism when the sender does not have priors about the environment he faces: probability distribution over possible states, the receiver's prior and outside option. Maximin utility approach, which is prevalent in literature on robust persuasion, has little bite in our setting without any external constraints. Moreover, it restricts the sender's attention to a specific worst-case prior of the receiver. Instead, we use a minimax loss approach, in which the sender is striving for the best *irrespective* of the receiver's prior. With this approach our setting can be interpreted as a problem of a sender facing a population of receivers with unknown characteristics.

In a model with binary state space the optimal mechanism is complicated and involves a continuum of messages, but it admits a simple representation as a randomisation over two-message mechanisms. We characterise properties of the minimax loss, and demonstrate that having perfect information about the receiver's characteristics reduces loss from $\mathbf{\Omega}$ to zero, while having perfect information about the state space reduces it just to $1/e$. Surprisingly, to achieve this loss the sender does not need to know probabilities of the states, but can be equally well-off if he is sure that the receiver has an accurate knowledge of them. We show that small uncertainty makes the sender more truthful than in the perfect information case, but larger uncertainty can lead to more lying. This indicates that introducing some small noise into the environment might improve the quality of communication.

Finally, we show that as the number of states approaches infinity the sender's performance drops to zero. This result, however, arises in the enviromnens when the optimal mechanism excludes specific, randomly chosen by nature, states from the support of adoption recommendation. For example, the sender might find it optimal to recommend adoption for all states between zero and one, but exclude a set of points like $\{1/23, 1/17, 1/6, 4/5\}$. As such mechanisms do not look reasonable, we restrict our attention to cut-off mechanisms. With this restriction the results from binary model carry over to an arbitrary number of states, producing similar persuasion strategy and identical minimax loss.

Appendix A: Proofs for Binary Model

A1: Notation

In this section we introduce notation, which we will use throughout the proofs. Denote

$$\begin{aligned}\underline{\phi}(\lambda) &= \underline{\alpha} + (1-\underline{\alpha})\lambda, \\ \overline{\phi}(\lambda) &= \overline{\alpha} + (1-\overline{\alpha})\lambda \end{aligned} \tag{16}$$

and

$$C(\lambda) = 1 - \ln \underline{\phi}(\overline{\lambda}) - \ln\left(\frac{\overline{\phi}(\lambda)}{\underline{\phi}(\lambda)}\right). \tag{17}$$

Moreover, we define functions

$$\underline{\mu}(\lambda) \equiv \max\left\{\underline{\lambda}, \frac{1}{1-\overline{\alpha}}\left(e^{-C(\lambda)} - \overline{\alpha}\right)\right\} \tag{18}$$

and

$$H(\lambda) \equiv \ln\left(\frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\lambda)}\right) - \max\{e^{-C(\lambda)}, \overline{\phi}(\underline{\lambda})\}\left(1 + \ln\left(\frac{1}{\overline{\phi}(\underline{\lambda})}\min\{e^{-C(\lambda)}, \overline{\phi}(\underline{\lambda})\}\right)\right). \tag{19}$$

A2: Proof of the Main Theorem

We are going to prove several Lemmata and finally proceed with the proof of the main result. We define the lower and upper bounds of the support of the strategy of the nature as $(\underline{\alpha}_F, \overline{\alpha}_F)$ and $(\underline{\lambda}_F, \overline{\lambda}_F)$. Note that our restrictions imply that $\underline{\alpha} \leq \underline{\alpha}_F \leq \overline{\alpha}_F \leq \overline{\alpha}$ and $\underline{\lambda} \leq \underline{\lambda}_F \leq \overline{\lambda}_F \leq \overline{\lambda}$. Correspondingly, the lower and the upper bounds of the support of the strategy of the sender are $(\underline{\mu}_G, \hat{\mu}_G)$. For a fixed strategy of the sender the expected loss of nature is given by

$$L_F(\alpha,\lambda) = \int_{\underline{\mu}_G}^{\hat{\mu}_G} L(\mu;\alpha,\lambda)dG(\mu)$$

The following Lemma states that for any λ nature chooses either $\alpha = \underline{\alpha}$ or $\alpha = \overline{\alpha}$. That is, with respect to the probability of the good state nature's strategy has widest support consisting of two points.

Lemma 5. *There is a function $\hat{\alpha}(\lambda) : [\underline{\lambda}_F, \overline{\lambda}_F] \to \{\underline{\alpha}, \overline{\alpha}\}$ such that for all $\alpha \in [\underline{\alpha}, \overline{\alpha}]$*

$$L_F(\hat{\alpha}(\lambda), \lambda) \geq L_F(\alpha, \lambda).$$

Moreover, there is $\lambda_0 \in [\underline{\lambda}_F, \overline{\lambda}_F)$ such that $\hat{\alpha}(\lambda) = \overline{\alpha}$ for all $\lambda < \lambda_0$ and $\hat{\alpha}(\lambda) = \underline{\alpha}$ for all $\lambda > \lambda_0$. Moreover, $lambda_0 > \underline{\mu}_G$.

Proof. Fix the strategy of the sender G. Then, the expected loss of nature is given by

$$L_F(\alpha,\lambda) =[1 - G(\lambda)](\alpha + (1-\alpha)\lambda) + (1-\alpha)\int_{\underline{\mu}_G}^{\lambda} (\lambda - \mu)dG(\mu),$$

which is linear in α for any λ. Thus, for given G and λ the expected loss of nature $L_F(\alpha, \lambda)$ is maximized by $\alpha \in \{\underline{\alpha}, \overline{\alpha}\}$. Moreover,

$$\frac{\partial L_F(\alpha,\lambda)}{\partial \alpha} = (1-\lambda)[1-G(\lambda)] - \int_{\underline{\mu}_G}^{\lambda} (\lambda-\mu)dG(\mu) = (1-\lambda) - \int_{\underline{\mu}_G}^{\lambda} (1-\mu)dG(\mu),$$

which is decreasing in λ. Thus, either α jumps downward from $\overline{\alpha}$ to $\underline{\alpha}$ at some interior point λ_0 or it equals to either $\underline{\alpha}$ or $\overline{\alpha}$ over the whole support (in which case $\lambda_0 = \underline{\lambda}_F$ or $\lambda_0 = \overline{\lambda}_F$ respectively).

Note that as at the point $\lambda = \underline{\mu}_G$ we have $\frac{\partial L_F}{\partial \alpha} > 0$ it must be the case that $\lambda_0 > \underline{\mu}_G$.

Moreover, at $\lambda = \overline{\lambda}_F$

$$\frac{\partial L_F(\alpha, \lambda)}{\partial \alpha} = (1 - \overline{\lambda}_F) - \int_{\underline{\mu}_G}^{\overline{\lambda}_F} (1 - \mu) dG(\mu) < 0$$

as playing $\mu > \overline{\lambda}_F$ is dominated by playing $\mu = \overline{\lambda}_F$ for the sender, so $G(\overline{\lambda}_F) = 1$.

$\square$

It is useful to show that nature does not play λ_0 with a positive probability.

Lemma 6. *If $\lambda_0 < \overline{\lambda}_F$, then in equilibrium $F(\lambda)$ is continuous at $\lambda = \lambda_0$.*

Proof. Suppose that nature plays λ_0 with positive probability. We show that the sender prefers $\mu = \lambda_0$ to $\mu = \lambda_0 + \varepsilon$. For any function $a(x)$ we define $a(x^-) = \sup\{a(t) : t < x\}$. Consider loss of the sender from choosing $\mu = \lambda_0$:

$$\int_{\underline{\lambda}_F}^{\lambda_0^-} [\overline{\alpha} + (1 - \overline{\alpha})\lambda] dF(\lambda) + [F(\lambda_0) - F(\lambda_0^-)]\alpha(\lambda_0) \times 0 + \int_{\lambda_0}^{\overline{\lambda}_F} (1 - \underline{\alpha})(\lambda - \lambda_0) dF(\lambda).$$

The expected loss from playing $\mu = \lambda_0 + \varepsilon$ equals to

$$\int_{\underline{\lambda}_F}^{\lambda_0^-} [\overline{\alpha} + (1 - \overline{\alpha})\lambda] dF(\lambda) + [F(\lambda_0) - F(\lambda_0^-)][\alpha(\lambda_0) + (1 - \alpha(\lambda_0))\lambda_0] +$$
$$+ \int_{\lambda_0}^{\lambda_0+\varepsilon} [\overline{\alpha} + (1 - \overline{\alpha})\lambda] dF(\lambda) + \int_{\lambda_0+\varepsilon}^{\overline{\lambda}_F} (1 - \underline{\alpha})(\lambda - \lambda_0 - \varepsilon) dF(\lambda).$$

The difference between the two losses is

$$[F(\lambda_0) - F(\lambda_0^-)][\alpha(\lambda_0) + (1 - \alpha(\lambda_0))\lambda_0] +$$
$$\int_{\lambda_0}^{\lambda_0+\varepsilon} [\overline{\alpha} + (1 - \overline{\alpha})\lambda - (1 - \underline{\alpha})(\lambda - \lambda_0)] dF(\lambda) - \int_{\lambda_0+\varepsilon}^{\overline{\lambda}_F} (1 - \underline{\alpha})\varepsilon dF(\lambda),$$

which is positive for ε small enough. Thus, the sender prefers not to play μ in some neighbourhood above λ_0. Then, any point in that neighbourhood gives nature a higher payoff than

λ_0, arriving to contradiction. $\square$

Lemma 5 defines $\hat{\alpha}(\cdot)$ at all points except λ_0. Lemma 6 states that the measure of this point is zero, so the minimax loss is unaffected by the choice of $\hat{\alpha}(\lambda_0)$. We define function at this point as $\hat{\alpha}(\lambda_0) = \underline{\alpha}$, so that $\hat{\alpha}$ is now uniquely defined on the full domain $[\underline{\alpha}, \overline{\alpha}]$. This allows us to redefine the nature's loss as

$$L_F(\lambda) = \int_{\underline{\mu}_G}^{\overline{\mu}_G} L(\mu, \hat{\alpha}(\lambda), \lambda) dG(\mu).$$

For a given strategy of nature we define the expected loss of the sender as

$$L_G(\mu) = \int_{\underline{\lambda}_F}^{\overline{\lambda}_F} L(\mu, \hat{\alpha}(\lambda), \lambda) dF(\lambda).$$

The following Lemma characterises supports of the equilibrium strategies.

Lemma 7. *In equilibrium,*

1. *$F(\lambda)$ is continuous on $[\underline{\lambda}_F, \overline{\lambda}_F)$;*

2. *$G(\mu)$ is continuous on $(\underline{\mu}_G, \overline{\mu}_G]$;*

3. *Distributions $G(\mu)$ and $F(\lambda)$ have the same compact support with $\underline{\lambda}_F = \underline{\mu}_G \equiv \underline{\mu}$ and $\overline{\lambda}_F = \overline{\mu}_G \equiv \overline{\mu}$. Moreover, $\overline{\mu} = \overline{\lambda}$.*

Proof. 1. Suppose, that the marginal distribution of the strategy of the nature has an atom at some λ_1, which is not the upper bound of its support. Suppose, that $\hat{\alpha}$ is constant (either $\underline{\alpha}$ or $\overline{\alpha}$) in the neighbourhood of λ_1. Then there exists $\varepsilon_1 > 0$ such that for all $\varepsilon \leq \varepsilon_1$ $L_G(\lambda_1) < L_G(\lambda_1 + \varepsilon)$. Thus the sender would not play μ just above λ_1 and $G(\lambda_1 + \varepsilon_1) - G(\lambda_1) = 0$. Then we have that $L_F(\lambda_1) < L_F(\lambda_1 + \varepsilon_1/2)$, as loss, defined by either case of equation (7), is increasing in λ and probability of having each case does not depend on λ if $\lambda \in (\lambda_1, \lambda_1 + \varepsilon_1)$. Contradiction. Now consider the case when

$\hat{\alpha}$ changes from $\overline{\alpha}$ to $\underline{\alpha}$ at λ_1. Again, as $(1-\overline{\alpha})(\lambda-\mu) < \underline{\alpha} + (1-\underline{\alpha})\lambda$ we have that $L_G(\lambda_1) < L_G(\lambda_1 + \varepsilon)$ and the proof follows the same steps as in the previous case.

2. Suppose that G has an atom at some point μ_1. Then, there exists $\varepsilon_1 > 0$ such that for all $\varepsilon \leq \varepsilon_1$ we have that

$$L_F(\mu_1) < L_F(\mu_1 - \varepsilon). \tag{20}$$

Now consider two cases. First, suppose that $\mu_1 = \overline{\lambda}_F$. From (20) we get that nature must play $\overline{\lambda}_F$ with probability zero, hence by playing μ_1 the sender never persuades the receiver, thus μ_1 cannot be played by the sender with positive probability. Second, suppose that $\mu_1 < \lambda_F$. From (20) it follows that points at and just above μ_1 are dominated from nature's perspective with points above μ_1, hence there exists $\varepsilon_2 > 0$ such that $F(\mu_1+\varepsilon_2)-F(\mu_1) = 0$. Then, we get that $L_G(\mu_1) > L_G(\mu_1+\varepsilon_2/2)$, regardless of whether $\hat{\alpha}$ is constant or jumps down from $\overline{\alpha}$ to $\underline{\alpha}$ at μ_1. Thus, the sender cannot play μ_1 with positive probability, a contradiction.

3. Suppose, that for some $a < b$ we have that $G(b) - G(a) = 0$ but $F(b) - F(a) > 0$. Then note that for any $c \in [a, b)$ we have that $L_F(b) > L_F(c)$ (as $\hat{\alpha}(b) \leq \hat{\alpha}(c)$), and thus no values in $[a, b)$ can be played with positive probability. Similarly, if $F(b) - F(a) = 0$, we have that the sender has lower losses at a than in any other point in the gap. Thus, supports must coincide. Now, if both supports have the same gap $[a, b)$ we have that $L_F((a + b)/2) > L_F(a)$, so there is a profitable deviation. Thus, both supports coincide and convex. Now, suppose that the upper bound of both supports $\overline{\mu} < \overline{\lambda}$. Note that $L_F(\overline{\lambda}) > L_F(\overline{\mu})$, thus there is a profitable deviation, hence $\overline{\mu} = \overline{\lambda}$.

□

Now we are ready to prove the main theorem.

Proof of Theorem 1. First, note that from Lemma 7 we have that $\overline{\mu} = \overline{\lambda}$, as well as the continuity of strategies in the interior of common support. Second, in the mixed strategy equilibrium both the sender and nature should be indifferent between all actions in the support of their distributions.

Denote the lower bound of the equilibrium support as $\underline{\mu} \geq \underline{\lambda}$. From Lemma 5 there exists $\lambda_0 \in (\underline{\mu}, \overline{\lambda}]$ such that $\hat{\alpha}(\lambda) = \overline{\alpha}$ for all $\lambda < \lambda_0$ and $\hat{\alpha}(\lambda) = \underline{\alpha}$ for all $\lambda \geq \lambda_0$. At the end of the proof we show that λ_0 is determined by equation (41).

Strategy of nature. We start by characterizing the strategy of the nature for given $\underline{\mu}$ and $\lambda_0 > \underline{\mu}$ (this inequality will be verified later using Lemma C.1). Consider the equilibrium loss of the sender. Using the indifference condition for $\mu < \lambda_0$ we obtain that

$$\begin{aligned}L_G(\mu) &= \int_{\lambda_0}^{\overline{\lambda}} (1-\underline{\alpha})(\lambda-\mu)dF(\lambda) + \int_{\mu}^{\lambda_0} (1-\overline{\alpha})(\lambda-\mu)dF(\lambda) + \int_{\underline{\mu}}^{\mu} [\overline{\alpha} + (1-\overline{\alpha})\lambda]dF(\lambda) \\ &= \int_{\lambda_0}^{\overline{\lambda}} (1-\underline{\alpha})(\lambda-\mu)dF(\lambda) + \int_{\underline{\mu}}^{\lambda_0} (1-\overline{\alpha})\lambda dF(\lambda) - \mu(1-\overline{\alpha})[F(\lambda_0)-F(\mu)] + \overline{\alpha}F(\mu).\end{aligned}$$

By setting $\mu = \underline{\mu}$ and using the fact that the sender must be indifferent across all μ we get

$$\begin{aligned}&\int_{\lambda_0}^{\overline{\lambda}} (1-\underline{\alpha})(\lambda-\mu)dF(\lambda) - \mu(1-\overline{\alpha})[F(\lambda_0)-F(\mu)] + \overline{\alpha}F(\mu) \\ &\qquad = \int_{\lambda_0}^{\overline{\lambda}} (1-\underline{\alpha})(\lambda-\underline{\mu})dF(\lambda) - \underline{\mu}(1-\overline{\alpha})F(\lambda_0),\end{aligned}$$

which simplifies to

$$[1-F(\lambda_0)](1-\underline{\alpha})(\mu-\underline{\mu}) + F(\lambda_0)(1-\overline{\alpha})(\mu-\underline{\mu}) = [\overline{\alpha} + (1-\overline{\alpha})\mu]F(\mu).$$

Thus,

$$F(\mu) = \frac{B(\mu-\underline{\mu})}{\overline{\alpha} + (1-\overline{\alpha})\mu}.$$

where $B = [1-F(\lambda_0)](1-\underline{\alpha}) + F(\lambda_0)(1-\overline{\alpha})$.

Next, we consider the case $\mu \geq \lambda_0$. The expected loss of the sender is given by

$$L_G(\mu) = \int_{\mu}^{\overline{\lambda}} (1-\underline{\alpha})(\lambda - \mu)dF(\lambda) + \int_{\lambda_0}^{\mu} [\underline{\alpha} + (1-\underline{\alpha})\lambda]dF(\lambda) + \int_{\underline{\mu}}^{\lambda_0} [\overline{\alpha} + (1-\overline{\alpha})\lambda]dF(\lambda).$$

By plugging in $\mu = \lambda_0$ and using the indifference we obtain:

$$\int_{\mu}^{\overline{\lambda}} (1-\underline{\alpha})(\lambda - \mu)dF(\lambda) + \int_{\lambda_0}^{\mu} [\underline{\alpha} + (1-\underline{\alpha})\lambda]dF(\lambda) = \int_{\lambda_0}^{\overline{\lambda}} (1-\underline{\alpha})(\lambda - \lambda_0)dF(\lambda),$$

which simplifies to

$$[1 - F(\lambda_0)](1-\underline{\alpha})(\mu - \lambda_0) = [\underline{\alpha} + (|1-\underline{\alpha})\mu]F(\mu),$$

which can be rewritten as

$$F(\mu) = 1 - \frac{A}{\underline{\alpha} + (1-\underline{\alpha})\mu}.$$

where $A = [1 - F(\lambda_0)][\underline{\alpha} + (1 - al)\lambda_0$ By using $\overline{\mu} = \overline{\lambda}$ we have that

$$F(\overline{\lambda}) = 1 - \frac{A}{\underline{\alpha} + (1-\underline{\alpha})\overline{\lambda}}.$$

Now, we use the result from Lemma 7 that the distribution function must be continuous at λ_0. This gives

$$F(\lambda_0) = \frac{B(\lambda_0 - \underline{\mu})}{\overline{\alpha} + (1-\overline{\alpha})\lambda_0} = 1 - \frac{A}{\underline{\alpha} + (1-\underline{\alpha})\lambda_0}.$$

Recall that $B = (1-\underline{\alpha}) - (\overline{\alpha} - \underline{\alpha})F(\lambda_0)$. Then, by solving for $F(\lambda_0)$ and A and using our notation $\underline{\phi}$ and $\overline{\phi}$ we obtain

$$A = \frac{\overline{\phi}(\underline{\mu})\underline{\phi}(\lambda_0)}{\underline{\phi}(\lambda_0) - \underline{\phi}(\underline{\mu}) + \overline{\phi}(\underline{\mu})}, \quad B = (1-\underline{\alpha})\frac{\overline{\phi}(\lambda_0)}{\underline{\phi}(\lambda_0) - \underline{\phi}(\underline{\mu}) + \overline{\phi}(\underline{\mu})} \tag{21}$$

and

$$F(\lambda_0) = \frac{\underline{\phi}(\lambda_0) - \underline{\phi}(\underline{\mu})}{\underline{\phi}(\lambda_0) - \underline{\phi}(\underline{\mu}) + \overline{\phi}(\underline{\mu})}.$$

It remains to show that $F(\lambda_0) \in [0,1]$. If λ_0 solves indeed solves equation (41), which we are going to verify later, then following the result of Lemma C.1 we have that $\lambda_0 > \underline{\mu}$. Then, using $\underline{\phi}' > 0$ we obtain

$$0 < A < \frac{\overline{\phi}(\underline{\mu})\underline{\phi}(\lambda_0)}{\overline{\phi}(\underline{\mu})} = \underline{\phi}(\lambda_0).$$

This implies that $F(\lambda_0) = 1 - \frac{A}{\underline{\phi}(\lambda_0)} \in (0,1)$.

Strategy of the sender. Next we derive the equilibrium strategy of the sender G. From Lemma 7 we know that G is continuous on $(\underline{\mu}, \overline{\lambda}]$. Consider the equilibrium loss of the nature that plays $\lambda \geq \lambda_0$[13]

$$\begin{aligned} L_F(\lambda) =& (1-G(\lambda))(\overline{\alpha} + (1-\overline{\alpha})\lambda) + \int_{\underline{\mu}}^{\lambda} (1-\overline{\alpha})(\lambda - \mu)dG(\mu) = \\ =& (1-G(\lambda))(\overline{\alpha} + (1-\overline{\alpha})\lambda) + (1-\overline{\alpha})\int_{\underline{\mu}}^{\lambda} G(\mu)d\mu. \end{aligned}$$

Similarly we can rewrite the loss for the case of $\lambda < \lambda_0$, so we get that

$$L_F(\lambda) = \begin{cases} (1-G(\lambda))(\overline{\alpha} + (1-\overline{\alpha})\lambda) + (1-\overline{\alpha})\int_{\underline{\mu}}^{\lambda} G(\mu)d\mu, & \lambda < \lambda_0 \\ (1-G(\lambda))(\underline{\alpha} + (1-\underline{\alpha})\lambda) + (1-\underline{\alpha})\int_{\underline{\mu}}^{\lambda} G(\mu)d\mu, & \lambda \geq \lambda_0 \end{cases}. \tag{22}$$

Consider the case of $\lambda < \lambda_0$. Define a differentiable function $J(\lambda) = \int_{\underline{\mu}}^{\lambda} G(\mu)d\mu$. Then, we can rewrite equation (22) as

$$[1 - J'(\lambda)]\overline{\phi}(\lambda) + (1-\overline{\alpha})J(\lambda) = \overline{L}. \tag{23}$$

Solving homogeneous equation $J'(\lambda)\overline{\phi}(\lambda) = (1-\overline{\alpha})J(\lambda)$ gives $J(\lambda) = c(\lambda)\overline{\phi}(\lambda)$. Plugging in

[13] Here we apply integration by parts to Riemann-Stiltjes integral, using the fact that $(\lambda - \mu)$ is continuous and G is non-decreasing, see Theorem 21.67 in Hewitt and Stromberg (2013).

back to the original equation gives $[1 - c'(\lambda)\overline{\phi}(\lambda)]\overline{\phi}(\lambda) = \overline{L}$. Thus,

$$c'(\lambda) = \frac{\overline{\phi}(\lambda) - \overline{L}}{[\overline{\phi}(\lambda)]^2},$$

which gives a solution (by replacing argument λ with μ):

$$G(\mu) = J'(\mu) = \ln \overline{\phi}(\mu) + C_0.$$

Similarly, for the case $\lambda \geq \lambda_0$ we obtain

$$G(\mu) = \ln \underline{\phi}(\mu) + C_1.$$

Using the continuity of G at the upper bound of the support, $G(\overline{\lambda}) = 1$, we get that $C_1 = 1 - \underline{\phi}(\overline{\lambda})$. To determine C_0 we use the fact that $G(\mu)$ is continuous at $\mu = \lambda_0$ as $\lambda_0 > \underline{\mu}$, so

$$\ln \overline{\phi}(\lambda_0) + C_0 = \ln \underline{\phi}(\lambda_0) + 1 - \ln \underline{\phi}(\overline{\lambda})$$

and therefore $C_0 = C(\lambda_0)$, where function $C(\cdot)$ is determined by equation (17). Thus, we obtain

$$G(\mu) = \begin{cases} \ln \overline{\phi}(\mu) + C(\lambda_0), & \mu < \lambda_0 \\ \ln \underline{\phi}(\mu) + 1 - \ln \underline{\phi}(\overline{\lambda}), & \mu \geq \lambda_0 \end{cases}.$$

The equilibrium lower bound. Next we are ready to characterize $\underline{\mu}$ as a function of λ_0. By substituting $\underline{\lambda}$ into the obtained function $G(\mu)$ we have that $G(\underline{\lambda}) > 0$ if and only if $\ln \overline{\phi}(\underline{\lambda}) + C(\lambda_0) > 0$ or equivalently when $\overline{\phi}(\underline{\lambda}) > e^{-C(\lambda_0)}$. In this case $\underline{\mu} = \underline{\lambda}$. Otherwise $\underline{\mu}$ is determined by the solution of $\overline{\phi}(\underline{\mu}) = e^{-C(\lambda_0)}$ and since $C' > 0$ it is easy to see that $\underline{\mu} \geq \underline{\lambda}$ where the equality occurs if and only if $\overline{\phi}(\underline{\lambda}) = e^{-C(\lambda_0)}$.

Summing up we established that the lower bound is determined by equation (18):

$$\underline{\mu} = \underline{\mu}(\lambda_0) = \max\left\{\underline{\lambda}, \frac{1}{1-\overline{\alpha}}\left(e^{-C(\lambda_0)} - \overline{\alpha}\right)\right\}.$$

The lower bound is defined as a function of λ_0 which we are going to characterise using Lemma 5. This Lemma also guarantees that $\underline{\mu}(\lambda_0) < \lambda_0$.

Characterization of λ_0. It remains to show that λ_0, i.e. the point where α jumps from the highest to the lowest possible value is defined by equation (41), i.e. $H(\lambda_0) = 0$. Then, by Lemma 5 such λ_0 always exists and is uniquely defined, and moreover, the lower bound is also correctly defined, so that $\underline{\mu}(\lambda_0) < \lambda_0 \leq \overline{\lambda}$.

Note that the loss of nature is defined by

$$L_F(\alpha, \lambda) = (1 - G(\lambda))(\alpha + (1-\alpha)\lambda) + (1-\alpha)\int_{\underline{\mu}}^{\lambda} G(\mu)d\mu.$$

Lemma 5 states that for $\lambda = \lambda_0$ nature is indifferent between all possible values of α (and prefers $\overline{\alpha}$ above and $\underline{\alpha}$ below this point, see equation 22). Thus, $\partial L_F/\partial\alpha = 0$ at $\lambda = \lambda_0$ which gives

$$(1 - G(\lambda_0))(1 - \lambda_0) = \int_{\underline{\mu}(\lambda_0)}^{\lambda_0} G(\mu)d\mu. \tag{24}$$

Let us first compute the left-hand side of (24)

$$\begin{aligned}(1 - G(\lambda_0))(1 - \lambda_0) &= (1 - (\ln\underline{\phi}(\lambda_0) + 1 - \ln\underline{\phi}(\overline{\lambda})))(1 - \lambda_0) = \\ &= (1 - \ln\overline{\phi}(\lambda_0) - C(\lambda_0))(1 - \lambda_0) = \\ &= (\ln\underline{\phi}(\overline{\lambda}) - \ln\underline{\phi}(\lambda_0))(1 - \lambda_0).\end{aligned}$$

The right-hand side of (24) can be rewritten as

$$\begin{aligned}
\int_{\underline{\mu}(\lambda_0)}^{\lambda_0} G(\mu)d\mu &= \int_{\underline{\mu}(\lambda_0)}^{\lambda_0} (\overline{\phi}(\mu) + C(\lambda_0))d\mu \\
&= \int_{\underline{\mu}(\lambda_0)}^{\lambda_0} (\overline{\phi}(\mu) + 1)d\mu - (1 - C(\lambda_0))(\lambda_0 - \underline{\mu}(\lambda_0)) \\
&= \frac{1}{1-\overline{\alpha}} \int_{\underline{\mu}}^{\hat{\mu}_0} d[\overline{\phi}(\mu) \ln \overline{\phi}(\mu)] - (1 - C(\lambda_0))(\lambda_0 - \underline{\mu}(\lambda_0)) \\
&= \frac{1}{1-\overline{\alpha}} \left[\overline{\phi}(\mu) \ln \overline{\phi}(\mu)\right] |_{\underline{\mu}(\lambda_0)}^{\lambda_0} - (1 - C(\lambda_0))(\lambda_0 - \underline{\mu}(\lambda_0)).
\end{aligned}$$

To characterize λ_0 we have to consider two cases determining the lower bound $\underline{\mu}(\lambda_0)$.

Case 1: $\overline{\phi}(\underline{\lambda}) > e^{-C(\lambda_0)}$. First, suppose that λ_0 satisfies $\overline{\phi}(\underline{\lambda}) > e^{-C(\lambda_0)}$ and $\underline{\mu}(\lambda_0) = \underline{\lambda}$. Then, equation (24) results in

$$(1 - C(\lambda_0) - \ln \overline{\phi}(\lambda_0))(1 - \lambda_0) = \frac{1}{1-\overline{\alpha}} \overline{\phi}(\lambda_0) \ln(\overline{\phi}(\lambda_0)) - \frac{\overline{\phi}(\underline{\lambda}) \ln \overline{\phi}(\underline{\lambda})}{1-\overline{\alpha}} - (1 - C(\lambda_0))(\lambda_0 - \underline{\lambda}).$$

By simplifying further we have that

$$\begin{aligned}
(1 - \underline{\lambda})(1 - C(\lambda_0)) &= \ln \overline{\phi}(\lambda_0) \left(1 - \hat{\mu}_0 + \frac{\overline{\alpha}}{1-\overline{\alpha}} + \hat{\mu}_0\right) - \frac{\overline{\phi}(\underline{\lambda}) \ln \overline{\phi}(\underline{\lambda})}{1-\overline{\alpha}} \\
&= \frac{1}{1-\overline{\alpha}} \ln \overline{\phi}(\lambda_0) - \frac{\overline{\phi}(\underline{\lambda}) \ln \overline{\phi}(\underline{\lambda})}{1-\overline{\alpha}}.
\end{aligned}$$

Notice that $(1 - \overline{\alpha})(1 - \underline{\lambda}) = (1 - \overline{\phi}(\underline{\lambda}))$. Using the formula for $C(\lambda_0)$ we come to the final equation

$$\ln\left(\frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\lambda_0)}\right) - \overline{\phi}(\underline{\lambda})(1 - C(\lambda_0) - \ln \overline{\phi}(\underline{\lambda})) = 0. \tag{25}$$

Notice that if $\overline{\phi}(\underline{\lambda}) > e^{-C(\lambda_0)}$, then equations (41) and (25) coincide.

Case 2: $\overline{\phi}(\underline{\lambda}) \leq e^{-C(\lambda_0)}$. Suppose that λ_0 satisfies $\overline{\phi}(\underline{\lambda}) \leq e^{-C(\lambda_0)}$ and $\underline{\phi}(\underline{\mu}(\lambda_0)) = e^{-C(\lambda_0)}$.

By plugging $\underline{\mu}(\lambda_0)$ into (24) we obtain:

$$\ln\left(\frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\lambda_0)}\right) - \overline{\phi}(\underline{\mu}(\lambda_0))(1 - C(\lambda_0) - \ln\overline{\phi}(\underline{\mu}(\lambda_0))) = 0.$$

Using that $\underline{\phi}(\underline{\mu}(\lambda_0)) = e^{-C(\lambda_0)}$ we have the final equation

$$\ln\left(\frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\lambda_0)}\right) - e^{-C(\lambda_0)} = 0. \tag{26}$$

Notice that if $\overline{\phi}(\underline{\lambda}) \leq e^{-C(\lambda_0)}$, then equations (41) and (26) coincide. So we showed that λ_0 is indeed determined by equation (41) and therefore according to Lemma C.1 there exists a unique λ_0 such that $\underline{\mu}(\lambda_0) < \lambda_0 \leq \overline{\lambda}$.

□

At the end of this section we explicitly present the equilibrium strategies derived in the proof of Theorem 1. For all values of parameters the sender's strategy which minimizes loss in the worst-case scenario is given by

$$G(\mu) = \begin{cases} \ln\overline{\phi}(\mu) + C(\lambda_0), & \mu < \lambda_0 \\ \ln\underline{\phi}(\mu) + 1 - \ln\underline{\phi}(\overline{\lambda}), & \mu \geq \lambda_0 \end{cases}, \tag{27}$$

with support $[\underline{\mu}(\lambda_0), \overline{\lambda}]$, where expressions for $\underline{\phi}(\cdot), \overline{\phi}(\cdot), C(\cdot), \underline{\mu}(\cdot)$ are given by equations (16) - (18) and λ_0 solves equation (41).

Corresponding strategy of the nature, which maximizes expected loss of the sender has the same support and is defined by

$$F(\lambda) = \begin{cases} \frac{B(\lambda-\underline{\mu}(\lambda_0))}{\overline{\phi}(\lambda)}, & \lambda < \lambda_0 \\ 1 - \frac{A}{\underline{\phi}(\lambda)}, & \lambda_0 \leq \lambda < \overline{\lambda} \\ 1 & \lambda \geq \overline{\lambda} \end{cases}, \tag{28}$$

where

$$A = \frac{\overline{\phi}(\underline{\mu}(\lambda_0))\underline{\phi}(\lambda_0)}{\underline{\phi}(\lambda_0) - \underline{\phi}(\underline{\mu}(\lambda_0)) + \overline{\phi}(\underline{\mu}(\lambda_0))} \quad \text{and} \quad B = (1-\underline{\alpha})\frac{\overline{\phi}(\lambda_0)}{\underline{\phi}(\lambda_0) - \underline{\phi}(\underline{\mu}(\lambda_0)) + \overline{\phi}(\underline{\mu}(\lambda_0))}.$$

Uniqueness. Note, that the mixed strategy equilibrium is uniquely defined for any support by constant expected payoff condition. Lemma 7 states that the upper bound of the support is uniquely defined. Then, for a given values of the upper bound and λ_0 the lower bound is uniquely defined by equation (18). Finally, the uniqueness of λ_0 is guaranteed by Lemma (C.1).

A3: Other Proofs

Proof of Lemma 1. Let $\beta < r$. Consider some mechanism $\{\mu_0, \mu_1\}$ and corresponding acceptance set A. Let $\tilde{\mu}$ be a mechanism which consists of two messages and sends message m^+ with probabilities $\tilde{\mu}_1(m^+) = \mu_1(A)$ and $\tilde{\mu}_0(m^+) = \mu_0(A)$ and message m^- with complimentary probabilities. For acceptance set, using (1), we have

$$\beta\mu_1(A) = \int_A P_\beta(m)d\mu_R(m) \geq r\mu_R(A)$$

Now, for the two message mechanism we have that

$$P_\beta(m^+) = \frac{\beta\tilde{\mu}_1(m^+)}{\beta\tilde{\mu}_1(m^+) + (1-\beta)\tilde{\mu}_0(m^+)} = \frac{\beta\mu_1(A)}{\mu_R(A)} \geq r$$

Therefore, message m^+ leads to acceptance. Thus, we can rewrite the sender's problem (3) as

$$\pi = \alpha\tilde{\mu}_1(m^+) + (1-\alpha)\tilde{\mu}_0(m^+) \qquad \text{s.t.} \quad P_\beta(m^+) \geq r$$

which gives the solution. Clearly, if $\beta \geq r$ the sender can reveal no information by sending m^+ with probability 1 in both states, leading to adoption. $\square$

Proof of Lemma 2. Suppose that the sender plays a mechanism $\mu_M = \{\mu_0, \mu_1\}$, where M is the message space. Nature randomises over parameters (α, β, r). We view the randomisation over (β, r) as a randomisation over the acceptance sets A defined in (2). Let $\mathcal{A}$ be a collection of all acceptance sets generated by different choices of (β, r). Let $\mathbb{F}$ be a strategy of the nature defined on all (α, β, r) and $F(A)$ be a corresponding 'marginal distribution' over $A \in \mathcal{A}$.

Let $M' = \cup_{\mathcal{A}} A$ – that is, M' is a set of messages which could potentially persuade at least some receiver. Let $\mathcal{M}'$ be a sigma-algebra on M'. We claim that $\mu_0 << \mu_1$ (μ_0 is absolutely continuous with respect to μ_1 on $\mathcal{M}'$). Indeed, if there exists a set $B \in \mathcal{M}'$ such that $\mu_0(B) > 0$ and $\mu_1(B) = 0$ then $\forall A \in \mathcal{A}$ we have $B \not\subset A$ as any $m \in B$ reveals that $\omega = 0$.

We define

$$q(m) = \int_{\mathcal{A}} \mathbb{I}(m \in A) dF(A)$$

and

$$\alpha^e(m) = \mathbb{E}(\alpha | m \in M') = \frac{1}{q(m)} \int_{\mathcal{A} \times [\underline{\alpha}, \overline{\alpha}]} \alpha \mathbb{I}(m \in A) d\mathbb{F}.$$

We can write down the expected payoff of the sender as

$$\begin{aligned}
\mathbb{E}_{\mathbb{F}} \pi(\mu_M; \alpha, \beta, r) &= E_{\mathbb{F}} \left[\alpha \int_M \mathbb{I}(m \in A) d\mu_1 + (1 - \alpha) \int_M \mathbb{I}(m \in A) d\mu_0 \right] \\
&= E_{\mathbb{F}} \left[\alpha \int_{M'} \mathbb{I}(m \in A) d\mu_1 + (1 - \alpha) \int_{M'} \mathbb{I}(m \in A) d\mu_0 \right] \\
&= \int_{M'} \alpha^e(m) q(m) d\mu_1 + \int_{M'} [1 - \alpha^e(m)] q(m) d\mu_0 \\
&= \int_{M'} \left(\alpha^e(m) + [1 - \alpha^e(m)] \frac{d\mu_0}{d\mu_1}(m) \right) q(m) d\mu_1,
\end{aligned}$$

where $\frac{d\mu_0}{d\mu_1}(m)$ is a Radon-Nikodym derivative of μ_0 with respect ot μ_1. We denote $t(m) = \frac{d\mu_0}{d\mu_1}(m)$. It is straightforward to verify that $t(m) \leq 1$ for any $m \in M'$.

Consider a binary mechanism $\mu^{t(m)}$ with two messages m^+ and m^-, where message m^+ is sent with probability 1 if $\omega = 1$ and with probability $t(m)$ if $\omega = 0$. Note that if m was in the acceptance set of some receiver under old mechanism, then m^+ is in the acceptance set under new mechanism. Indeed, the posterior of the receiver following message m^+ is

$$\tilde{p}(m) = \frac{\beta}{\beta + (1-\beta)t(m)}.$$

Consider a receiver with a corresponding acceptance set A. Now, for *any* set $C \subseteq A$ we have

$$\begin{aligned}
\int_C (\tilde{p}(m) - r)d\mu_R &= \int_C \tilde{p}(m)d\mu_R - r\mu_R(C) \\
&= \int_C \frac{\beta}{\beta + (1-\beta)t(m)} \left(\beta + (1-\beta)\frac{d\mu_0}{d\mu_1}(m)\right) d\mu_1 - r\mu_R(C) \\
&= \beta\mu_1(C) - r\mu_R(C) \\
&= \int_C P_\beta(m)d\mu_R - r\mu_R(C) \\
&= \int_C (P_\beta(m) - r)d\mu_R \geq 0.
\end{aligned}$$

As this must hold for any set C we conclude that $\tilde{p}(m) \geq r$ almost everywhere. Then, the expected payoff of such binary mechanism is

$$\pi(\mu^{t(m)}; \alpha, \beta, r) = E_{\mathbb{F}}[\alpha + (1-\alpha)t(m)]\mathbb{I}(m \in A) = \left\{\alpha^e(m) + [1-\alpha^e(m)]t(m)\right\} q(m)$$

Thus, we can rewrite the expected payoff from an arbitrary mechanism as

$$\mathbb{E}_{\mathbb{F}}\pi(\mu_M; \alpha, \beta, r) = \int_{M'} \mathbb{E}_{\mathbb{F}}\pi(\mu^{t(m)}; \alpha, \beta, r)d\mu_1 = \int_M \mathbb{E}_{\mathbb{F}}\pi(\mu^{t(m)}; \alpha, \beta, r)d\mu_1. \tag{29}$$

Therefore, a payoff of each mechanism can be represented as a randomisation over binary mechanisms. Let $\mathcal{G}$ be the set of all probability measures over persuasion mechanisms, and $\mathcal{G}_2$ be the set of all probability measures over binary message mechanisms. From (29) we obtain that

$$\sup_{\mu_M \in \mathcal{G}} \mathbb{E}_{\mathbb{F}} \pi(\mu_M; \alpha, \beta, r) \leq \sup_{\mu_M \in \mathcal{G}_2} \mathbb{E}_{\mathbb{F}} \pi(\mu_M; \alpha, \beta, r).$$

Moreover, as $\mathcal{G}_2 \subset \mathcal{G}$ we obtain that

$$\sup_{\mu_M \in \mathcal{G}} \mathbb{E}_{\mathbb{F}} \pi(\mu_M; \alpha, \beta, r) \geq \sup_{\mu_M \in \mathcal{G}_2} \mathbb{E}_{\mathbb{F}} \pi(\mu_M; \alpha, \beta, r).$$

These two inequalities imply that

$$\sup_{\mu_M \in \mathcal{G}} \mathbb{E}_{\mathbb{F}} \pi(\mu_M; \alpha, \beta, r) = \sup_{\mu_M \in \mathcal{G}_2} \mathbb{E}_{\mathbb{F}} \pi(\mu_M; \alpha, \beta, r).$$

Thus,

$$\inf_{\mu_M \in \mathcal{G}} \mathbb{E}_{\mathbb{F}} L(\mu_M; \alpha, \beta, r) = \inf_{\mu_M \in \mathcal{G}_2} \mathbb{E}_{\mathbb{F}} L(\mu_M; \alpha, \beta, r).$$

□

Proof of Proposition 1. Define the sender's loss from playing strategy $G' \in \mathcal{G}$ when nature plays $F \in \mathcal{F}$ as

$$\mathcal{L}(G', F') = \int \int L(\mu_M; \alpha, \beta, r) dF' dG'.$$

We say that $G', G'' \in \mathcal{G}$ are payoff equivalent for some $F' \in \mathcal{F}$ if

$$\mathcal{L}(G', F') = \mathcal{L}(G'', F').$$

We show that the set of equilibria consists of pairs (G', F), where G' is payoff equivalent to G.

First, we show the uniqueness of nature strategy F. Suppose for a contradiction that (G', F') is an equilibrium and $F' \neq F$. Theorem 1 implies that if $G' \in \mathcal{G}_2$, where $\mathcal{G}_2$ is the set of all probability measures over binary message mechanisms, then $G' = G$ and $F' = F$.

Suppose that G' does not belong to $\mathcal{G}_2$. Then, Lemma 2 implies that there exists a sender's strategy $G_2 \in \mathcal{G}_2$ such that

$$\mathcal{L}(G', F') = \mathcal{L}(G_2, F').$$

If $G_2 = G$, then by Theorem 1 (G', F') is an equilibrium if and only if $F' = F$. Otherwise, if $G_2 \neq G$, then (G_2, F') is not an equilibrium (as this would contradict uniqueness of equilibrium in the class of binary message mechanism strategies) – that is, either sender or nature has a profitable deviation. This, in turn, implies that (G', F') is not an equilibrium either, a contradiction.

It remains to show that (G', F) is an equilibrium if and only if G' is payoff equivalent to G. First, we prove necessity. Suppose that (G', F) is an equilibrium and G' is not payoff equivalent to G. Then, by Lemma 2 there exists some $G_2 \in \mathcal{G}$, $G_2 \neq G$, such that $\mathcal{L}(G', F) = \mathcal{L}(G_2, F)$. Since $G_2 \neq G$, then by Theorem 1 either the sender or nature has a profitable deviation implying that (G', F) is not an equilibrium, a contradiction. Second, we show sufficiency. Suppose that G' is payoff equivalent to G. Since (G, F) is an equilibrium, then (G', F) is also an equilibrium. $\square$

Proof of Lemma 3. The equilibrium losses are constant for all strategies of the nature played with positive probability. By plugging $\lambda = \overline{\lambda}$ into (22) and using the equilibrium strategy of the sender defined in (27) for $\mu > \lambda_0$ we have that

$$\begin{aligned}
\overline{L} &= (1 - \underline{\alpha}) \int_{\underline{\mu}(\lambda_0)}^{\overline{\lambda}} G(\mu) d\mu \\
&= (1 - \underline{\alpha}) \int_{\underline{\mu}(\lambda_0)}^{\lambda_0} G(\mu) d\mu + (1 - \underline{\alpha}) \int_{\lambda_0}^{\overline{\lambda}} (\underline{\phi}(\mu) + 1 - \ln \overline{\lambda}) d\mu.
\end{aligned}$$

By using the indifference condition (24) for the first integral and simplifying further we obtain the final formula for the equilibrium losses

$$
\begin{aligned}
\overline{L} &= (1-\underline{\alpha})(1-G(\lambda_0))(1-\lambda_0) + \int_{\hat{\mu}_0}^{\overline{\lambda}} d[\underline{\phi}(\mu)\ln\underline{\phi}(\mu)] - (\underline{\phi}(\overline{\lambda}) - \underline{\phi}(\lambda_0))\ln\overline{\lambda} \\
&= (1-\underline{\phi}(\lambda_0))(\ln\overline{\lambda} - \ln\underline{\phi}(\lambda_0)) + \left[\underline{\phi}(\mu)\ln\underline{\phi}(\mu)\right]\Big|_{\lambda_0}^{\overline{\lambda}} - (\underline{\phi}(\overline{\lambda}) - \underline{\phi}(\lambda_0))\ln\overline{\lambda} \\
&= \ln\underline{\phi}(\overline{\lambda}) - \ln\underline{\phi}(\lambda_0).
\end{aligned}
$$

□

Proof of Proposition 2. Part 1. Proof of this statement follows from Lemmata C.2-C.6, which derive partial derivatives of function H with respect to all the parameters and λ_0.

The derivative of minimax loss (9) with respect to $\overline{\lambda}$:

$$
\begin{aligned}
\frac{d\overline{L}}{d\overline{\lambda}} =& \frac{1-\underline{\alpha}}{\underline{\phi}(\overline{\lambda})} - \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)}\frac{\partial H}{\partial\overline{\lambda}} \Big/ \left(-\frac{\partial H}{\partial\lambda_0}\right) \geq \\
\geq& \frac{1-\underline{\alpha}}{\underline{\phi}(\overline{\lambda})} - \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)}\left(\frac{(1-\underline{\alpha})}{\underline{\phi}(\overline{\lambda})}(1-\overline{\phi}(\underline{\lambda}))\right) \Big/ \left(-\frac{\partial H}{\partial\lambda_0}\right) > \\
>& \frac{1-\underline{\alpha}}{\underline{\phi}(\overline{\lambda})} - \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)}\left(\frac{(1-\underline{\alpha})}{\underline{\phi}(\overline{\lambda})}(1-\overline{\phi}(\underline{\lambda}))\right) \Big/ \left(\frac{1-\overline{\alpha}}{\underline{\phi}(\lambda_0)}\right),
\end{aligned}
$$

where in the last inequality we used the fact that $-\frac{\partial H}{\partial\lambda_0} > \frac{1-\overline{\alpha}}{\underline{\phi}(\lambda_0)}$ (Lemma C.2). By simplifying further we obtain that

$$
\frac{d\overline{L}}{d\overline{\lambda}} > \frac{1-\underline{\alpha}}{\underline{\phi}(\overline{\lambda})}\left(1 - \frac{1-\underline{\alpha}}{1-\overline{\alpha}}(1-\overline{\phi}(\underline{\lambda}))\right) = \frac{1-\underline{\alpha}}{\underline{\phi}(\overline{\lambda})}\left(1-(1-\underline{\alpha})(1-\underline{\lambda})\right) = (1-\underline{\alpha})\frac{\underline{\phi}(\underline{\lambda})}{\underline{\phi}(\overline{\lambda})} > 0. \quad (30)
$$

The derivative of minimax loss (9) with respect to $\underline{\lambda}$:

$$
\frac{d\overline{L}}{d\underline{\lambda}} = -\frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)}\frac{\partial\lambda_0}{\partial\underline{\lambda}}.
$$

From Lemma C.2 we have $\frac{\partial H}{\partial \lambda_0} < 0$ and by Lemma C.4 we get $\frac{\partial H}{\partial \underline{\lambda}} \geq 0$, implying that $\frac{\partial \lambda_0}{\partial \underline{\lambda}} \geq 0$ and therefore $\frac{\partial \overline{L}}{\partial \underline{\lambda}} \leq 0$.

The derivative of minimax loss (9) with respect to $\overline{\alpha}$:

$$\frac{d\overline{L}}{d\overline{\alpha}} = -\frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)} \frac{\partial \lambda_0}{\partial \overline{\alpha}}.$$

From Lemma C.2 we have $\frac{\partial H}{\partial \lambda_0} < 0$ and by Lemma C.5 we get $\frac{\partial H}{\partial \underline{\alpha}} < 0$, implying that $\frac{\partial \lambda_0}{\partial \overline{\alpha}} < 0$ and therefore $\frac{\partial \overline{L}}{\partial \overline{\alpha}} > 0$.

The derivative of minimax loss (9) with respect to $\underline{\alpha}$:

$$\begin{aligned}
\frac{d\overline{L}}{d\underline{\alpha}} &= -\frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)} + \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)} \left(-\frac{\partial H}{\partial \underline{\alpha}}\right) \Big/ \left(-\frac{\partial H}{\partial \lambda_0}\right) \leq \\
&\leq -\frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)} + \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)} \left(\frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)}(1-\overline{\phi}(\underline{\lambda}))\right) \Big/ \left(-\frac{\partial H}{\partial \lambda_0}\right) < \\
&< -\frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)} + \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)} \left(\frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)}(1-\overline{\phi}(\underline{\lambda}))\right) \Big/ \left(\frac{1-\overline{\alpha}}{\underline{\phi}(\lambda_0)}\right),
\end{aligned}$$

where the first inequality follows from Lemma C.6 and the second inequality follows from Lemma C.2. By simplifying further we obtain

$$\begin{aligned}
\frac{d\overline{L}}{d\underline{\alpha}} &< \frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)} \left(-1 + \frac{1-\underline{\alpha}}{1-\overline{\alpha}}(1-\overline{\phi}(\underline{\lambda}))\right) \\
&= \frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)} \left(-1 + (1-\underline{\alpha})(1-\underline{\lambda})\right) < -\frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\lambda_0)} < 0. \qquad (31)
\end{aligned}$$

Part 2. We derive a minimal upper bound on the equilibrium losses $\overline{L}$ for all parameters $\underline{\alpha} \leq \overline{\alpha} \in [0,1)$ and $\underline{\lambda} \leq \overline{\lambda} \in [0,1]$.

From Part 1 we have that the losses are higher for a wider range of parameters, so the supremum of losses is reached when, $\underline{\lambda} = 0$, $\overline{\lambda} = 1$, $\underline{\alpha}$ tends to 0 and $\overline{\alpha}$ tends to 1. Thus,

$$\sup \overline{L} = \lim_{\overline{\alpha}\to 1} \overline{L} = -\lim_{\overline{\alpha}\to 1} \ln \lambda_0.$$

From Lemma C.2 and C.5 we have $\frac{\partial \lambda_0}{\partial \overline{\alpha}} < 0$.

First, we rule out that in a small neighborhood of $\overline{\alpha} = 1$ we have that $\lambda_0 < \frac{1}{e}$. Suppose it is the case. Then, in this neighborhood we have that $\underline{\mu}(\lambda_0) > 0$ and the equation (41) simplifies to

$$H(\lambda_0) = -\ln \lambda_0 - \frac{\overline{\phi}(\lambda_0)}{e\lambda_0} = 0.$$

Note that $\lim_{\overline{\alpha}\to 1} \overline{\phi}(\lambda_0) = 1$, which implies that

$$0 = \lim_{\overline{\alpha}\to 1} \left(\frac{H(\lambda_0)}{\overline{\phi}(\lambda_0)}\right) = \lim_{\overline{\alpha}\to 1} \left(-\frac{\ln \lambda_0}{\overline{\phi}(\lambda_0)}\right) - \frac{1}{e\lambda_0} < 1 - \frac{1}{e\lambda_0}.$$

We obtain that $\lambda_0 > \frac{1}{e}$ and arrive to a contradiction. Therefore it must be the case that in a small neighborhood of $\overline{\alpha} = 1$ $\lambda_0 \geq \frac{1}{e}$ and $\underline{\mu}(\lambda_0) = 0$. In this case the equation (41) simplifies to

$$H(\lambda_0) = -(1-\overline{\alpha}) \ln \lambda_0 - \overline{\alpha} \ln(\overline{\alpha} + (1-\overline{\alpha})\lambda_0) + \overline{\alpha} \ln \overline{\alpha} = 0.$$

By dividing this equation by $1 - \overline{\alpha}$ and taking the limit as $\overline{\alpha} \to 1$ we obtain that

$$-\lim_{\overline{\alpha}\to 1} \ln \lambda_0 = \lim_{\overline{\alpha}\to 1} \frac{\overline{\alpha} \ln(\overline{\alpha} + (1-\overline{\alpha})\lambda_0)}{1-\overline{\alpha}} - \lim_{\overline{\alpha}\to 1} \frac{\overline{\alpha} \ln \overline{\alpha}}{1-\overline{\alpha}}. \tag{32}$$

By applying the L'Hopitals's rule we obtain that

$$\lim_{\overline{\alpha}\to 1} \frac{\overline{\alpha} \ln(\overline{\alpha} + (1-\overline{\alpha})\lambda_0)}{1-\overline{\alpha}} = -(1-\lambda_0)$$

and

$$\lim_{\overline{\alpha}\to 1} \frac{\overline{\alpha}\ln\overline{\alpha}}{1-\overline{\alpha}} = -1.$$

Thus, equation (32) can be rewritten as

$$\lim_{\overline{\alpha}\to 1}(\lambda_0 + \ln\lambda_0) = 0.$$

Therefore, we obtain that $\lim_{\overline{\alpha}\to 1}\lambda_0 = \mathbf{\Omega} \approx 0.5671$ and

$$\sup\overline{L} = -\lim_{\overline{\alpha}\to 1}\ln\lambda_0 = \mathbf{\Omega}.$$

Part 3. Note that if $\underline{\lambda} = \overline{\lambda} = \lambda$, then the sender can choose $\mu = \lambda$ making loss given by (7) equal to zero regardless of nature strategy with respect to α.

Part 4. Suppose that $\underline{\alpha} = \overline{\alpha} = \alpha$. Then,

$$\underline{\phi}(x) = \overline{\phi}(x) = \phi(x) \equiv \alpha + (1-\alpha)x.$$

The equilibrium loss can be rewritten as

$$\overline{L} = \phi(\underline{\mu})\ln\left(\frac{\phi(\overline{\lambda})}{\phi(\underline{\mu})}\right).$$

First, suppose that $\underline{\mu} = \frac{1}{1-\alpha}\left(\frac{\phi(\overline{\lambda})}{e} - \alpha\right)$ (that is, $\overline{\lambda} > e\underline{\lambda}$). Then $\phi(\underline{\mu}) = \frac{\phi(\overline{\lambda})}{e}$. Thus,

$$\overline{L} = \frac{\phi(\overline{\lambda})}{e}.$$

Second, suppose that $\underline{\mu} = \underline{\lambda}$. Then,

$$\overline{L} = \phi(\underline{\lambda}) \ln \left(\frac{\phi(\overline{\lambda})}{\phi(\underline{\lambda})} \right). \tag{33}$$

The partial derivative of $\overline{L}$ with respect to α is given by

$$\frac{\partial \overline{L}}{\partial \alpha} = (1 - \overline{\lambda}) \ln \frac{\phi(\overline{\lambda})}{\phi(\underline{\lambda})} - \frac{\overline{\lambda} - \underline{\lambda}}{\phi(\overline{\lambda})}.$$

Note that

$$\frac{\partial^2 \overline{L}}{\partial \alpha^2} = -\frac{(\overline{\lambda} - \underline{\lambda})^2}{\phi(\underline{\lambda})[\phi(\overline{\lambda})]^2}.$$

Moreover, $\frac{\partial \overline{L}}{\partial \alpha}\big|_{\alpha=1} < 0$. Thus, $\frac{\partial \overline{L}}{\partial \alpha}$ is decreasing in α and turns negative for large values of α. Thus, the maximum loss is attained at some $\tilde{\alpha} > \hat{\alpha}_0$ where $\hat{\alpha}_0$ solves $\underline{\mu} = \underline{\lambda}$ or $\phi(\underline{\lambda}) = \frac{\phi_0(\overline{\lambda})}{e}$ – that is,

$$\hat{\alpha}_0 = \max \left\{ 0, \frac{\overline{\lambda} - e\underline{\lambda}}{\overline{\lambda} - e\underline{\lambda} + e - 1} \right\}.$$

Combining both cases, we have that loss is linearly increases for $\alpha \leq \hat{\alpha}_0$, increases and then decreases for $\alpha > \hat{\alpha}_0$, which gives the result in proposition.

Finally, note that if $\overline{\lambda} = 1$ loss in the first case is constant and equals $1/e$, while in the second case $L = -\phi(\underline{\lambda}) \ln[\phi(\underline{\lambda})] \leq 1/e$, which completes the proof.

□

Proof of Proposition 3. Part 1. We first consider the impact of small uncertainty in λ. That is, let $\alpha \in [\underline{\alpha}, \overline{\alpha}]$ and let $\lambda \in [\hat{\lambda} - \varepsilon/2, \hat{\lambda} + \varepsilon/2]$ where $\hat{\lambda}$ is the initial value and $\varepsilon \to 0$. Note that from equation (17) we obtain

$$\lim_{\varepsilon \to 0} C(\lambda_0) = \lim_{\varepsilon \to 0} [1 - \ln \underline{\phi}(\underline{\lambda}) - \ln \overline{\phi}(\lambda_0) + \ln \underline{\phi}(\lambda_0)] = 1 - \ln \overline{\phi}(\hat{\lambda}).$$

Therefore,

$$\lim_{\varepsilon\to 0} e^{-C(\lambda_0)} = \frac{1}{e}\overline{\phi}(\hat{\lambda}) < \overline{\phi}(\hat{\lambda}) = \lim_{\varepsilon\to 0}\overline{\phi}(\underline{\lambda}).$$

Therefore, for ε small enough we are in the case $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$.

Equation (9) states that $\overline{L} = \ln\frac{\phi(\overline{\lambda})}{\underline{\phi}(\lambda_0)}$ and therefore

$$\frac{\partial\overline{L}}{\partial\varepsilon} = \frac{1-\underline{\alpha}}{\underline{\phi}(\overline{\lambda})}\frac{1}{2} - \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)}\frac{d\lambda_0}{d\varepsilon}. \tag{34}$$

Next, we proceed with deriving $\frac{d\lambda_0}{d\varepsilon}$. Using Lemmata C.3 and C.4 we obtain

$$\begin{aligned}\frac{\partial H}{\partial\varepsilon} &= \frac{1}{2}\left(\frac{\partial H}{\partial\overline{\lambda}} - \frac{\partial H}{\partial\underline{\lambda}}\right)\\ &= \frac{1}{2}\left(\frac{1-\underline{\alpha}}{\underline{\phi}(\overline{\lambda})}(1-\overline{\phi}(\underline{\lambda})) - (1-\overline{\alpha})(C(\lambda_0)+\ln\overline{\phi}(\underline{\lambda}))\right),\end{aligned}$$

which gives

$$\lim_{\varepsilon\to 0}\frac{\partial H}{\partial\varepsilon} = \frac{1}{2}\frac{(1-\overline{\alpha})[2(1-\underline{\phi}(\hat{\lambda}))-1]}{\underline{\phi}(\hat{\lambda})}.$$

From Lemma C.2 we have that

$$\frac{\partial H}{\partial\lambda_0} = \frac{\overline{\alpha}-\underline{\alpha}}{\overline{\phi}(\lambda_0)\underline{\phi}(\lambda_0)}\overline{\phi}(\underline{\lambda}) - \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)},$$

which gives $\lim_{\varepsilon\to 0}\frac{\partial H}{\partial\lambda_0} = -\frac{1-\overline{\alpha}}{\underline{\phi}(\hat{\lambda})}$. As this limit is bounded away form zero, we get

$$\lim_{\varepsilon\to 0} = \frac{d\lambda_0}{d\varepsilon} = \frac{1}{2} - \underline{\phi}(\hat{\lambda}).$$

By plugging this result into (34) we obtain the final result

$$\lim_{\varepsilon\to 0}\frac{\partial\overline{L}}{\partial\varepsilon} = \frac{1-\underline{\alpha}}{\underline{\phi}(\overline{\lambda})}\frac{1}{2} - \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)}\left[\frac{1}{2} - \underline{\phi}(\hat{\lambda})\right] = 1-\underline{\alpha}.$$

Note that from part 3 of Proposition 2 we have that $\overline{L}(\varepsilon = 0) = 0$. Therefore in a small neighborhood of $\varepsilon = 0$ we have that

$$\overline{L} = (1 - \underline{\alpha})\varepsilon + o(\varepsilon).$$

Part2. Now we derive the behaviour of $\overline{L}$ in a neighbourhood of some $\hat{\alpha}$. We will separately consider three cases in which *i)* $\hat{\alpha} < \hat{\alpha}_0$, *ii)* $\hat{\alpha} > \hat{\alpha}_0$ and *iii)* $\hat{\alpha} = \hat{\alpha}_0$, where $\hat{\alpha}_0$ is defined by

$$\hat{\alpha}_0 \equiv \max\left\{0, \frac{\overline{\lambda} - e\underline{\lambda}}{\overline{\lambda} - e\underline{\lambda} + e - 1}\right\}, \tag{35}$$

i.e. solves $e(\hat{\alpha} + (1 - \hat{\alpha})\underline{\lambda}) = \hat{\alpha} + (1 - \hat{\alpha})\overline{\lambda}$. We also introduce notation

$$\hat{\phi}(x) = \hat{\alpha} + (1 - \hat{\alpha})x.$$

Case 1: $\hat{\alpha} < \hat{\alpha}_0$. Note that in this case $\lim_{\varepsilon\to 0} C(\lambda_0) = 1 - \ln\hat{\phi}(\overline{\lambda})$, and thus $\lim_{\varepsilon\to 0} e^{-C(\lambda_0)} = \frac{1}{e}\hat{\phi}(\overline{\lambda}) < \hat{\phi}(\underline{\lambda})$, where the last inequality follows form $\hat{\alpha} < \hat{\alpha}_0$. Next, we proceed with deriving $\frac{d\lambda_0}{d\varepsilon}$. Using Lemmata C.5 and C.6 we obtain

$$\begin{aligned}
\frac{\partial H}{\partial \varepsilon} &= \frac{1}{2}\left(\frac{\partial H}{\partial \overline{\alpha}} - \frac{\partial H}{\partial \underline{\alpha}}\right) \\
&= \frac{1}{2}\left(-e^{-C(\lambda_0)}\frac{1 - \lambda_0}{\overline{\phi}(\lambda_0)} + \frac{\overline{\lambda} - \lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)}(1 - e^{-C(\lambda_0)})\right).
\end{aligned}$$

From Lemma C.2 we have

$$\frac{\partial H}{\partial \lambda_0} = e^{-C(\lambda_0)}\frac{\overline{\alpha} - \underline{\alpha}}{\overline{\phi}(\lambda_0)\underline{\phi}(\lambda_0)} - -\frac{1 - \underline{\alpha}}{\underline{\phi}(\lambda_0)}$$

and therefore

$$\frac{d\lambda_0}{d\varepsilon} = \frac{1}{2}\left(-e^{-C(\lambda_0)}\frac{1-\lambda_0}{\overline{\phi}(\lambda_0)} + \frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)}(1-e^{-C(\lambda_0)})\right)\Bigg/\left(\frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)} - e^{-C(\lambda_0)}\frac{\overline{\alpha}-\underline{\alpha}}{\overline{\phi}(\lambda_0)\underline{\phi}(\lambda_0)}\right).$$

Using the fact that $\lim_{\varepsilon\to 0} C(\lambda_0) = \hat{\phi}(\overline{\lambda})$ we obtain

$$\lim_{\varepsilon\to 0}\frac{d\lambda_0}{d\varepsilon} = \frac{1}{2(1-\hat{\alpha})}\left(\frac{\overline{\lambda}-\lambda_0}{\hat{\phi}(\overline{\lambda})} - \frac{1}{e}\left[\hat{\phi}(\overline{\lambda})(1-\lambda_0)+\overline{\lambda}-\lambda_0\right]\right)$$

Recall that $\overline{L} = \frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\lambda_0)}$, see equation (9). Thus,

$$\begin{aligned}\frac{d\overline{L}}{d\varepsilon} &= \frac{1}{2}\frac{\partial\overline{L}}{\partial\overline{\alpha}} - \frac{1}{2}\frac{\partial\overline{L}}{\partial\underline{\alpha}} + \frac{\partial\overline{L}}{\partial\lambda_0}\frac{d\lambda_0}{d\varepsilon}\\ &= \frac{1}{2}\frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)} - \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)}\frac{d\lambda_0}{d\varepsilon}.\end{aligned}$$

Note that the limit of equation (19) is

$$\lim_{\varepsilon\to 0}\ln\hat{\phi}(\overline{\lambda}) - \ln\hat{\phi}(\lambda_0) - \frac{1}{e}\hat{\phi}(\overline{\lambda}),$$

which using $H(\lambda_0)=0$ gives

$$\hat{\phi}(\lambda_0) = \hat{\phi}(\overline{\lambda})e^{-\hat{\phi}(\overline{\lambda})/e}. \tag{36}$$

Now, we are ready to derive the limit of $d\overline{L}/d\varepsilon$

$$
\begin{aligned}
\lim_{\varepsilon\to 0}\frac{d\overline{L}}{d\varepsilon} &= \frac{1}{2}\frac{\overline{\lambda}-\lambda_0}{\hat{\phi}(\overline{\lambda})\hat{\phi}(\lambda_0)} - \frac{1}{2\hat{\phi}(\lambda_0)}\left(\frac{\overline{\lambda}-\lambda_0}{\hat{\phi}(\overline{\lambda})} - \frac{1}{e}\left[\hat{\phi}(\overline{\lambda})(1-\lambda_0)+\overline{\lambda}-\lambda_0\right]\right) \\
&= \frac{1}{2e\hat{\phi}(\lambda_0)}\left[\hat{\phi}(\overline{\lambda})(1-\lambda_0)+\overline{\lambda}-\lambda_0\right] \\
&= \frac{1}{2e\hat{\phi}(\lambda_0)(1-\hat{\alpha})}\left[2\hat{\phi}(\overline{\lambda})-(1+\hat{\phi}(\overline{\lambda}))\hat{\phi}(\lambda_0)\right] \\
&= \frac{1}{(1-\hat{\alpha})e}\left[e^{\hat{\phi}(\overline{\lambda})/e}-\frac{1+\hat{\phi}(\overline{\lambda})}{2}\right]. \qquad (37)
\end{aligned}
$$

We will show that limit (37) is lower than 1 for all $\hat{\alpha}\leq\hat{\alpha}_0$. To do so, we first show that (37) increases in α and attains its maximum at $\hat{\alpha}=\hat{\alpha}_0$. The partial derivative of (37) with respect to $\hat{\alpha}$ is higher than 0 since

$$
\frac{\partial}{\partial\alpha}\left[\frac{1}{(1-\hat{\alpha})e}\left(e^{\hat{\phi}(\overline{\lambda})/e}-\frac{1+\hat{\phi}(\overline{\lambda})}{2}\right)\right] = \frac{1}{e(1-\hat{\alpha})^2}\left(e^{\hat{\phi}(\overline{\lambda})/e}(e+1-\hat{\phi}(\overline{\lambda}))-e\right) > 0.
$$

This implies that if limit (37) is lower than 1 for $\hat{\alpha}=\hat{\alpha}_0$, then it is lower than 1 for all $\hat{\alpha}<\hat{\alpha}_0$. Since (37) depends on $\underline{\lambda}$ only through $\hat{\alpha}_0$ and $\hat{\alpha}_0$ given by (35) decreases in $\underline{\lambda}$, we have that expression (37) can be bounded from above by the case in which $\underline{\lambda}=0$. By plugging in $\hat{\alpha}=\hat{\alpha}_0|_{\underline{\lambda}=0}$ we obtain

$$
\begin{aligned}
\lim_{\varepsilon\to 0}\frac{d\overline{L}}{d\varepsilon}\Big|_{\hat{\alpha}=\hat{\alpha}_0} &= \frac{1}{(1-\hat{\alpha}_0)e}\left(e^{\hat{\phi}(\overline{\lambda})/e}-\frac{1+\hat{\phi}(\overline{\lambda})}{2}\right) \\
&= \frac{\overline{\lambda}+e-1}{(e-1)e}\left(e^{\hat{\phi}(\overline{\lambda})/e}-\frac{1+\hat{\phi}(\overline{\lambda})}{2}\right).
\end{aligned}
$$

By taking the firs and second derivatives of the limit obtain:

$$
\frac{\partial}{\partial\overline{\lambda}}\left(\lim_{\varepsilon\to 0}\frac{d\overline{L}}{d\varepsilon}\Big|_{\hat{\alpha}=\hat{\alpha}_0}\right) = \frac{2e^{\lambda/\lambda+e-1}(2e-2+\lambda)-(\lambda+e-1)(e-1)}{2(\overline{\lambda}+e-1)(e-1)e} \qquad (38)
$$

and

$$\frac{\partial^2}{\partial\overline{\lambda}^2}\left(\lim_{\varepsilon\to 0}\frac{d\overline{L}}{d\varepsilon}\Big|_{\hat{\alpha}=\hat{\alpha}_0}\right) = \frac{(e-1)e^{-\frac{e-1}{\lambda+e-1}}}{(\lambda+e-1)^3} \geq 0$$

Thus, we conclude that derivative (38) is largest when $\overline{\lambda} = 1$ and by plugging this value to (38) verify that

$$\frac{\partial}{\partial\overline{\lambda}}\left(\lim_{\varepsilon\to 0}\frac{d\overline{L}}{d\varepsilon}\Big|_{\hat{\alpha}=\hat{\alpha}_0}\right)\Bigg|_{\overline{\lambda}=1} > 0$$

Thus, we conclude that (37) reaches its maximum value for $\hat{\alpha} = \hat{\alpha}_0$ and $[\underline{\lambda}, \overline{\lambda}] = [0, 1]$. Therefore, the derivative of minmax losses with respect to ε at $\hat{\alpha} < \hat{\alpha}_0$ when $\varepsilon \to 0$ for any $[\underline{\lambda}, \overline{\lambda}]$ is lower than 1 since

$$\lim_{\varepsilon\to 0}\frac{d\overline{L}}{d\varepsilon} \leq \frac{1}{e-1}\left(e^{1/e}-1\right) \approx 0.2588 < 1/2.$$

Note that when $\varepsilon = 0$ loss is given by equation (33). Thus, we conclude that

$$\overline{L} = \overline{L}_0 + k_1\varepsilon + o(\varepsilon).$$

with $k_1 \leq \frac{1}{e-1}\left(e^{1/e}-1\right) < 1/2$ given by equation (37) and $\overline{L}_0 = \phi(\underline{\lambda})\ln\left(\frac{\phi(\overline{\lambda})}{\phi(\underline{\lambda})}\right)$.

Case 2: $\hat{\alpha} > \hat{\alpha}_0$. Note that in this case we have $\overline{\phi}(\underline{\lambda}) > e^{-C(\lambda_0)}$. We proceed with deriving $\frac{d\lambda_0}{d\varepsilon}$. Using Lemmata C.5 and C.6 we obtain

$$\begin{aligned}
\frac{\partial H}{\partial\varepsilon} &= \frac{1}{2}\left(\frac{\partial H}{\partial\overline{\alpha}} - \frac{\partial H}{\partial\underline{\alpha}}\right) \\
&= (1-\underline{\lambda})(C(\lambda_0) + \ln\overline{\phi}(\underline{\lambda})) - \overline{\phi}(\underline{\lambda})\frac{1-\lambda_0}{\overline{\phi}(\lambda_0)} + \frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)}(1-\overline{\phi}(\underline{\lambda}))
\end{aligned}$$

and

$$\lim_{\varepsilon\to 0}\frac{\partial H}{\partial\varepsilon} = \frac{1}{2}\left[(1-\underline{\lambda})\left(1-\ln\frac{\hat{\phi}(\overline{\lambda})}{\hat{\phi}(\underline{\lambda})}\right) - \hat{\phi}(\underline{\lambda})\frac{1-\lambda_0}{\hat{\phi}(\lambda_0)} + \frac{\overline{\lambda}-\lambda_0}{\hat{\phi}(\overline{\lambda})\hat{\phi}(\lambda_0)}(1-\hat{\phi}(\underline{\lambda}))\right].$$

From Lemma C.2 we have

$$\frac{\partial H}{\partial \lambda_0} = \overline{\phi}(\underline{\lambda})\frac{\overline{\alpha} - \underline{\alpha}}{\overline{\phi}(\lambda_0)\underline{\phi}(\lambda_0)} - \frac{1 - \underline{\alpha}}{\underline{\phi}(\lambda_0)}$$

and

$$\lim_{\varepsilon \to 0} \frac{\partial H}{\partial \lambda_0} = -\frac{1 - \hat{\alpha}}{\hat{\phi}(\lambda_0)}.$$

Thus, we get that

$$\lim_{\varepsilon \to 0} \frac{d\lambda_0}{d\varepsilon} = \frac{\hat{\phi}(\lambda_0)}{2(1 - \hat{\alpha})}\left[(1 - \underline{\lambda})\left(1 - \ln \frac{\hat{\phi}(\overline{\lambda})}{\hat{\phi}(\underline{\lambda})}\right) - \hat{\phi}(\underline{\lambda})\frac{1 - \lambda_0}{\hat{\phi}(\lambda_0)} + \frac{\overline{\lambda} - \lambda_0}{\hat{\phi}(\overline{\lambda})\hat{\phi}(\lambda_0)}(1 - \hat{\phi}(\underline{\lambda}))\right].$$

Recall that $\overline{L} = \frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\lambda_0)}$, see equation (9). Thus,

$$\begin{aligned}
\frac{d\overline{L}}{d\varepsilon} &= \frac{1}{2}\frac{\partial \overline{L}}{\partial \overline{\alpha}} - \frac{1}{2}\frac{\partial \overline{L}}{\partial \underline{\alpha}} + \frac{\partial \overline{L}}{\partial \lambda_0}\frac{d\lambda_0}{d\varepsilon} \\
&= \frac{1}{2}\frac{\overline{\lambda} - \lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)} - \frac{1 - \underline{\alpha}}{\underline{\phi}(\lambda_0)}\frac{d\lambda_0}{d\varepsilon}.
\end{aligned}$$

Thus, we obtain

$$\begin{aligned}
\lim_{\varepsilon \to 0} \frac{d\overline{L}}{d\varepsilon} &= \frac{1}{2}\left[\hat{\phi}(\underline{\lambda})\frac{1 - \lambda_0}{\hat{\phi}(\lambda_0)} + \frac{\overline{\lambda} - \lambda_0}{\hat{\phi}(\overline{\lambda})\hat{\phi}(\lambda_0)}\hat{\phi}(\underline{\lambda}) - (1 - \underline{\lambda})\left(1 - \ln \frac{\hat{\phi}(\overline{\lambda})}{\hat{\phi}(\underline{\lambda})}\right)\right] = \\
&= \frac{1}{2(1 - \hat{\alpha})}\left[\hat{\phi}(\underline{\lambda})\frac{1 - \hat{\phi}(\lambda_0)}{\hat{\phi}(\lambda_0)} + \frac{\hat{\phi}(\overline{\lambda}) - \hat{\phi}(\lambda_0)}{\hat{\phi}(\overline{\lambda})\hat{\phi}(\lambda_0)}\hat{\phi}(\underline{\lambda}) - [1 - \hat{\phi}(\underline{\lambda})]\left(1 - \ln \frac{\hat{\phi}(\overline{\lambda})}{\hat{\phi}(\underline{\lambda})}\right)\right] \\
&= \frac{1}{2(1 - \hat{\alpha})}\left[2\frac{\hat{\phi}(\underline{\lambda})}{\hat{\phi}(\lambda_0)} - \frac{\hat{\phi}(\underline{\lambda})[1 + \hat{\phi}(\overline{\lambda})]}{\hat{\phi}(\overline{\lambda})} - [1 - \hat{\phi}(\underline{\lambda})]\left(1 - \ln \frac{\hat{\phi}(\overline{\lambda})}{\hat{\phi}(\underline{\lambda})}\right)\right] \\
&= \frac{1}{2(1 - \hat{\alpha})}\left[2\frac{\hat{\phi}(\underline{\lambda})}{\hat{\phi}(\lambda_0)} - \frac{\hat{\phi}(\underline{\lambda}) + \hat{\phi}(\overline{\lambda})}{\hat{\phi}(\overline{\lambda})} + [1 - \hat{\phi}(\underline{\lambda})]\ln \frac{\hat{\phi}(\overline{\lambda})}{\hat{\phi}(\underline{\lambda})}\right].
\end{aligned}$$

From solving $\lim_{\varepsilon\to 0} H(\lambda_0) = 0$ we get

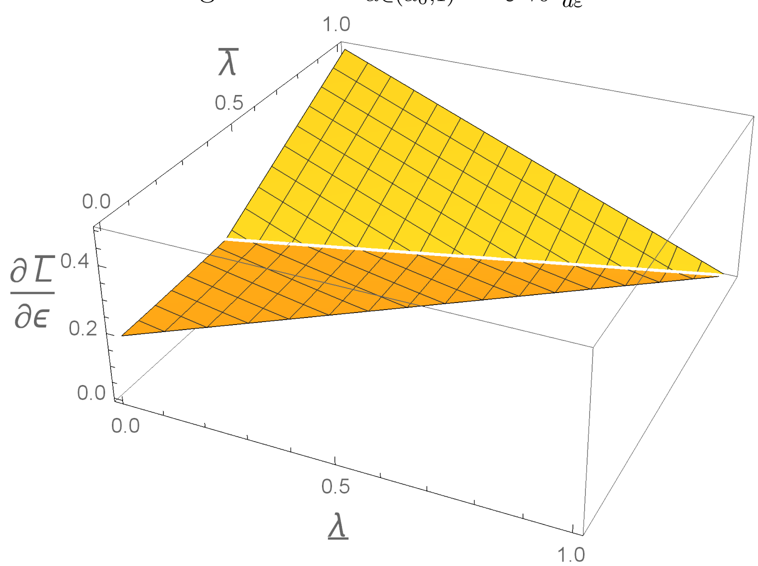
$$\hat{\phi}(\lambda_0) = \hat{\phi}(\overline{\lambda}) e^{-\hat{\phi}(\underline{\lambda}) \ln \frac{\hat{\phi}(\overline{\lambda})}{\hat{\phi}(\underline{\lambda})}},$$

which gives the final result

$$\lim_{\varepsilon\to 0} \frac{d\overline{L}}{d\varepsilon} = \frac{1}{2(1-\hat{\alpha})} \left[2\frac{\hat{\phi}(\underline{\lambda})}{\hat{\phi}(\overline{\lambda})} e^{\hat{\phi}(\underline{\lambda}) \ln \frac{\hat{\phi}(\overline{\lambda})}{\hat{\phi}(\underline{\lambda})}} - \frac{\hat{\phi}(\underline{\lambda}) + \hat{\phi}(\overline{\lambda})}{\hat{\phi}(\overline{\lambda})} + [1 - \hat{\phi}(\underline{\lambda})] \ln \frac{\hat{\phi}(\overline{\lambda})}{\hat{\phi}(\underline{\lambda})} \right]. \tag{39}$$

We analyse (39) numerically. It is a function of $\hat{\alpha}$ and parameters $\underline{\lambda}, \overline{\lambda}$. The function is concave in $\hat{\alpha}$, thus we plug in either $\hat{\alpha} = \hat{\alpha}_0$ or $\hat{\alpha} = 1$ and check how the maximum of these two candidate maxima depends on parameters. The results are presented in Figure 5

Figure 5: $\max_{\hat{\alpha}\in(\hat{\alpha}_0,1)} \lim_{\varepsilon\to 0} \frac{d\overline{L}}{d\varepsilon}$



Note that the maximum is attained at the point $\underline{\lambda} = 0$ and$\underline{\lambda} = 1$. Plugging it to (39) and letting $\hat{\alpha} \to 1$ we obtain

$$\lim \hat{\alpha} \to 1 \left(\lim_{\varepsilon\to 0} \frac{d\overline{L}}{d\varepsilon} = \right) = \frac{1}{2}$$

Thus, we conclude that

$$\overline{L} = \overline{L}_0 + k_2\varepsilon + o(\varepsilon).$$

with $k_2 \leq \frac{1}{2}$ given by equation (39) and $\overline{L}_0 = \phi(\underline{\lambda}) \ln\left(\frac{\phi(\overline{\lambda})}{\phi(\underline{\lambda})}\right)$.

Case 3: $\hat{\alpha} = \hat{\alpha}_0$. As loss function is continuously differentiable in α we ogtain that loss is presented as $\overline{L} = k_3\varepsilon + o(\varepsilon)$ with $k_3 \leq \max\{k_1, k_2\} < 1/2$. □

Proof of Lemma 4. By plugging in $\lambda = \overline{\lambda}$ to equation (22) we have that the equilibrium minmax loss can be written as

$$\overline{L} = (1 - \underline{\alpha}) \int_{\underline{\mu}(\lambda_0)}^{\overline{\lambda}} G(\mu) d\mu, \tag{40}$$

where $\underline{\mu}(\lambda_0)$ is given by (18). Thus, by integrating $\mathbb{E}_G\mu$ by parts and using (40) we find that

$$\begin{aligned}
\mathbb{E}_G\mu &= \int_{\underline{\mu}(\lambda_0)}^{\overline{\lambda}} \mu dG(\mu) + G(\underline{\mu}(\lambda_0))\underline{\mu}(\lambda_0) \\
&= \mu G(\mu)\Big|_{\underline{\mu}(\lambda_0)}^{\overline{\lambda}} - \int_{\underline{\mu}(\lambda_0)}^{\overline{\lambda}} G(\mu) d\mu + G(\underline{\mu}(\lambda_0))\underline{\mu}(\lambda_0) \\
&= \overline{\lambda} - \int_{\underline{\mu}(\lambda_0)}^{\overline{\lambda}} G(\mu) d\mu \\
&= \overline{\lambda} - \frac{\overline{L}}{1 - \underline{\alpha}}.
\end{aligned}$$

□

Proof of Proposition 4. To explore the sign of derivative of $\mathbb{E}_G\mu$ with respect to any parameter of the model we use the relation of $\mathbb{E}_G\mu$ and the minimax loss $\overline{L}$ derived in Lemma 4 and the properties of the loss derivatives from Proposition 2 (Part 1).

Equation (40) implies that the derivative of $\mathbb{E}_G\mu$ with respect to $\overline{\lambda}$ is given by

$$\frac{d\mathbb{E}_G\mu}{d\overline{\lambda}} = 1 - \frac{1}{1-\underline{\alpha}}\frac{d\overline{L}}{d\overline{\lambda}}.$$

We build up on the proof of Proposition 2, which uses Lemmata C.2 and C.3, to evaluate $\frac{d\overline{L}}{d\overline{\lambda}}$.

For $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$ we obtain

$$\begin{aligned}
\frac{d\mathbb{E}_G\mu}{d\overline{\lambda}} &= 1 - \frac{1}{1-\underline{\alpha}}\frac{d\overline{L}}{d\overline{\lambda}} \\
&= 1 - \frac{1}{\underline{\phi}(\overline{\lambda})} + \frac{1}{\overline{\phi}(\lambda_0)}\frac{d\lambda_0}{d\overline{\lambda}} \\
&= 1 - \frac{1}{\underline{\phi}(\overline{\lambda})}\left(1 - \frac{(1-\underline{\alpha})[1-\overline{\phi}(\underline{\lambda})]}{1-\underline{\alpha}-(\overline{\alpha}-\underline{\alpha})\frac{\overline{\phi}(\underline{\lambda})}{\overline{\phi}(\lambda_0)}}\right) \\
&= \frac{(1-\underline{\alpha})[1-\overline{\phi}(\underline{\lambda})] - [1-\underline{\phi}(\overline{\lambda})]\left(1-\underline{\alpha}-(\overline{\alpha}-\underline{\alpha})\frac{\overline{\phi}(\underline{\lambda})}{\overline{\phi}(\lambda_0)}\right)}{\left(1-\underline{\alpha}-(\overline{\alpha}-\underline{\alpha})\frac{\overline{\phi}(\underline{\lambda})}{\overline{\phi}(\lambda_0)}\right)\underline{\phi}(\overline{\lambda})}.
\end{aligned}$$

As denominator of this expression is always positive, we focus of the numerator.

$$\begin{aligned}
\frac{d\mathbb{E}_G\mu}{d\overline{\lambda}} &\propto (1-\underline{\alpha})[1-\overline{\phi}(\underline{\lambda})] - [1-\underline{\phi}(\overline{\lambda})]\left(1-\underline{\alpha}-(\overline{\alpha}-\underline{\alpha})\frac{\overline{\phi}(\underline{\lambda})}{\overline{\phi}(\lambda_0)}\right) \\
&= (1-\underline{\alpha})[1-\overline{\phi}(\underline{\lambda})] - (1-\underline{\alpha})[1-\underline{\phi}(\overline{\lambda})] + [1-\underline{\phi}(\overline{\lambda})](\overline{\alpha}-\underline{\alpha})\frac{\overline{\phi}(\underline{\lambda})}{\overline{\phi}(\lambda_0)} \\
&= (1-\underline{\alpha})[\underline{\phi}(\overline{\lambda}) - \overline{\phi}(\underline{\lambda})] + (1-\underline{\alpha})(1-\overline{\lambda})(\overline{\alpha}-\underline{\alpha})\frac{\overline{\phi}(\underline{\lambda})}{\overline{\phi}(\lambda_0)} \\
&= (1-\underline{\alpha})[\underline{\phi}(\overline{\lambda}) - \overline{\phi}(\underline{\lambda})] + (1-\underline{\alpha})(\overline{\phi}(\underline{\lambda}) - \underline{\phi}(\underline{\lambda}))\frac{\overline{\phi}(\underline{\lambda})}{\overline{\phi}(\lambda_0)}.
\end{aligned}$$

This implies that

$$\begin{aligned}
\frac{d\mathbb{E}_G\mu}{d\overline{\lambda}} &\propto [\underline{\phi}(\overline{\lambda}) - \overline{\phi}(\underline{\lambda})]\overline{\phi}(\lambda_0) + [\overline{\phi}(\overline{\lambda}) - \underline{\phi}(\underline{\lambda})]\overline{\phi}(\underline{\lambda}) \\
&= \underline{\phi}(\overline{\lambda})[\overline{\phi}(\lambda_0) - \overline{\phi}(\underline{\lambda})] + \overline{\phi}(\underline{\lambda})[\overline{\phi}(\overline{\lambda}) - \overline{\phi}(\lambda_0)] > 0,
\end{aligned}$$

since $\underline{\lambda} < \lambda_0 < \overline{\lambda}$ and $\overline{\phi}(\cdot)$ is a strictly increasing function.

Case $\overline{\phi}(\underline{\lambda}) < e^{-C(\lambda_0)}$ follows similar steps, so we present it in less detail.

$$\begin{aligned}
\frac{d\mathbb{E}_G\mu}{d\overline{\lambda}} &= 1 - \frac{1}{\underline{\phi}(\overline{\lambda})}\left(1 - \frac{(1-\underline{\alpha})\left[1-e^{-C(\lambda_0)}\right]}{1-\underline{\alpha}-(\overline{\alpha}-\underline{\alpha})\frac{e^{-C(\lambda_0)}}{\overline{\phi}(\lambda_0)}}\right)\\
&\propto (1-\underline{\alpha})\left[1-e^{-C(\lambda_0)}\right] - (1-\underline{\alpha})[1-\underline{\phi}(\overline{\lambda})] + [1-\underline{\phi}(\overline{\lambda})](\overline{\alpha}-\underline{\alpha})\frac{e^{-C(\lambda_0)}}{\overline{\phi}(\lambda_0)}\\
&\propto [\underline{\phi}(\overline{\lambda}) - e^{-C(\lambda_0)}]\overline{\phi}(\lambda_0) + [\overline{\phi}(\overline{\lambda}) - \underline{\phi}(\underline{\lambda})]e^{-C(\lambda_0)}\\
&= \underline{\phi}(\overline{\lambda})[\overline{\phi}(\lambda_0) - e^{-C(\lambda_0)}] + e^{-C(\lambda_0)}[\overline{\phi}(\overline{\lambda}) - \overline{\phi}(\lambda_0)] > 0
\end{aligned}$$

where the last inequality follows from $\overline{\phi}(\lambda_0) > e^{-C(\lambda_0)}$, which was derived in (42). Thus, we conclude that

$$\frac{d\mathbb{E}_G\mu}{d\overline{\lambda}} > 0.$$

From Proposition 2 we have that $\overline{L}$ weakly decreases in $\underline{\lambda}$. Thus, the derivative of $\mathbb{E}_G\mu$ with respect to $\underline{\lambda}$ is

$$\frac{d\mathbb{E}_G\mu}{d\underline{\lambda}} = -\frac{1}{1-\underline{\alpha}}\frac{d\overline{L}}{d\underline{\lambda}} \geq 0.$$

Also, from Proposition 2 we have that $\overline{L}$ increases in $\overline{\alpha}$. Therefore, derivative of $\mathbb{E}_G\mu$ with respect to $\overline{\alpha}$ is

$$\frac{d\mathbb{E}_G\mu}{d\overline{\alpha}} = -\frac{1}{1-\underline{\alpha}}\frac{d\overline{L}}{d\overline{\alpha}} < 0.$$

The derivative of $\mathbb{E}_G\mu$ with respect to $\underline{\alpha}$ is given by

$$\frac{d\mathbb{E}_G\mu}{d\underline{\alpha}} = -\frac{1}{(1-\underline{\alpha})^2}\left(\overline{L} + (1-\underline{\alpha})\frac{d\overline{L}}{d\underline{\alpha}}\right).$$

The inequality (31) implies that

$$\overline{L} + (1-\underline{\alpha})\frac{d\overline{L}}{d\underline{\alpha}} < \overline{L} - \frac{(1-\underline{\alpha})(\overline{\lambda}-\lambda_0)}{\underline{\phi}(\lambda_0)} = \ln\frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\lambda_0)} + 1 - \frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\lambda_0)} \leq 0,$$

since $\ln(x) \leq x - 1$ for all $x > 0$. Therefore, $d\mathbb{E}_G\mu/d\underline{\alpha} > 0$. □

A4: The Informed Receiver

Proof of Proposition 5. We start with proving a series of claims about the supports and continuity of equilibrium distribution functions. Let

$$L_F(\alpha) = \int_{\underline{\mu}}^{\overline{\mu}} L_r(\mu; \alpha) dG(\mu)$$

and

$$L_G(\mu) = \int_{\underline{\alpha}_F}^{\overline{\alpha}_F} L_r(\mu; \alpha) dF(\alpha)$$

Claim 1: $\underline{\alpha}_F = \lambda_r^{-1}(\underline{\mu})$ and $\overline{\alpha}_F = \lambda_r^{-1}(\overline{\mu}) = r$. Suppose that for some $b > a$ we have $G(b) - G(a) = 0$ but $F(\lambda_r^{-1}(b)) - F(\lambda_r^{-1}(a)) > 0$. Then L_F is increasing on this interval which contradicts the indifference. Similarly if for some $b > a$ we have $F(b) - F(a) = 0$ but $G(\lambda(b)) - F(\lambda(a)) > 0$ we have that L_G is increasing on this interval which contradicts the indifference. If both distribution functions have a mutual gap over some (a, b) and $(\lambda_r^{-1}(a), \lambda_r^{-1}(b))$ we have that $L_F((a + b)/2) > L_F(a)$, so nature has a profitable deviation.

Claim 2: F is continuous of $[\underline{\alpha}_F, \overline{\alpha}_F)$ and $\overline{\alpha}_F = r$. Suppose that F has an atom at $\alpha_1 < r$ and let $\lambda_1 \equiv \lambda_r(\alpha_1)$. Note that in this case there exists ε_1 such that for all $\varepsilon < \varepsilon_1$ $L_G(\lambda_1) < L_G(\lambda_1 + \varepsilon)$ (as there is higher probability to end up in the first rather than the second case). Then there exists a gap: $G(\lambda_1 + \varepsilon_1) - G(\lambda_1) = 0$, and loss of the nature is increasing in that gap: $L_F(\alpha_1) < L_F(\lambda_r^{-1}(\lambda_1 + \varepsilon_1/2))$ – a contradiction with optimality of F. Finally, note nature's loss is decreasing in α at the upper bound, so the upper bound must be equal to r, which together with Claim 1 implies $\overline{\mu} = 1$.

Claim 3: G is continuous of $[\underline{\mu}, \overline{\mu}]$. Suppose that G has an atom at $\mu_1 \geq \underline{\mu} > 0$. Then there exist ε_1 such that for all $\varepsilon < \varepsilon_1$ we have $L_F(\lambda_r^{-1}(\mu_1)) < L_F(\lambda_r^{-1}(\mu_1) - \varepsilon)$ and thus a gap in support of F must exist just below $\lambda_r^{-1}(\mu_1)$. As both first and second case are decreasing in μ the sender has lower losses over that gap than at μ_1 and thus a profitable deviation exists. It remains to check the case $\underline{\mu} = 0$. As supports of F and G must coincide (Claim 1) we obtain that $\alpha = 0$ is in the support of nature. As $\lambda_r(0) = 0$ this choice of α generates zero loss, while nature could obtain positive loss by choosing higher α, which contradicts the indifference condition.

Claims 1-3 imply that the equilibrium strategy of the sender G is determined on $[\underline{\mu}, 1]$ and the equilibrium strategy of the nature F is determined on $[\underline{\alpha}_F, r]$, where the lower bounds satisfy $\lambda_r(\underline{\alpha}_F) = \underline{\mu}$. In the equilibrium the sender and the nature must be indifferent between playing any strategy in their equilibrium supports.

We first solve for the equilibrium strategy of the sender G by exploring the indifference condition of the nature. Then we characterize the equilibrium strategy of the nature by exploring the indifference condition of the sender.

Strategy of the sender. We consider non-trivial case of $r < 1$. To simplify algebraic expressions define $\tau = \frac{1-r}{r} > 0$. In what follows, it is useful to note by taking the derivative of (11) with respect to α we obtain

$$\frac{\partial \lambda_r}{\partial \alpha} = \frac{\tau}{(1-\alpha)^2} = \frac{(\tau + \lambda)^2}{\tau}.$$

Claim 3 implies that G is continuous on $[\underline{\mu}, 1]$. Therefore, the objective function of the nature that plays $\alpha \in [\alpha_F, r]$ is given by

$$L_F(\alpha) = [1 - G(\lambda_r(\alpha))]\frac{\alpha}{r} + (1-\alpha)\int_{\underline{\mu}}^{\lambda_r(\alpha)} G(\mu) d\mu.$$

Note that the first term represents the expected losses of (12) from not persuading and the second term represents the expected losses from revealing too much information.

By taking derivative with respect to α we get

$$0 = -g(\lambda)\frac{\partial \lambda_r}{\partial \alpha}\frac{\alpha}{r} + (1 - G(\lambda))\frac{1}{r} - \int_{\underline{\mu}}^{\lambda} G(\mu)d\mu + (1-\alpha)G(\lambda)\frac{\partial \lambda_r}{\partial \alpha}$$

Using (11), expression for $\frac{\partial \lambda_r}{\partial \alpha}$ and changing the variable $\lambda = \lambda_r(\alpha)$ we obtain

$$0 = -g(\lambda)\frac{(t+\lambda)\lambda}{tr} + \frac{1 - G(\lambda)}{r} - \int_{\underline{\mu}}^{\lambda} G(\mu)d\mu + G(\lambda)(t+\lambda).$$

To solve for G we take one more derivative with respect to λ:

$$\begin{aligned} 0 &= -\frac{1}{tr}(g'(\lambda)(t+\lambda)\lambda + g(\lambda)(t+2\lambda)) - \frac{g(\lambda)}{r} - G(\lambda) + G(\lambda) + g(\lambda)(t+\lambda) \\ &= \frac{t+\lambda}{tr}\left[g'(\lambda)\lambda + g(\lambda)(2-tr)\right] \end{aligned}$$

Therefore,

$$\frac{g'(\lambda)}{g(\lambda)} = -\frac{2-tr}{\lambda} = -\frac{1+r}{\lambda}.$$

By integrating on both sides we arrive to equation

$$\log g(\lambda) = -(1+r)\log\lambda + \log A_0,$$

where A_0 is some constant. This holds true if and only if the density function takes the form of $g(\mu) = \frac{A_0}{\mu^{1+r}}$. Thus, the strategy of the sender is given by

$$G(\mu) = -\frac{A_0}{r\mu^r} + A_1.$$

It remains to solve for the parameters A_0, A_1. From Claims 2 and 3 we have that $G(1) = 1$, which implies $-\frac{A_0}{r} + A_1 = 1$. Note that the lower bound $\underline{\mu}$ must satisfy $G(\underline{\mu}) = 0$, so $A_1 = \frac{A_0}{r\underline{\mu}^r}$.

We can express A_0 and A_1 as functions of $\underline{\mu}$. By plugging $A_0 = A_1 r \underline{\mu}^r$ into $G(1) = 1$ we obtain that

$$A_0 = \frac{r\underline{\mu}^r}{1-\underline{\mu}^r} \quad \text{and} \quad A_1 = \frac{1}{1-\underline{\mu}^r}.$$

In order to characterize $\underline{\mu}$ we exploit the optimality of $G(\mu)$ by plugging it back to the derivative of $L_F(\alpha)$:

$$\begin{aligned}
0 &= -\frac{A_0(t+\lambda)}{(1-r)\lambda^r} + \frac{1}{r}\left(1 - A_1 + \frac{A_0}{r\lambda^r}\right) - A_1(\lambda - \underline{\mu}) + \frac{A_0}{r(1-r)}\left(\frac{\lambda}{\lambda^r} - \frac{\underline{\mu}}{\underline{\mu}^r}\right) \\
&+ A_1(t+\lambda) - \frac{A_0(t+\lambda)}{r\lambda^r} \\
&= \frac{1}{r} - A_1 + A_1\underline{\mu} - \frac{A_0}{(1-r)r}\frac{\underline{\mu}}{\underline{\mu}^r}.
\end{aligned}$$

By using the expression for A_0 and A_1 derived above we find that

$$\frac{1}{r} = A_1\left(1 + \underline{\mu}\frac{r}{1-r}\right) = \frac{1}{1-\underline{\mu}^r}\left(1 + \underline{\mu}\frac{r}{1-r}\right),$$

which implies that

$$\underline{\mu}^r + \frac{r^2}{1-r}\underline{\mu} - (1-r) = 0.$$

Next we show that the solution of this equation exists and is unique. Define function $\kappa(\mu) = \mu^r + \frac{r^2}{1-r}\mu - (1-r)$. Note that $\kappa(\mu)$ increases in μ, $\kappa(0) = -(1-r) < 0$ and $\kappa(1) = r + r^2/(1-r) > 0$. Therefore there exists a unique $\underline{\mu}$ that solves $\kappa(\underline{\mu}) = 0$.

Strategy of nature. It remains to characterize the equilibrium strategy of the nature by exploring the problem of the sender. To simplify algebra we define $z = \frac{r}{1-r}$ and

$$\zeta(\mu) = \frac{z\mu}{1+z\mu}.$$

In what follows, it is useful to calculate the derivatives of ζ with respect to μ

$$\zeta' = \frac{z}{(1+z\mu)^2} = z(1-\zeta)^2,$$
$$\zeta'' = -2z(1-\zeta)\zeta' = -2z^2(1-\zeta)^3.$$

Using $\lambda_r(\underline{\alpha}_F) = \underline{\mu}$ we obtain

$$\underline{\alpha}_F = \frac{\underline{\mu}}{t+\underline{\mu}} = \frac{r\underline{\mu}}{1-r+r\underline{\mu}}.$$

The losses of the sender who plays μ in response to the nature's strategy $F(\alpha) : [\underline{\alpha}_F, r]$ is given by

$$\begin{aligned} L_G(\mu) &= (1-\mu)(1-F(r))(1-r) + \int_\zeta^r \left(\alpha\frac{1-r}{r} - (1-\alpha)\mu\right) dF(\alpha) + \int_{\underline{\alpha}_F}^\zeta \frac{\alpha}{r} dF(\alpha) \\ &= (1-\mu)(1-r) - \left(\frac{1-r}{r} + \mu\right) \int_\zeta^r F(\alpha)d\alpha + \frac{1}{r}\int_{\underline{\alpha}_F}^\zeta \alpha dF(\alpha). \end{aligned}$$

Since $L_G(\mu)$ is constant for all μ in the equilibrium support we have that by taking the derivative of the expression above we obtain

$$0 = -(1-F(r))(1-r) - \int_\zeta^r (1-\alpha)dF(\alpha) + \frac{1}{r}\zeta f(\zeta)\zeta'.$$

To solve for f we differentiate this expression one more time with respect to μ using the expression for ζ' and ζ'' that we derived earlier

$$\begin{aligned} 0 &= r(1-\zeta)f(\zeta)\zeta' + f'(\zeta)(\zeta')^2\zeta + f(\zeta)\frac{d\zeta\zeta'}{d\mu} \\ &= r(1-\zeta)f(\zeta)\zeta' + f'(\zeta)(\zeta')^2\zeta + f(\zeta)((\zeta')^2 + \zeta\zeta'') \\ &= f(\zeta)(r(1-\zeta)\zeta' + (\zeta')^2 + \zeta\zeta'') + f'(\zeta)(\zeta')^2\zeta \\ &= f(\zeta)(rz(1-\zeta)^3 + z^2(1-\zeta)^4 - 2z^2(1-\zeta)^3\zeta) + f'(\zeta)z^2(1-\zeta)^4\zeta. \end{aligned}$$

By dividing the resulting expression by $z^2(1-\zeta)^3$ we obtain

$$
\begin{aligned}
0 &= f(\zeta)(r/z + (1-\zeta) - 2\zeta) + f'(\zeta)(1-\zeta)\zeta \\
&= f(\zeta)(2 - r - 3\zeta) + f'(\zeta)(1-\zeta)\zeta.
\end{aligned}
$$

Therefore

$$\frac{f'(\zeta)}{f(\zeta)} = -\frac{2-r-3\zeta}{(1-\zeta)\zeta}.$$

We integrate this expression from both sides

$$
\begin{aligned}
\log f(\zeta) &= -(2-r)\log(\zeta) - (1+r)\log(1-\zeta) + \log C_0 \\
&= -\log(\zeta^{2-r}(1-\zeta)^{1+r}) + \log C_0
\end{aligned}
$$

and solve for the density function

$$f(\alpha) = \frac{C_0}{\alpha^{2-r}(1-\alpha)^{1+r}}$$

which implies that the strategy of the nature has the following functional form

$$F(\alpha) = -\frac{C_0}{(1-r)r}\frac{r-\alpha}{\alpha}\left(\frac{\alpha}{1-\alpha}\right)^r + C_1.$$

It remains to characterize coefficients C_0 and C_1. We exploit the optimality of $F(\alpha)$ and plug it back into the indifference condition

$$(1-r)(1-F(r)) = \int_\zeta^r (1-\alpha)dF(\alpha) - \frac{z}{r}\zeta f(\zeta)(1-\zeta)^2 \frac{C_0}{1-r}\left(\frac{r}{1-r}\right)^{r-1}.$$

Since $F(r) = 1$, we have that

$$1 - C_1 = \frac{C_0}{(1-r)r}\left(\frac{r}{1-r}\right)^r.$$

The definition of the lower bound $F(\underline{\alpha}_F) = 0$ implies that

$$C_1 = \frac{C_0}{(1-r)r}\frac{r-\underline{\alpha}_F}{\underline{\alpha}_F}\left(\frac{\underline{\alpha}_F}{1-\underline{\alpha}_F}\right)^r.$$

By adding last two equations and using the fact that $\frac{\underline{\alpha}_F}{1-\underline{\alpha}_F} = \frac{r}{1-r}\underline{\mu}$ we obtain

$$\begin{aligned}1 &= \frac{C_0}{(1-r)r}\left(\left(\frac{r}{1-r}\right)^r + \frac{r-\underline{\alpha}_F}{\underline{\alpha}_F}\left(\frac{\underline{\alpha}_F}{1-\underline{\alpha}_F}\right)^r\right)\\ &= \frac{C_0}{(1-r)r}\left(\frac{r}{1-r}\right)^r\left(1 + \frac{(1-r)(1-\underline{\mu})}{\underline{\mu}}\underline{\mu}^r\right),\end{aligned}$$

which, in turn, implies that

$$C_0 = \frac{r}{\left(\frac{r}{1-r}\right)^r}\frac{1-r-\underline{\mu}^r}{(1-\underline{\mu}^r)^2} \quad \text{and} \quad C_1 = \frac{C_0}{r}\left(\frac{r}{1-r}\right)^r\underline{\mu}^{r-1}(1-\underline{\mu}).$$

$\square$

Proof of Proposition 6. The objective function of the nature that plays $\alpha \in [\alpha_F, r]$ is given by

$$L_F(\alpha) = [1 - G(\lambda_r(\alpha))]\frac{\alpha}{r} + (1-\alpha)\int_{\underline{\mu}}^{\lambda_r(\alpha)} G(\mu)d\mu.$$

By plugging r into $L_F(\alpha)$ we calculate the minmax loss

$$\begin{aligned}
\overline{L}_r &= (1-r)\int_{\underline{\mu}}^{1} G(\mu)du \\
&= (1-r)\left(A_1 - \frac{A_0}{(1-r)r} - \left(A_1\underline{\mu} - \frac{A_0}{(1-r)r}\frac{\underline{\mu}}{\underline{\mu}^r}\right)\right) \\
&= (1-r)\left(A_1 - \frac{A_0}{(1-r)r} - \left(A_1 - \frac{1}{r}\right)\right) \\
&= \frac{1}{r}\left(1-r-A_0\right).
\end{aligned}$$

From the definition of A_0 and equation (13)

$$\begin{aligned}
\overline{L}_r &= \frac{1}{r}\left(1-r-\frac{r\underline{\mu}^r}{1-\underline{\mu}^r}\right) = \frac{1}{r} - \frac{1}{1-\underline{\mu}^r} \\
&= \frac{1}{r} - \frac{1-r}{r}\frac{1}{1-r+r\underline{\mu}} \\
&= \frac{\underline{\mu}}{1-r+r\underline{\mu}} = \frac{\underline{\alpha}_F}{r}.
\end{aligned}$$

Next we show that $\overline{L}_r$ decreases in r. From the implicit function theorem applied to equation (13) we have that

$$\frac{d\underline{\mu}}{dr} = -\frac{1+\frac{r(2-r)}{(1-r)^2}}{rx^{r-1}+\frac{r^2}{1-r}} < 0,$$

so the lower bound $\underline{\mu}$ decreases in r. Consequently, since $\frac{r}{1-r}\underline{\mu}$ decreases in r we have that

$$\underline{\alpha}_F = \frac{\underline{\mu}}{\frac{1-r}{r}+\underline{\mu}} = 1 - \frac{1}{1+\frac{r}{1-r}\underline{\mu}_G}$$

decreases in r as well. Therefore, $\overline{L}_r = \frac{\underline{\alpha}_F}{r}$ is higher for lower r.

The upper bound of $\overline{L}_r$ can be found by taking the limit of $\overline{L}_r$ when r tends to 0. we explore the limit of $\underline{\mu}$ when r tends to 0. By rearranging (13) and taking the logs we obtain

$$\lim_{r\to+0}\log\underline{\mu} = \lim_{r\to+0}\frac{\log\left(1-r-\frac{r^2}{1-r}\underline{\mu}\right)}{r}.$$

Using the fact that $\underline{\mu}$ is bounded for all $r>0$, we have that $\lim_{r\to+0}\log\underline{\mu} = \lim_{r\to+0}\log(1-r)/r = -1$. Thus, $\underline{\mu}$ converges to $1/e$ when r tends to 0. Finally we obtain that

$$\lim_{r\to+0}\overline{L}_r = \lim_{r\to+0}\frac{\underline{\mu}}{1-r+r\underline{\mu}} = \frac{1}{e}.$$

$\square$

Appendix B: Proof for Model with Large Number of States

Proof of Theorem 2. Suppose that malicious nature uniformly randomizes between $\binom{n}{\lfloor\sqrt{n}\rfloor}$ possible types of the receiver. Receiver j's prior belief about the state of the world is characterized by

$$\beta^j_w = \begin{cases} 1 - \lfloor\sqrt{n}\rfloor\delta - (n - \lfloor\sqrt{n}\rfloor)\varepsilon, & \text{if } \omega = 1, \\ \delta, & \text{if } \omega \in S^j, \\ \varepsilon, & \text{otherwise,} \end{cases}$$

where $S^j \subset \Omega$ consisting of $\lfloor\sqrt{n}\rfloor$ elements from Ω. We define[14]

$$\delta = \frac{1}{\lfloor\sqrt{n}\rfloor n^{1/2}} \quad \text{and} \quad \varepsilon = \frac{1}{(n - \lfloor\sqrt{n}\rfloor)n^{7/2}}.$$

Suppose that there are two messages m_1 and m_2 which persuade the receiver (i.e. give expected utility higher than r), sent with probabilities $\mu_\omega(m_i)$ with $\mu_\omega(m_1) + \mu_\omega(m_2) \leq 1$. Then, a message m^+ sent with probability $\mu_\omega(m^+) = \mu_\omega(m_1) + \mu_\omega(m_2)$ also persuades the receiver. Thus, we can focus on two-message mechanisms.

To determine the nature's choice of the outside option r, consider the expected payoff of receiver j with $S^j = \left\{\frac{n-\lfloor\sqrt{n}\rfloor}{n}, \ldots, \frac{n-1}{n}\right\}$, after receiving a message m^+ from the mechanism which sends this message if and only if $\omega \in \Omega \setminus S^j$. Then,

$$\mathbb{E}(u_R | m^+, a = 1) = \frac{1 - \lfloor\sqrt{n}\rfloor\delta - (n - \lfloor\sqrt{n}\rfloor)\varepsilon + \varepsilon \sum_{\omega \in \Omega \setminus (S^j \cup 1)} \omega}{1 - \lfloor\sqrt{n}\rfloor\delta}.$$

[14]Note that the receiver's prior is well-defined for all $n \geq 2$ since $\beta^j_1 = 1 - \frac{1}{n^{1/2}} - \frac{1}{n^{7/2}}$ increases in n and is equal to $1 - 9/(8\sqrt{(2)}) > 0$ for $n = 2$, which implies that $\beta^j_1 \in (0, 1)$.

Define the receiver's outside option (for all types j) as

$$r \equiv \frac{1 - \lfloor\sqrt{n}\rfloor\delta - (n - \lfloor\sqrt{n}\rfloor)\varepsilon}{1 - \lfloor\sqrt{n}\rfloor\delta}.$$

Clearly, $\mathbb{E}(u_R|m^+, a = 1) > r$, as r is equivalent to the expected payoff in the case if all the bottom $n - \lfloor\sqrt{n}\rfloor$ were replaced with zero.

Now, consider an expected payoff of a receiver who faces a mechanism which sends message m^+ with probability $p > 0$ in a 'mine state' upon receiving such message:

$$\mathbb{E}(u_R|\mu(m^+|\omega \in S^j) > 0, m^+, a = 1) < \frac{1 - \lfloor\sqrt{n}\rfloor\delta - (n - \lfloor\sqrt{n}\rfloor)\varepsilon + \frac{n-1}{n}\delta p}{1 - \lfloor\sqrt{n}\rfloor\delta - (n - \lfloor\sqrt{n}\rfloor)\varepsilon + \delta p} \equiv \rho.$$

The right-hand-side, which we denote as ρ, is the payoff of the receiver with the most optimistic prior, i.e. with $\beta_{(n-1)/n} = \delta$, upon receiving message m^+ from a mechanism which sends this message only in $\omega = 1$ with probability 1 and in $\omega = \frac{n-1}{n}$ with probability p. Alternatively, one can interpret the right-hand-side as the upper bound of the payoff of the receiver who faces a mechanism sending m^+ at least in one of mine states with probability p. Now we will show that for our choice of ε and δ we have $\rho < r$ for $p = 1$, i.e. sending m^+ just in one mine state with probability 1 is sufficient for rejection. By plugging in our expressions we obtain that

$$\rho = \frac{1 - \frac{1}{n^{1/2}} - \frac{1}{n^{7/2}} + \frac{n-1}{n^{3/2}\lfloor\sqrt{n}\rfloor}}{1 - \frac{1}{n^{1/2}} - \frac{1}{n^{7/2}} + \frac{1}{\lfloor\sqrt{n}\rfloor n^{1/2}}} < \frac{1 - \frac{1}{n^{1/2}} - \frac{1}{n^{7/2}} + \frac{n-1}{n^2}}{1 - \frac{1}{n^{1/2}} - \frac{1}{n^{7/2}} + \frac{1}{n}}.$$

Thus, $\rho < r$ whenever

$$\frac{1 - \frac{1}{n^{1/2}} - \frac{1}{n^{7/2}} + \frac{n-1}{n^2}}{1 - \frac{1}{n^{1/2}} - \frac{1}{n^{7/2}} + \frac{1}{n}} < \frac{1 - \frac{1}{n^{1/2}} - \frac{1}{n^{7/2}}}{1 - \frac{1}{n^{1/2}}}$$

which is equivalent to

$$\frac{-n^{9/2} + n^3 + n^2 + n^{3/2} + n + n^{1/2} + 1}{n^3\left(n^{7/2} - n^3 + n^{5/2} - 1\right)} < 0.$$

The denominator is clearly positive as

$$\left(n^{7/2}-n^3+n^{5/2}-1\right)=\left(n^{1/2}-1\right)n^3+n^{5/2}-1$$

and $n \geq 2$. Now we are going to work with the numerator:

$$\xi(n)=-n^{9/2}+n^3+n^2+n^{3/2}+n+n^{1/2}+1.$$

We have that $\xi(2)=15-13\sqrt{2}<0$. Moreover, for any $n \geq 2$ we have that

$$\begin{aligned}\xi'(n)&=-\frac{9}{2}n^{7/2}+3n^2+2n+\frac{3}{2}n^{1/2}+1+\frac{1}{2}n^{-1/2}\\&<-\frac{9}{2}n^{7/2}+7n^2+1\\&=\left(-\frac{9}{2}n^{3/2}+7\right)n^2+1\\&<-2n^2+1<0.\end{aligned}$$

Thus, $\xi(n)$ is decreasing, and, therefore, is always negative, which implies that for our choice of ε and δ we have $\rho<r$, meaning that message m^+ does not persuade the receiver if sent with probability 1 in a mine state. Thus, that there exists $\bar{p}\in(0,1)$ such that if the sender sends message m^+ in a 'mine' state with probability $p>\bar{p}$, then the receiver rejects. Since the receiver is indifferent between adopting and rejecting for $p=\bar{p}$ we have that

$$\bar{p}=\frac{\varepsilon n(n-\lfloor\sqrt{n}\rfloor)(1-\lfloor\sqrt{n}\rfloor\delta-(n-\lfloor\sqrt{n}\rfloor)\varepsilon)}{\delta(1-\lfloor\sqrt{n}\rfloor\delta-n(n-\lfloor\sqrt{n}\rfloor)\varepsilon)}.$$

Note that $\lim_{n\to\infty} n\bar{p}=0$.

Next we show that the maximal payoff that the uninformed sender can obtain approaches 0 when n tends to ∞. Consider two cases.

First, suppose that the sender sends message m^+ in state $\omega=1$ with probability 1 and in some other states with probabilities not higher than $\bar{p}$. Then, the probability to persuade the receiver is not higher than $\alpha_1 + \sum_{\omega\in(\Omega\setminus 1)} \alpha_\omega \bar{p} = \alpha_1 + (1-\alpha_1)\bar{p}$. Since $\alpha_1 \to 0$ and $\bar{p} \to 0$ when n tends to ∞, then the probability to persuade the receiver approaches 0 when $n \to \infty$.

Second, suppose that the sender sends m^+ in $k < n$ states with probability strictly larger than $\bar{p}$. Then, the payoff of the uninformed sender is not larger than

$$\begin{aligned}
&\alpha_1 + (n-k)p + \frac{\overline{A}}{n} k \binom{n-k}{\lfloor\sqrt{n}\rfloor} \Big/ \binom{n}{\lfloor\sqrt{n}\rfloor} \\
&= \alpha_1 + (n-k)p + \frac{\overline{A}}{n} k \left(1 - \frac{\lfloor\sqrt{n}\rfloor}{n}\right)\left(1 - \frac{\lfloor\sqrt{n}\rfloor}{n-1}\right)\left(1 - \frac{\lfloor\sqrt{n}\rfloor}{n-k+1}\right) \\
&< \alpha_1 + (n-k)p + \frac{\overline{A}}{n} k \left(1 - \frac{\lfloor\sqrt{n}\rfloor}{n}\right)^k \equiv \Psi.
\end{aligned}$$

If $k \leq \lfloor\sqrt{n}\rfloor$, then $\Psi \leq \alpha_1 + np + \overline{A}\frac{\lfloor\sqrt{n}\rfloor}{n} \to 0$ when n tends to 0. If $\lfloor\sqrt{n}\rfloor < k < n - \lfloor\sqrt{n}\rfloor$, then since $k\left(1 - \frac{\lfloor\sqrt{n}\rfloor}{n}\right)^k$ is a weakly decreasing function for all $k \geq \lfloor\sqrt{n}\rfloor + 1$, we have that

$$\begin{aligned}
\lim_{n\to\infty} \Psi &\leq \lim_{n\to\infty} \left(\alpha_1 + (n-k)p + \overline{A}\frac{\lfloor\sqrt{n}\rfloor + 1}{n}\left(1 - \frac{\lfloor\sqrt{n}\rfloor}{n}\right)^{\lfloor\sqrt{n}\rfloor+1}\right) \\
&= \overline{A} \lim_{n\to\infty} \frac{\lfloor\sqrt{n}\rfloor + 1}{n} e^{-\frac{\lfloor\sqrt{n}\rfloor}{n}(\lfloor\sqrt{n}\rfloor+1)} = 0.
\end{aligned}$$

We conclude that the expected payoff of the uninformed sender approaches 0 as the number of states approaches infinity.

The informed sender can persuade the receiver in all states except the 'mine' states, as pooling all non-'mine' states with the top states delivers utility r to the receiver. Thus, using the fact that nature randomises uniformly across all $\binom{n}{\lfloor\sqrt{n}\rfloor}$ choices, the expected payoff of informed sender can be written as

$$\alpha_1 + (1-\alpha_1)\left(1 - \frac{\lfloor\sqrt{n}\rfloor}{n}\right)$$

which approaches 1 as $n \to \infty$.

Thus, as the number of states approaches infinity, the payoff of the informed sender approaches 1, while the payoff of the uninformed sender approaches 0, meaning that $\overline{L} \to 1$.

$\square$

Appendix C: Technical Proofs

In Lemma C.1 we establish that there exists a unique solution for

$$H(\lambda) = 0, \tag{41}$$

which we denote λ^*. Then, in the proof of the main theorem, we will show that this solution is precisely the jump point in the optimal strategy of the sender described in Lemma 5, i.e. $\lambda_0 = \lambda^*$. Moreover, in equilibrium the lower bounds of the supports of F and G are defined by $\underline{\mu}(\lambda_0)$.

Lemma C.1. *There exists a unique solution λ^* of equation* (41) *such that $\underline{\mu}(\lambda^*) < \lambda^* \leq \overline{\lambda}$.*

Proof. In order to resolve the maximum and the minimum operators in (19), it is useful to define the following functions on $[\underline{\lambda}, \overline{\lambda}]$:

$$H_1(\lambda) \equiv \ln\left(\frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\lambda)}\right) - e^{-C(\lambda)},$$
$$H_2(\lambda) \equiv \ln\left(\frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\lambda)}\right) - \overline{\phi}(\underline{\lambda})(1 - C(\lambda) - \ln\overline{\phi}(\underline{\lambda})).$$

Note that $H(\lambda) = H_1(\lambda)$ for any λ satisfying $\overline{\phi}(\underline{\lambda}) < e^{-C(\lambda)}$ and $H(\lambda) = H_2(\lambda)$ for any λ satisfying $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda)}$. We will show that there is a unique λ^* that either satisfies $H_1(\lambda^*) = 0$ and $\overline{\phi}(\underline{\lambda}) < e^{-C(\lambda^*)}$ or $H_2(\lambda^*) = 0$ and $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda^*)}$.

We start by considering the degenerate case in which $\underline{\lambda} = \overline{\lambda}$. If $\underline{\lambda} = \overline{\lambda}$, then it is straightforward to see that $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\underline{\lambda})}$ and therefore $\underline{\lambda}$ solves (19) since $H(\underline{\lambda}) = H_2(\underline{\lambda}) = -\overline{\phi}(\underline{\lambda}) \ln\left(\frac{\overline{\phi}(\overline{\lambda})}{\overline{\phi}(\underline{\lambda})}\right) = 0$, which implies that $\lambda^* = \underline{\lambda} = \overline{\lambda}$.

In what follows we assume that $\underline{\lambda} < \overline{\lambda}$. We separate the problem in three cases and show the existence and uniqueness of the solution of (41) for every case.

Case 1: $e < \frac{\overline{\phi}(\overline{\lambda})}{\overline{\phi}(\underline{\lambda})}$. First, we show that in this case we have that $\overline{\phi}(\underline{\lambda}) < e^{-C(\lambda)}$ for all

$\lambda \in [\underline{\lambda}, \overline{\lambda}]$. Note that function $C(\cdot)$ defined in (17) weakly increases in λ since

$$\frac{dC}{d\lambda} = \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda)} - \frac{1-\overline{\alpha}}{\overline{\phi}(\lambda)} = \frac{\overline{\alpha}-\underline{\alpha}}{\underline{\phi}(\lambda)\overline{\phi}(\lambda)} \geq 0.$$

Consequently, $e^{-C(\lambda)}$ weakly decreases on $[\underline{\lambda}, \overline{\lambda}]$ and therefore for any $\lambda \in [\underline{\lambda}, \overline{\lambda}]$ we have that

$$e^{-C(\lambda)} \geq e^{-C(\overline{\lambda})} = \frac{\overline{\phi}(\overline{\lambda})}{e} > \overline{\phi}(\underline{\lambda}),$$

where the last inequality follows from the assumption $e < \frac{\overline{\phi}(\overline{\lambda})}{\overline{\phi}(\underline{\lambda})}$. This implies that for all $\lambda \in [\underline{\lambda}, \overline{\lambda}]$ we have that $H(\lambda) = H_1(\lambda)$ and $\underline{\mu}(\lambda) = \frac{1}{1-\overline{\alpha}}\left(e^{-C(\lambda)} - \overline{\alpha}\right)$.

Second, we show that there exists a unique $\lambda^* \in [\underline{\mu}(\lambda^*), \overline{\lambda}]$ that solves $H_1(\lambda) = 0$. Define t as a solution of $\underline{\phi}(t)e = \underline{\phi}(\overline{\lambda})$. It is easy to see that $t \in (\underline{\lambda}, \overline{\lambda})$ and is uniquely defined.[15] Note that $H_1(t) = 1 - \overline{\phi}(t) > 0$ and $H_1(\overline{\lambda}) = -e^{-C(\lambda_0)} < 0$. From Lemma C.2 we have that $H'(\lambda) < 0$ and therefore there exists unique $\lambda^* \in (t, \overline{\lambda})$ that solves $H_1(\lambda^*) = 0$.

Third, it remains to show that $\lambda^* > \mu(\lambda^*)$. Note that $\lambda^* > t$ implies that $\underline{\phi}(\lambda^*)e > \underline{\phi}(\overline{\lambda})$ and therefore

$$\overline{\phi}(\lambda^*) > \frac{\underline{\phi}(\overline{\lambda})}{e\underline{\phi}(\lambda^*)}\overline{\phi}(\lambda^*) = e^{-C(\lambda^*)} = \overline{\phi}(\underline{\mu}(\lambda^*)).$$

Since $\overline{\phi}(\cdot)$ is an increasing function we obtain that $\lambda^* > \underline{\mu}(\lambda^*)$.

Case 2: $\frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\underline{\lambda})} < e$. First, we show that in this case we have that $\overline{\phi}(\underline{\lambda}) > e^{-C(\lambda)}$ for all $\lambda \in [\underline{\lambda}, \overline{\lambda}]$. Since $C(\cdot)$ increases in λ we have that for any $\lambda \in [\underline{\lambda}, \overline{\lambda}]$

$$e^{-C(\lambda)} \leq e^{-C(\underline{\lambda})} = \frac{\underline{\phi}(\overline{\lambda})}{e\underline{\phi}(\underline{\lambda})}\overline{\phi}(\underline{\lambda}) < \overline{\phi}(\underline{\lambda}).$$

This implies that $H(\lambda) = H_2(\lambda)$ for all $\lambda \in [\underline{\lambda}, \overline{\lambda}]$ and the lower bound of the equilibrium distribution defined in (18) is $\underline{\mu}(\lambda) = \underline{\lambda}$.

[15]Since $\frac{\alpha+(1-\alpha)\overline{\lambda}}{\alpha+(1-\alpha)\underline{\lambda}}$ strictly decreases in α we have that $\frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\underline{\lambda})} > \frac{\overline{\phi}(\overline{\lambda})}{\overline{\phi}(\underline{\lambda})}$. By assumption $e < \frac{\overline{\phi}(\overline{\lambda})}{\overline{\phi}(\underline{\lambda})}$ we obtain that $\underline{\phi}(\underline{\lambda})e < \underline{\phi}(\overline{\lambda})$. Thus, since $\underline{\phi}(\overline{\lambda})e > \underline{\phi}(\overline{\lambda})$, we obtain that there is unique $t \in (\underline{\lambda}, \overline{\lambda})$ that solves $\underline{\phi}(t)e = \underline{\phi}(\overline{\lambda})$.

Second, we show in this case there exists a unique $\lambda^* \in (\underline{\lambda}, \overline{\lambda})$ that solves $H_2(\lambda) = 0$. Note that $H_2(\underline{\lambda}) = (1 - \overline{\phi}(\underline{\lambda})) \ln\left(\frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\underline{\lambda})}\right) > 0$, $H_2(\overline{\lambda}) = -\overline{\phi}(\underline{\lambda}) \ln\left(\frac{\overline{\phi}(\overline{\lambda})}{\overline{\phi}(\underline{\lambda})}\right) < 0$. From Lemma C.2 we have that $H_2(\lambda)$ strictly decreases on $[\underline{\lambda}, \overline{\lambda}]$ and therefore there exists a unique $\lambda^* \in (\underline{\lambda}, \overline{\lambda})$ that solves $H_2(\lambda) = 0$.

Case 3: $\frac{\overline{\phi}(\overline{\lambda})}{\overline{\phi}(\underline{\lambda})} \leq e \leq \frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\underline{\lambda})}$. First, we show that in this case there exists $\tilde{\lambda} \in [t, \overline{\lambda}]$ such that $e^{-C(\lambda)} \geq (<) \overline{\phi}(\underline{\lambda})$ if and only if $\lambda \leq (>) \tilde{\lambda}$. To show this we note that $e^{-C(\underline{\lambda})} = \frac{\underline{\phi}(\overline{\lambda})}{e\underline{\phi}(\underline{\lambda})} \overline{\phi}(\underline{\lambda}) \geq \overline{\phi}(\underline{\lambda})$ and $e^{-C(\overline{\lambda})} = \overline{\phi}(\overline{\lambda})/e \leq \overline{\phi}(\underline{\lambda})$. Therefore, since $C(\cdot)$ strictly increases on $[\underline{\lambda}, \overline{\lambda}]$, there exists a unique $\tilde{\lambda}$ that solves $e^{-C(\tilde{\lambda})} = \overline{\phi}(\underline{\lambda})$. It remains show that $\tilde{\lambda} > t$. By using the definition of t we have that

$$e^{-C(t)} = \frac{\underline{\phi}(\overline{\lambda})}{e\underline{\phi}(t)} \overline{\phi}(t) = \overline{\phi}(t) > \overline{\phi}(\underline{\lambda}) = e^{-C(\tilde{\lambda})}.$$

Since $C(\cdot)$ is a strictly increasing function on $[\underline{\lambda}, \overline{\lambda}]$ we have that $\tilde{\lambda} > t$.

Second, for what follows, it is useful to show that $H_2(\lambda) \geq H_1(\lambda)$ for all $\lambda \in [\underline{\lambda}, \overline{\lambda}]$ and $H_2(\lambda) = H_1(\lambda)$ only at $\lambda = \tilde{\lambda}$. To see this we explore the difference between functions $H_1(\cdot)$ and $H_2(\cdot)$ on $[\underline{\lambda}, \overline{\lambda}]$ – that is,

$$H_1(\lambda) - H_2(\lambda) = \overline{\phi}(\underline{\lambda})(1 - C(\lambda) - \ln \overline{\phi}(\underline{\lambda})) - e^{-C(\lambda)}.$$

By taking the derivative of $H_1(\cdot) - H_2(\cdot)$ we obtain that

$$\frac{d(H_1(\lambda) - H_2(\lambda))}{d\lambda} = (e^{-C(\lambda)} - \overline{\phi}(\underline{\lambda}))C'(\lambda),$$

which implies that $H_1(\lambda) - H_2(\lambda)$ is a hump-shaped function that reaches its maximum that is equal to 0 at $\lambda = \tilde{\lambda}$. Therefore, we can conclude that $H_2(\lambda) \geq H_1(\lambda)$ for all $\lambda \in [\underline{\lambda}, \overline{\lambda}]$ and $H_2(\lambda) = H_1(\lambda)$ only at $\lambda = \tilde{\lambda}$.

Third, we are ready to show that there exists a solution of (41). From the analysis of case 1 and case 2 we have established that equations $H_1(\lambda) = 0$ and $H_2(\lambda) = 0$ have unique roots on $[\underline{\lambda}, \overline{\lambda}]$ that we denote by λ_1 and λ_2 respectively. Moreover, $\lambda_1 > t$ and $\lambda_2 > \underline{\lambda}$.

Suppose for a contradiction that there is no solution of (41), which implies that $\lambda_1 > \tilde{\lambda}$ and $\lambda_2 < \tilde{\lambda}$ hold simultaneously. If $\lambda_1 > \tilde{\lambda}$, then $H_2(\tilde{\lambda}) = H_1(\tilde{\lambda}) > 0$ as $H_1(\cdot)$ strictly decreases on $[t, \overline{\lambda}]$. Since $H_2(\cdot)$ is a strictly decreasing function on $[\underline{\lambda}, \overline{\lambda}]$ and $H_2(\tilde{\lambda}) > 0$ we have that $\lambda_2 > \tilde{\lambda}$, a contradiction. If $\lambda_2 < \tilde{\lambda}$, then it must be that $H_1(\tilde{\lambda}) = H_2(\tilde{\lambda}) < 0$. Since H_1 is a strictly decreasing function and $H_1(t) > 0$ we have that $t < \lambda_1 < \tilde{\lambda}$ and we arrive to a contradiction. This implies that there is at least one root of (41) on $[\underline{\lambda}, \overline{\lambda}]$.

Forth, we show the uniqueness of the solution of (41) on $[\underline{\lambda}, \overline{\lambda}]$. Suppose for a contradiction that there exist $\lambda_1^* \neq \lambda_2^*$ such that $\tilde{\lambda} > \lambda_1^* > t$, $H_1(\lambda_1^*) = 0$ and $\lambda_2^* \geq \tilde{\lambda}$, $H_2(\lambda_2^*) = 0$.

$$0 = H_2(\lambda_2^*) \leq H_2(\tilde{\lambda}) = H_1(\tilde{\lambda}) < H_1(\lambda_1^*),$$

so we arrive to a contradiction.

Finally, it remains to prove that $\lambda^* > \underline{\mu}(\lambda^*)$. If $H_2(\lambda^*) = 0$, then $\lambda^* \geq \tilde{\lambda}$ and therefore $\underline{\mu}(\lambda^*) = \underline{\lambda}$. Since $H_2(\underline{\lambda}) > 0$ we have that $\lambda^* > \underline{\lambda}$. If $H_1(\lambda^*) = 0$, then $\lambda^* \leq \tilde{\lambda}$ and $\underline{\mu}(\lambda^*) = \frac{1}{1-\overline{\alpha}}\left(e^{-C(\lambda^*)} - \overline{\alpha}\right)$. Since $\lambda^* > t$ we have that $\underline{\phi}(\lambda^*)e > \underline{\phi}(\overline{\lambda})$ and therefore $\overline{\phi}(\lambda^*) > \overline{\phi}(\underline{\mu}(\lambda^*))$ (by the same argument made in case 1.) Therefore, we conclude that that $\lambda^* > \underline{\mu}(\lambda^*)$.

□

Lemma C.2. $\frac{\partial H}{\partial \lambda_0} = \frac{\overline{\alpha}-\underline{\alpha}}{\overline{\phi}(\lambda_0)\underline{\phi}(\lambda_0)} \max\left\{\overline{\phi}(\underline{\lambda}), e^{-C(\lambda_0)}\right\} - \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)} < -\frac{1-\overline{\alpha}}{\underline{\phi}(\lambda_0)} < 0.$

Proof. First, suppose that $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$. Then

$$\begin{aligned}
\frac{\partial H}{\partial \lambda_0} &= \overline{\phi}(\underline{\lambda})\frac{\partial C}{\partial \lambda_0} - \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)} \\
&= \overline{\phi}(\underline{\lambda})\left(\frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)} - \frac{1-\overline{\alpha}}{\overline{\phi}(\lambda_0)}\right) - \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)} \\
&= \underbrace{\frac{\overline{\phi}(\underline{\lambda})}{\overline{\phi}(\lambda_0)}}_{<1}\frac{\overline{\alpha}-\underline{\alpha}}{\underline{\phi}(\lambda_0)} - \frac{1-\underline{\alpha}}{\underline{\phi}(\lambda_0)} < -\frac{1-\overline{\alpha}}{\underline{\phi}(\lambda_0)} < 0.
\end{aligned}$$

Second, for the case $\overline{\phi}(\underline{\lambda}) < e^{-C(\lambda_0)}$ we use the fact that

$$\begin{aligned} C(\lambda_0) &= 1 - \ln \underline{\phi}(\overline{\lambda}) - \ln \overline{\phi}(\lambda_0) + \ln \underline{\phi}(\lambda_0) \\ &= 1 - \overline{L} - \ln \overline{\phi}(\lambda_0) \\ &> - \ln \overline{\phi}(\lambda_0) \end{aligned} \tag{42}$$

and thus $e^{-C(\lambda_0)} < \overline{\phi}(\lambda_0)$. Using this inequality we obtain that

$$\begin{aligned} \frac{\partial H}{\partial \lambda_0} &= e^{-C(\lambda_0)} \frac{\overline{\alpha} - \underline{\alpha}}{\overline{\phi}(\lambda_0)\underline{\phi}(\lambda_0)} - \frac{1 - \underline{\alpha}}{\underline{\phi}(\lambda_0)} \\ &< \frac{\overline{\alpha} - \underline{\alpha}}{\underline{\phi}(\lambda_0)} - \frac{1 - \underline{\alpha}}{\underline{\phi}(\lambda_0)} < -\frac{1 - \overline{\alpha}}{\underline{\phi}(\lambda_0)} < 0. \end{aligned}$$

So we get that for any parameters $\frac{\partial H}{\partial \lambda_0} < -\frac{1-\overline{\alpha}}{\underline{\phi}(\lambda_0)} < 0$. $\square$

Lemma C.3. $\frac{\partial H}{\partial \overline{\lambda}} = \frac{1-\underline{\alpha}}{\underline{\phi}(\overline{\lambda})} \left[1 - \max\left\{\overline{\phi}(\underline{\lambda}), e^{-C(\lambda_0)}\right\}\right]$.

Proof. First, consider the case in which $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$. Using the partial derivative $\partial C/\partial \overline{\lambda} = -(1 - \underline{\alpha})/\underline{\phi}(\overline{\lambda})$ we find that

$$\frac{\partial H}{\partial \overline{\lambda}} = \frac{1 - \underline{\alpha}}{\underline{\phi}(\overline{\lambda})} + \overline{\phi}(\underline{\lambda}) \frac{\partial C}{\partial \overline{\lambda}} = \frac{(1 - \underline{\alpha})}{\underline{\phi}(\overline{\lambda})}(1 - \overline{\phi}(\underline{\lambda})).$$

Second, assume that $\overline{\phi}(\underline{\lambda}) < e^{-C(\lambda_0)}$, then

$$\frac{\partial H}{\partial \overline{\lambda}} = \frac{1 - \underline{\alpha}}{\underline{\phi}(\overline{\lambda})} + e^{-C(\lambda_0)} \frac{\partial C}{\partial \overline{\lambda}} = \frac{(1 - \underline{\alpha})}{\underline{\phi}(\overline{\lambda})} \left(1 - e^{-C(\lambda_0)}\right).$$

$\square$

Lemma C.4. $\frac{\partial H}{\partial \underline{\lambda}} = (1 - \overline{\alpha}) \max\left\{\ln \overline{\phi}(\underline{\lambda}) + C(\lambda_0), 0\right\}$.

Proof. Suppose that $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$. Then, the partial derivative of H with respect to $\underline{\lambda}$ is

$$\frac{\partial H}{\partial \underline{\lambda}} = (1-\overline{\alpha})(C(\lambda_0) + \ln \overline{\phi}(\underline{\lambda})).$$

Next suppose that $\overline{\phi}(\underline{\lambda}) < e^{-C(\lambda_0)}$, then $H(\lambda_0)$ does not depend on $\underline{\lambda}$, so we have that $\frac{\partial H}{\partial \underline{\lambda}} = 0$. □

Lemma C.5.

$$\frac{\partial H}{\partial \overline{\alpha}} = \begin{cases} -e^{-C(\lambda_0)} \frac{1-\lambda_0}{\overline{\phi}(\lambda_0)}, & \overline{\phi}(\underline{\lambda}) < e^{-C(\lambda_0)} \\ (1-\underline{\lambda})(C(\lambda_0) + \ln \overline{\phi}(\underline{\lambda})) - \overline{\phi}(\underline{\lambda}) \frac{1-\lambda_0}{\overline{\phi}(\lambda_0)}, & \overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)} \end{cases}.$$

Moreover, $\frac{\partial H}{\partial \overline{\alpha}} \leq 0$.

Proof. We start with the case $\overline{\phi}(\underline{\lambda}) < e^{-C(\lambda_0)}$. The partial derivative of $H(\lambda_0)$ with respect to $\overline{\alpha}$ is given by

$$\frac{\partial H}{\partial \overline{\alpha}} = e^{-C(\lambda_0)} \frac{\partial C}{\partial \overline{\alpha}} = -e^{-C(\lambda_0)} \frac{1-\lambda_0}{\overline{\phi}(\lambda_0)} < 0 \tag{43}$$

Suppose that $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$. The partial derivative of $H(\lambda_0)$ with respect to $\overline{\alpha}$ is

$$\begin{aligned} \frac{\partial H}{\partial \overline{\alpha}} &= -(1-\underline{\lambda})(1 - C(\lambda_0) - \ln \overline{\phi}(\underline{\lambda})) - \overline{\phi}(\underline{\lambda}) \left(-\frac{\partial C}{\partial \overline{\alpha}} - \frac{1-\underline{\lambda}}{\overline{\phi}(\underline{\lambda})} \right) \\ &= (1-\underline{\lambda})(C(\lambda_0) + \ln \overline{\phi}(\underline{\lambda})) - \overline{\phi}(\underline{\lambda}) \frac{1-\lambda_0}{\overline{\phi}(\lambda_0)}. \end{aligned} \tag{44}$$

It remains to establish the sign of $\frac{\partial H}{\partial \overline{\alpha}}$ in this case. We will prove that $\frac{\partial H}{\partial \overline{\alpha}} < 0$ in the following steps:

1. We show that if $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$, then at the neighbourhood of any point $\overline{\alpha}$ where $\frac{\partial H}{\partial \overline{\alpha}} \geq 0$ it must be the case that $\frac{d}{d\overline{\alpha}} \left(\frac{\partial H}{\partial \overline{\alpha}} \right) > 0$. That is, there may be at most one point of intersection with zero.

2. We show that for $\overline{\alpha}$ in the neighbourhood of 1 we are in the case $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$ and we have $\lim_{\overline{\alpha}\to 1} \frac{\partial H}{\partial \overline{\alpha}} = 0$, meaning that $\frac{\partial H}{\partial \overline{\alpha}}$ is negative for high $\overline{\alpha}$.

3. Finally, we show that $\frac{\partial H}{\partial \overline{\alpha}}$ is continuous when transitioning between cases $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$ and $\overline{\phi}(\underline{\lambda}) < e^{-C(\lambda_0)}$, concluding that it is negative for all $\overline{\alpha} < 1$.

Claim 1: if $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$, then at the neighbourhood of any point $\overline{\alpha}$ where $\frac{\partial H}{\partial \overline{\alpha}} \geq 0$ it must be the case that $\frac{d}{d\overline{\alpha}}\left(\frac{\partial H}{\partial \overline{\alpha}}\right) > 0$.

Note that $\frac{\partial H}{\partial \overline{\alpha}}$ depends on $\overline{\alpha}$ directly and via λ_0. Thus, we obtains that

$$\frac{d}{d\overline{\alpha}}\left(\frac{\partial H}{\partial \overline{\alpha}}\right) = \frac{\partial^2 H}{\partial \overline{\alpha}\partial \lambda_0}\frac{d\lambda_0}{d\overline{\alpha}} + \frac{\partial^2 H}{\partial \overline{\alpha}^2}. \tag{45}$$

Firstly, we the second derivative of $H(\lambda_0)$ with respect to $\overline{\alpha}$

$$\begin{aligned}
\frac{\partial^2 H}{\partial \overline{\alpha}^2} &= (1-\underline{\lambda})\left(\frac{\partial C}{\partial \overline{\alpha}} + \frac{(1-\underline{\lambda})}{\overline{\phi}(\underline{\lambda})}\right) - \frac{(1-\underline{\lambda})(1-\lambda_0)}{\overline{\phi}(\lambda_0)} + \overline{\phi}(\underline{\lambda})\frac{(1-\lambda_0)^2}{\overline{\phi}^2(\lambda_0)} \\
&= \frac{1}{(1-\overline{\alpha})^2}(1-\overline{\phi}(\underline{\lambda}))\left(-\frac{1-\overline{\phi}(\lambda_0)}{\overline{\phi}(\lambda_0)} + \frac{1-\overline{\phi}(\underline{\lambda})}{\overline{\phi}(\underline{\lambda})}\right) \\
&+ \frac{1}{(1-\overline{\alpha})^2}(1-\overline{\phi}(\lambda_0))\left(-\frac{1-\overline{\phi}(\underline{\lambda})}{\overline{\phi}(\lambda_0)} + \frac{\overline{\phi}(\underline{\lambda})(1-\overline{\phi}(\lambda_0))}{\overline{\phi}^2(\lambda_0)}\right) \\
&= \frac{1}{(1-\overline{\alpha})^2}\frac{\overline{\phi}(\lambda_0)-\overline{\phi}(\underline{\lambda})}{\overline{\phi}(\lambda_0)}\left(\frac{1-\overline{\phi}(\underline{\lambda})}{\overline{\phi}(\underline{\lambda})} - \frac{1-\overline{\phi}(\lambda_0)}{\overline{\phi}(\lambda_0)}\right) \\
&= \frac{1}{(1-\overline{\alpha})^2}\frac{(\overline{\phi}(\lambda_0)-\overline{\phi}(\underline{\lambda}))^2}{\overline{\phi}^2(\lambda_0)\overline{\phi}(\underline{\lambda})} > 0.
\end{aligned}$$

Secondly, we compute the cross-partial derivative of $H(\lambda_0)$ with respect to λ_0 and $\overline{\alpha}$ – that is,

$$\frac{\partial^2 H}{\partial \overline{\alpha}\partial \lambda_0} = \frac{\overline{\phi}(\underline{\lambda}}{\overline{\phi}(\lambda_0)\underline{\phi}(\lambda_0)} + \frac{(\overline{\alpha}-\underline{\alpha})(\lambda_0-\underline{\lambda})}{\underline{\phi}(\lambda_0)[\overline{\phi}(\lambda_0)]^2},$$

which is positive and bounded.

Thirdly, note that as $\frac{d\lambda_0}{d\overline{\alpha}} = -\frac{\partial H/\partial \overline{\alpha}}{\partial H/\partial \lambda_0}$, $\frac{\partial H}{\partial \overline{\alpha}}$ is a continuous function and $\frac{\partial H}{\partial \lambda_0} \leq 0$. Thus, assuming $\frac{\partial H}{\partial \overline{\alpha}} \geq 0$, from equation (45) we have that

$$\frac{d}{d\overline{\alpha}}\left(\frac{\partial H}{\partial \overline{\alpha}}\right) = \frac{\partial^2 H}{\partial \overline{\alpha} \partial \lambda_0}\frac{d\lambda_0}{d\overline{\alpha}} + \frac{\partial^2 H}{\partial \overline{\alpha}^2} > 0,$$

which completes the proof of claim 1.

Claim 2: for $\overline{\alpha}$ in the neighbourhood of 1 we are in the case $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$ and we have $\lim_{\overline{\alpha}\to 1} \frac{\partial H}{\partial \overline{\alpha}} = 0$.

We prove the first part of the claim by contradiction. Suppose that for some $\varepsilon > 0$ we have that for all $\overline{\alpha} \in [1-\varepsilon, 1]$ inequality $\overline{\phi}(\underline{\lambda}) < e^{-C(\lambda_0)}$ holds. Note that $H(\lambda_0) = 0$, Thus,

$$0 = \frac{H(\lambda_0)}{\overline{\phi}(\underline{\lambda})} = \ln\left(\frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\lambda_0)}\right) \Big/ \overline{\phi}(\underline{\lambda}) - \frac{e^{-C(\lambda_0)}}{\overline{\phi}(\underline{\lambda})} < 1 - \frac{e^{-C(\lambda_0)}}{\overline{\phi}(\underline{\lambda})},$$

which implies $\overline{\phi}(\underline{\lambda}) > e^{-C(\lambda_0)}$, so we arrive at contradiction.

By rewriting equation (41) we obtain

$$\begin{aligned}
0 = H(\lambda_0) &= \ln\frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\lambda_0)} - \overline{\phi}(\underline{\lambda})\left(1 - C(\lambda_0) - \ln\overline{\phi}(\underline{\lambda})\right) \\
&= \ln\frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\lambda_0)} - \overline{\phi}(\underline{\lambda})\left(\ln\frac{\overline{\phi}(\lambda_0)}{\overline{\phi}(\underline{\lambda})} + \ln\frac{\underline{\phi}(\overline{\lambda})}{\underline{\phi}(\lambda_0)}\right) \\
&= \overline{L}(1 - \overline{\phi}(\underline{\lambda})) - \overline{\phi}(\underline{\lambda})\ln\frac{\overline{\phi}(\lambda_0)}{\overline{\phi}(\underline{\lambda})}.
\end{aligned}$$

We divide by $1 - \overline{\alpha}$ and take the limit when $\overline{\alpha} \to 1$

$$\begin{aligned}
0 &= (1 - \underline{\lambda})\lim_{\overline{\alpha}\to 1}\overline{L} - \lim_{\overline{\alpha}\to 1}\frac{\ln\frac{\overline{\phi}(\lambda_0)}{\overline{\phi}(\underline{\lambda})}}{1 - \overline{\alpha}} \\
&= (1 - \underline{\lambda})\lim_{\overline{\alpha}\to 1}\overline{L} - \lim_{\overline{\alpha}\to 1}(\lambda_0 - \underline{\lambda}).
\end{aligned}$$

By taking the limit of $\frac{\partial H}{\partial \overline{\alpha}}$ when $\overline{\alpha} \to 1$, we have that

$$\begin{aligned}\lim_{\overline{\alpha}\to 1} \frac{\partial H}{\partial \overline{\alpha}} &= \lim_{\overline{\alpha}\to 1} \left[(1-\underline{\lambda}) \left(1 - \frac{\overline{L}}{\overline{\phi}(\underline{\lambda})} \right) - \overline{\phi}(\underline{\lambda}) \frac{1-\lambda_0}{\overline{\phi}(\lambda_0)} \right] \\ &= (1-\underline{\lambda}) \lim_{\overline{\alpha}\to 1} (1-\overline{L}) - \lim_{\overline{\alpha}\to 1} (1-\lambda_0) \\ &= \lim_{\overline{\alpha}\to 1} (\lambda_0 - \underline{\lambda}) - (1-\underline{\lambda}) \lim_{\overline{\alpha}\to 1} \overline{L} = 0,\end{aligned}$$

completing the second part of the claim.

Claim 3: $\frac{\partial H}{\partial \overline{\alpha}}$ is continuous when transitioning between cases $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$ and $\overline{\phi}(\underline{\lambda}) < e^{-C(\lambda_0)}$, concluding that it is negative for all $\overline{\alpha} < 1$.

By comparing $\frac{\partial H}{\partial \overline{\alpha}}$ for both cases at $\overline{\phi}(\underline{\lambda}) = e^{-C(\lambda_0)}$ we see that values coincide, so function $\frac{\partial H}{\partial \overline{\alpha}}$ is continuous.

Now, suppose there is $\overline{\alpha}$ such that $\frac{\partial H}{\partial \overline{\alpha}} > 0$. From (43) we know that then it must be that $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$, as otherwise the derivative would be negative. From Claim 1 we know that if $\frac{\partial H}{\partial \overline{\alpha}} > 0$ and $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$ then $\frac{\partial H}{\partial \overline{\alpha}}$ must be increasing in $\overline{\alpha}$. However, from Claim 2 we have that for $\overline{\alpha}$ high enough $\frac{\partial H}{\partial \overline{\alpha}} < 0$ and $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$, arriving to a contradiction due to continuity of $\frac{\partial H}{\partial \overline{\alpha}}$ within each case and when transitioning between the cases. Thus, $\frac{\partial H}{\partial \overline{\alpha}} < 0$ for all $\overline{\alpha} < 1$.

$\square$

Lemma C.6. $\frac{\partial H}{\partial \underline{\alpha}} = -\frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)} \left[1 - \max\left\{\overline{\phi}(\underline{\lambda}), e^{-C(\lambda_0)}\right\}\right]$.

Proof. Consider the case $\overline{\phi}(\underline{\lambda}) \geq e^{-C(\lambda_0)}$. The partial derivative of H with respect to $\underline{\alpha}$ is

$$\frac{\partial H}{\partial \underline{\alpha}} = -\frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)} + \overline{\phi}(\underline{\lambda}) \frac{\partial C}{\partial \underline{\alpha}} = -\frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)} (1 - \overline{\phi}(\underline{\lambda})).$$

Next suppose that $\overline{\phi}(\underline{\lambda}) < e^{-C(\lambda_0)}$, then

$$\frac{\partial H}{\partial \underline{\alpha}} = -\frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)} + e^{-C(\lambda_0)} \frac{\partial C}{\partial \underline{\alpha}} = -\frac{\overline{\lambda}-\lambda_0}{\underline{\phi}(\overline{\lambda})\underline{\phi}(\lambda_0)} \left(1 - e^{-C(\lambda_0)}\right),$$

which gives the result of the Lemma. $\square$

References

Alonso, R. and O. Camara (2016): "Bayesian persuasion with heterogeneous priors," *Journal of Economic Theory*, 165, 672–706.

Babichenko, Y., I. Talgam-Cohen, H. Xu, and K. Zabarnyi (2021): "Regret-minimizing bayesian persuasion," *Proceedings of the 22nd ACM Conference on Economics and Computation.*

Bergemann, D. and S. Morris (2019): "Information design: A unified perspective," *Journal of Economic Literature*, 57, 44–95.

Bergemann, D. and K. Schlag (2011): "Robust monopoly pricing," *Journal of Economic Theory*, 146, 2527–2543.

Bergemann, D. and K. H. Schlag (2008): "Pricing without priors," *Journal of the European Economic Association*, 6, 560–569.

Best, J. and D. Quigley (2020): "Persuasion for the long run," *Available at SSRN 2908115.*

Dworczak, P. and A. Pavan (2020): "Preparing for the worst but hoping for the best: Robust (bayesian) persuasion," *CEPR Discussion Paper No. DP15017.*

Guo, Y. and E. Shmaya (2019a): "The interval structure of optimal disclosure," *Econometrica*, 87, 653–675.

——— (2019b): "Robust monopoly regulation," *arXiv preprint arXiv:1910.04260.*

Hewitt, E. and K. Stromberg (2013): *Real and abstract analysis: a modern treatment of the theory of functions of a real variable*, Springer-Verlag.

Hu, J. and X. Weng (2021): "Robust persuasion of a privately informed receiver," *Economic Theory*, 72, 909–953.

Kamenica, E. (2019): "Bayesian persuasion and information design," *Annual Review of Economics*, 11, 249–272.

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

Kolotilin, A., T. Mylovanov, A. Zapechelnyuk, and M. Li (2017): "Persuasion of a privately informed receiver," *Econometrica*, 85, 1949–1964.

Kosterina, S. (2022): "Persuasion with unknown beliefs," *Theoretical Economics*, forthcoming.

Laclau, M. and L. Renou (2017): "Public persuasion," *Working Paper*.

Manski, C. F. (2004): "Statistical treatment rules for heterogeneous populations," *Econometrica*, 72, 1221–1246.

Parakhonyak, A. and A. Sobolev (2015): "Non-Reservation Price Equilibrium and Search without Priors," *The Economic Journal*, 125, 887–909.

Parakhonyak, A. and N. Vikander (2019): "Information design through scarcity and social learning," *mimeo*.

Schlag, K. and A. Sobolev (2022): "Robust search: How to learn from the past without priors?" *Working paper*.

Schlag, K. H. and A. Zapechelnyuk (2021): "Robust sequential search," *Theoretical Economics*, 16, 1431–1470.