The Hidden Value of Lying: Evasion of Guilt in Expert Advice

Kiryl Khalmetski*

March 31, 2018

Abstract

I develop a model of strategic communication between an uninformed receiver and a partially informed sender who is averse to lying. The sender's cost of lying is endogenous, depending on the receiver's *beliefs* induced by the sender's message, rather than on its exogenous formulation. Such preferences lead to the endogenous emergence of evasive communication, i.e., avoidance to make explicit claims about the state of the world. In turn, this gives rise to specific predictions regarding welfare implications of several conventional policies. In particular, prohibition of lying (i.e., of explicit falsification) may lead to a decrease in the receiver's welfare under certain conditions. Besides, dealing with ex-ante less informed sender can be beneficial to the receiver.

Keywords: guilt aversion, information transmission, experts, psychological game theory.

JEL codes: D82, D83, D84, C72, L51.

1 Introduction

Consumers often lack sufficient knowledge to optimally make certain purchase or investment decisions. In these cases, they must rely on more sophisticated experts, such as financial advisors, doctors, and consultants. However, incentives affecting these experts can be inconsistent with what consumers want: truthful, unbiased advice that helps them to choose the most suitable option. One common example is the commissions that financial advisors receive if their clients buy specific products, independently of whether these

---

*University of Cologne, Department of Economics, Albertus-Magnus-Platz, 50923 Cologne, Germany (e-mail: kiryl.khalmetski@uni-koeln.de). I am thankful to Matthias Blonski, Ruben Enikolopov, Guido Friebel, Andreas Hackethal, Roman Inderst, Michael Kosfeld, Axel Ockenfels and Ferdinand von Siemens for helpful comments. Financial support of the German Research Foundation (DFG) through the Research Unit "Design & Behavior" (FOR 1371) is gratefully acknowledged.

products match consumer needs (Inderst and Ottaviani 2012). Even doctors are often incentivized to provide a specific medical treatment (Gruber et al. 1999). The scope of potential fraud is large enough that there are extensive regulations aimed at mitigating conflict of interest, or prosecuting fraudulent advice. For example, the UK Financial Services Authority has implemented bans on commissions paid to independent financial advisors by product providers (Collinson 2012).

Still, even in the presence of clear financial incentives for biased advice, consumers considerably rely on this service in practice.[1] Thereby, they rely also on the indirect costs arising for the expert from deceiving the consumer. For example, deception can lead to reputational loss (Bolton et al. 2007), reclamation costs (Inderst and Ottaviani 2013) or psychological costs, which arise from intrinsic concern for the well-being of the other party (McGuire 2000).

The present paper examines how the expert's incentive to avoid deceiving the consumer, which countervails his monetary conflict of interest, affects the informativeness of advice. A distinctive feature of the modeling approach is that the expert's cost of deception depends not on the message formulation *per se*, but rather on the receiver's beliefs associated with the message. In particular, the expert suffers a utility loss if the beliefs that his message induces do not match the realized outcome. Formally, this corresponds to the concept of guilt aversion (Battigalli and Dufwenberg 2007).[2] A notable implication of these preferences is the endogenous emergence of evasive communication in equilibrium (i.e., avoidance to make explicit claims about the state of the world). As a result, some conventional policies aimed at increasing transparency (such as prohibition of lying) can ultimately backfire.

In my model, the expert and the consumer are called "the sender" (he) and "the receiver" (she), respectively. With some probability, the sender observes the state of the world, which can be either good or bad, while with the remaining probability he remains uninformed. Then, the sender sends a message to the receiver out of an arbitrarily large message space. Finally, the receiver must decide between a risky action (investment) and a riskless action (abstaining), with the former having a positive payoff for her only in the good state.

The sender is biased to always induce investment independently from the state of the world, while at the same time being sensitive to guilt toward the receiver. Guilt is determined by the discrepancy between the receiver's payoff expectation conditional on the sender's message, and the ex post receiver's payoff. The sender's guilt sensitivity is unobservable to the receiver. Thus, the sender is characterized by both the information that he has observed and his guilt sensitivity (the latter referred to as the sender's type).

There are two robust equilibria in this game. In the "lying" equilibrium, the only

[1] For example, a large online survey by Chater et al. (2010) shows that nearly 58 percent of purchasers of investment products are influenced by advisors.

[2] Regarding guilt aversion, empirical evidence is documented in Guerra and Zizzo (2004), Charness and Dufwenberg (2006), Reuben et al. (2009), Khalmetski et al. (2015) and Khalmetski (2016), among others. See also Ellingsen et al. (2010) and Vanberg (2008) for conflicting evidence; see Khalmetski et al. (2015) for a discussion.

message leading to investment is the one which is sent by all types observing the good state. Hence, for types in other informational states the only way to induce investment is to pool with these types on their message. Such equilibrium arises if all other possible messages are not sufficiently credible to induce investment. In the other type of equilibrium, the "evasion" equilibrium, sender types who observe the bad state of the world do not lie explicitly by pooling with the types who observe the good state, but rather send an "evasive" message pooling with the uninformed types (who themselves also prefer to separate from the types observing the good state). The evasive message affects the receiver's equilibrium beliefs less intensively than the highest message in the lying equilibrium, and hence is less costly for the sender in terms of expected guilt.

This equilibrium structure has an important interaction with policy aimed at lying prohibition, e.g., introducing a sufficiently high fine for lying. Under plausible assumptions on verifiability, the only type of lying which can be effectively prohibited by a regulator is the explicit lying in the lying equilibrium, while the "evasive" lying in the evasion equilibrium can not be verified. Hence, the policy eliminates the lying equilibrium, while the evasion equilibrium is fully robust to it. As a result, there can be a shift from the lying to the evasion equilibrium, which can be beneficial or detrimental to the receiver, depending on the degree of conflict of interest. In particular, the policy can backfire by providing the sender a psychologically cheap way for deception through the evasive message, and hence increasing the rate of deception of sender types who observe the bad state. Besides, lying prohibition can preclude any efficient communication of uninformed types due to excessive pooling of bad types on the evasive message, which leads to the loss of its credibility.

Further, another conventional policy, that of mitigating the monetary conflict of interest between the sender and the receiver, can backfire due to the effect of guilt aversion. In particular, if the ex ante share of uninformed senders is sufficiently high, then banning sender commissions destroys all equilibria except for the babbling equilibrium where the receiver always invests independently of the message (the worst possible equilibrium from the receiver's perspective). This occurs due to the fact that uncertain types become motivated primarily by the risk of letting down the receiver, and thus, begin to excessively pool with types who have observed the bad state. This, in turn, leads to investment even after the lowest message and hence precludes emergence of any informative equilibrium.

I also consider the welfare effects of the policy aimed at increasing the quality of the sender's private information. One of the results is that dealing with an (ex ante) *less* informed sender can be preferable for the receiver. This occurs due to the fact that the ex-ante probability that the sender is uninformed affects the receiver's beliefs conditional on the message, and hence the expected guilt of the sender. As a result, an (ex ante) more knowledgeable sender can be, at the same time, more prone to deception. Under certain parameter values, this effect is sufficiently strong to outweigh the positive effect of higher quality of the sender's private information.

Finally, I compare my results with those obtained within a purely outcome-based model and find that most effects (e.g., potentially negative welfare effects of a lying prohibition policy) can be explained only by belief-dependent preferences, like guilt aversion.

My study relates to several strands of literature. The aversion to lying is experimentally documented by Gneezy (2005) and analyzed by Kartik et al. (2007) and Kartik (2009). In the latter papers, the cost of lying is modeled as exogenous, depending only on how much the exogenously given formulation of a message quantitatively deviates from the truth. While such approach can address a broad range of situations (like reporting company profits to shareholders), there are limits to its applicability. First, in some cases, states of the world, which can be reported, cannot be ranked quantitatively (e.g., possible diagnoses of a patient), so that different possible lies cannot be compared by severity based only on message formulations. Second, there are many ways in which the expert can manipulate or mitigate explicit message formulations while conveying the same meaning (e.g., euphemisms). Finally, the expert may avoid disclosing some of his information, leaving the receiver in uncertainty, which can be still harmful to a degree comparable to that of explicit lying. The psychological costs of such behavior cannot be analyzed based on message formulation, since no message is used. In contrast, my approach based on guilt aversion provides a more widely applicable measure of deception which can be used to analyze all these cases: the difference between expectations induced by the advice and the actually realized outcome. Besides, it allows a natural explanation of such empirical phenomena as vague or evasive communication, while yielding specific policy predictions.

Khalmetski et al. (2017) consider a structurally similar setting (both theoretically and experimentally), where a sender, who may be either informed or uninformed about a binary state of the world, sends a message to a risk averse receiver. They also distinguish evasion and lying strategies (terming them as "evasive" and "direct" lying, respectively), showing that senders may use evasion to sidestep higher psychological costs from outright lying. At the same time, their model uses a reduced-form approach by exogenously assuming that the (fixed) intrinsic cost of evasive lying is lower than the cost of direct lying. In the current setting, this basic feature of the model is derived endogenously, which allows for further theoretical implications.

The role of guilt aversion in communication (in particular, promise keeping) was studied by Charness and Dufwenberg (2006). However, in their setting communication serves mainly as a commitment device for a guilt-averse agent, who can use it to coordinate on specific equilibria. The current setting is different in that communication also resolves information asymmetry between the sender and the receiver. More relatedly, Loginova (2012) and Battigalli et al. (2013) study communication of private information with a guilt-averse sender. However, in both of these settings, the sender is always informed about the state of the world so that there is no scope for strategic evasion as a means to mitigate guilt.[3]

[3] A different type of belief-dependent preferences in the context of communication games - concern for being believed as telling the truth - has been modelled in Abeler et al. (2016), Dufwenberg and Dufwenberg (2016), Khalmetski and Sliwka (2017) and Gneezy et al. (2018)

	Good state	Bad state
Invest	P, F	c, F
Abstain	$0, 0$	$0, 0$

Figure 1: Payoff matrix of the game.

The problem of strategic evasion (pooling of informed and uninformed types) has been analyzed thus far mainly in verifiable disclosure settings (Dye 1985, Dziuda 2011), where the sender cannot misreport the observed information, but can only conceal it. Austen-Smith (1994) studies evasive communication in a mixed setting, where an informed sender can choose any message, while the uninformed sender cannot conceal the fact that he is uninformed from the receiver. In contrast to these studies, I show that, once the sender has belief-dependent preferences, a credible evasion can emerge even with completely unrestricted communication.

The remainder of this paper is organized as follows. Section 2 presents the model. Section 3 analyzes equilibria, including their welfare comparison. Section 4 considers the effects of policy interventions. Section 5 compares the results with those arising within a purely outcome-based model. Finally, Section 6 presents conclusions. All proofs are in the Appendix.

2 The model

2.1 Baseline setting

I consider a game between two players, the sender (he) and the receiver (she). There are two possible states of the world $\sigma \in \{G, B\}$ ("good" and "bad," respectively), each occurring with prior probability 0.5.[4] The timing of the game is as follows. In stage 1, nature chooses the state of the world, which is privately observed by the sender with probability $\kappa \in (0, 1)$. That is, there are three possible states of sender information $i^s \in \{G', B', N'\}$ (later termed information states), where G' corresponds to observation of G, B' to observation of B, and N' to no information. In stage 2 the sender sends a message m to the receiver about the state of the world out of a sufficiently large message space M (for now, I impose no structure on the message space, i.e., the exogenous formulation of messages is completely irrelevant). In the final stage, the receiver takes a binary action $x \in \{I, A\}$ ("invest" or "abstain," respectively) and the payoffs are realized. The payoff matrix is given in Figure 1.

Regarding the payoffs, I assume:

1) $F > 0$: the sender prefers investment independently of the state;

2) $P > 0$, $c < 0$: the receiver prefers to invest only in the good state.

Thus, there is monetary conflict of interest between the sender and the receiver in the bad state of the world. In terms of applications, a financial advisor can be, for instance, monetarily biased toward recommending investment in a specific financial product, which

[4] This restriction is made for ease of exposition and does not affect the generality of the results.

allows him to receive a higher commission (independently of whether this product fits the receiver's needs). In a similar way, a doctor can be incentivized by a pharmaceutical company to prescribe its products to patients.

I also make the following assumption about the payoffs.

Assumption 1 *Investment is ex ante profitable for the receiver, i.e.,* $P > -c$.

This restriction is necessary to generate evasive communication in equilibrium (considered in Subsection 3.3). Otherwise, evasion, i.e., pretending to be uninformed, cannot induce investment.[5]

The assumption $0 < \kappa < 1$ is also crucial for our setting (to ensure the credibility of evasive communication), and reflects the fact that the sender might sometimes fail to adequately address the receiver's investment problem (while being aware of this fact). For example, a doctor might not always be able to detect the true cause of a patient's symptoms (and hence, to recommend the right medical treatment), due to the complexity of the patient's case, lack of specific experience or competence, or noisy information from diagnostic tests.

Let us consider the utilities of the players. Denote by $U^r(x)$ the ex post utility of the receiver from action x. Her expected utility conditional on observing message m and investment is:[6,7]

$$E^r[U^r(I)|m] = \eta(m)P + (1-\eta(m))c, \tag{1}$$

where $\eta(m) \equiv \Pr^r[G|m]$, which is called below the persuasiveness of the message (in the sense of how persuasive is the message in inducing investment). The receiver's utility from abstaining, $U^r(A)$, is always 0 (see Figure 1).

Regarding the sender's utility, the core assumption is that the sender is guilt-averse, i.e., he dislikes to be responsible for letting down the receiver (Battigalli and Dufwenberg 2007).[8] In the current model setting, this arises if the sender's message induces overly high expectations relative to the eventually realized outcome. Specifically, the receiver is let down whenever her expected utility (conditional on the sender's message) is higher than her ex post payoff:

$$D^r(m,x) = \max[0, E^r[U^r(x)|m] - U^r(x)], \tag{2}$$

where $D^r(m)$ is the magnitude of letting down. In turn, the expected guilt of the sender in information state i^s is

$$G^s_{i^s}(\theta, m, x) = \theta E[D^r(m,x)|i^s], \tag{3}$$

where θ is the sender's sensitivity toward guilt. Finally, the total expected utility of the

[5] In this case, the only (robust) equilibrium is the lying equilibrium, considered in Subsection 3.2.

[6] Although formally this specification corresponds to risk neutrality, risk aversion does not qualitatively change any of the subsequent results as far as investment is still ex ante profitable.

[7] Both here and below, the upper index r refers to the receiver and s to the sender.

[8] This concept originates from psychological game theory, which presumes that utility can depend on beliefs *per se* (Geanakoplos et al. 1989, Battigalli and Dufwenberg 2009).

sender in information state i^s, denoted by $U^s_{i^s}(\theta, m, x)$, is assumed to be additive in the monetary and guilt components:

$$U^s_{i^s}(\theta, m, x) = F \cdot 1_I - G^s_{i^s}(\theta, m, x), \tag{4}$$

where 1_I is an indicator function, equal to 1 if the receiver invests and zero otherwise.

Note that $D^r(m, x) = 0$ if the good state is realized ex post, since the receiver cannot be let down by the highest possible outcome (i.e., $E^r[U^r(x)|m] \leq P$). Consequently, the sender expects non-zero guilt given investment only if the bad state of the world is realized, which implies

$$U^s_{i^s}(\theta, m, I) = F - \theta\lambda_{i^s}(E^s E^r[U^r(I)|m] - c), \tag{5}$$

where λ_{i^s} is the probability of the bad state expected by the sender in state i^s. If the receiver does not invest conditional on the message, then her outcome is no longer stochastic (being zero in all states), so that

$$E^r(U^r(A)|m) = U^r(A) = 0, \tag{6}$$

implying

$$D^r(m, A) = G^s_{i^s}(\theta, m, A) = U^s_{i^s}(\theta, m, A) = 0. \tag{7}$$

Hence, the sender never expects guilt if the receiver abstains, independent of the sent message.[9]

The sender's guilt aversion coefficient θ is a random variable, unknown to the receiver, distributed uniformly on the interval $(0, \bar{\theta}]$. This assumption serves to reflect the uncertainty of the receiver about the trustworthiness of the sender, which is widely heterogeneous in the population, as documented by many experimental studies (e.g., Charness and Dufwenberg 2006). Hence, the sender is characterized by both the information which he has observed i^s and his sensitivity to guilt θ. In what follows, I refer to θ as the sender's "type".

Note that lying costs defined in the current model may relate not only to psychological guilt. The term $D^r(m, x)$, more generally, is supposed to reflect the receiver's dissatisfaction with advice, separately from her dissatisfaction with the payoff itself - e.g., $D^r(m, x) = 0$ if the receiver expects to get a zero payoff after advice and eventually gets it. In turn, the receiver's dissatisfaction with advice can naturally lead to other costly consequences for the sender besides psychological costs, for example, reputational losses. Naturally, our definition still leaves aside some other aspects of lying costs, like aversion to induce a wrong decision of the receiver with own advice (due to, for instance, altrusitic concerns for the receiver). In particular, by (7) the sender bears no costs if the receiver

[9] One could argue that the receiver can still be dissatisfied with advice in this case, if she finds out that she has lost profitable investment opportunities by abstaining in the good state. At the same time, in many settings the state of the world is not observable for the receiver *per se* (but only through the ex post payoff). For instance, the receiver might never realize whether some innovative product fits her preferences unless she really tries the product. In such cases, the property in (7) is empiricially plausible.

abstains, even if the sender knows that it is a suboptimal decision. At the same time, it is plausible to assume that in many settings the advisor is disciplined by the receiver's *perception* of the quality of advice, rather that its actual quality.[10] These are the settings which fit to our model and which are thus at the focus of the current analysis.

2.2 Equilibrium concept

The equilibrium outcome is characterized by

1. the strategy of the receiver specifying whether to invest or abstain conditional on each possible message;

2. the strategy of the sender specifying which message to send conditional on the information state and the sensitivity to guilt;

3. the receiver's belief about the state of the world conditional on each message $\eta(m)$;[11]

4. all higher-order beliefs about the state of the world conditional on each message.

I apply the solution concept of pure strategy perfect Bayesian equilibrium, which implies that the sender's and receiver's equilibrium strategies should maximize the respective expected utility functions given equilibrium beliefs; the receiver's first-order beliefs are derived by Bayes rule whenever possible; higher-order beliefs are correct (Battigalli and Dufwenberg 2009).

Let us specify the optimal receiver strategy. She prefers to invest if and only if the expected utility from investment is larger than the utility from abstaining, i.e.,[12]

$$\begin{aligned} E^r[U^r(I)|m] &\geq E^r[U^r(A)|m] \\ &\Leftrightarrow \eta(m) \geq \frac{-c}{P-c} \equiv \underline{\eta}. \end{aligned} \tag{8}$$

The sender chooses the message, which maximizes his utility. Given (5) and the receiver's strategy specified above, the sender utility is

$$U^s_{i^s}(\theta, m, x) = \begin{cases} 0 & \text{if } \eta(m) < \underline{\eta}, \\ F - \theta\lambda_{i^s}(E^s E^r[U^r(I)|m] - c) & \\ \quad = F - \theta\lambda_{i^s}\eta(m)(P-c) & \text{if } \eta(m) \geq \underline{\eta}, \end{cases} \tag{9}$$

where the equality on the RHS follows from the consistency of the sender's second-order beliefs in equilibrium (i.e., $E^s E^r[U^r(I)|m] = E^r[U^r(I)|m] = \eta(m)(P-c)+c$). Thus,

[10] See Abeler et al. (2016) and Gneezy et al. (2018) for empirical evidence that repuational concerns play one of the central roles in communication settings.

[11] As becomes clear later in this subsection, the receiver's beliefs about the state of the world are sufficient to determine both optimal sender and receiver strategy, hence we do not need to additionally specify the receiver's beliefs about the sender's type θ and about his information state i^s.

[12] Hence, I assume that the receiver prefers investment over abstaining conditional on equal utility, which precludes mixed strategies. However, analysis of equilibria with mixed strategies of the receiver does not produce qualitatively different results.

the sender faces a tradeoff between inducing investment by being sufficiently persuasive with his message (to ensure $\eta(m) \geq \underline{\eta}$), and at the same time keeping the receiver's expectations low to mitigate guilt ($\eta(m)$ enters negatively in the sender's utility function once the receiver invests). Below, given (9), I denote the sender's utility function as $U^s_{i^s}(\theta, \eta(m), x)$.

The receiver's equilibrium beliefs are determined by the sender's messaging strategy through Bayes rule.

Lemma 1 *The persuasiveness of any equilibrium message m is*

$$\eta(m) \equiv \Pr[G|m] = \frac{\Pr[m|G']\kappa + \Pr[m|N'](1-\kappa)}{(\Pr[m|G'] + \Pr[m|B'])\kappa + 2\Pr[m|N'](1-\kappa)}.$$

Note that $\Pr[m|i^s]$ is determined by the sender's strategy in information state i^s, i.e., by the fraction of types who send the message conditional on this information state, while κ denotes the prior probability that the sender is informed.

I also impose the following restriction on out-of-equilibrium beliefs:

Assumption 2 *There always exists at least one out-of-equilibrium message m, with $\eta(m) < \underline{\eta}$.*

This assumption can be interpreted as if there always exists an opportunity for the sender to convince the receiver not to invest. Given that the sender is monetarily biased in the opposite direction, this assumption is intuitively reasonable. It rules out equilibria where the receiver invests independently of what the sender says (including out-of-equilibrium messages). Such equilibria are economically irrelevant, since there is no intuitive reason why the receiver would then refer to the sender in the first place.

3 Equilibrium analysis

3.1 Preliminaries

First, let us observe that in any possible equilibrium there exists a message leading to investment.

Lemma 2 *There is no equilibrium where all messages lead to abstaining.*

The intuition is that at least one equilibrium message must induce beliefs (i.e., the probability of the good state conditional on the message) not lower than 0.5 — otherwise, there is a contradiction to the prior of 0.5. Then, the receiver should invest after this message by Assumption 1.

Next, those types who indeed observe the good state always induce investment:

Lemma 3 *In any equilibrium, if $i^s = G'$, then all sender types induce investment.*

Indeed, if the sender observes the good state, his anticipated guilt from inducing investment is zero, since he knows that the receiver is not going to be let down. Hence, his expected utility from any message leading to investment is F, which is larger than the zero utility from inducing abstaining. Finally, in any equilibrium, there is a possibility for the sender to send an investment-inducing message by Lemma 2.

In contrast to this case, whenever the sender does not observe the good state with certainty (i.e., $i^s \neq G'$), his expected probability of the bad state, and hence the expected guilt from inducing investment, is strictly positive. One can show that the message strategy of such types has a cutoff structure in any equilibrium.

Lemma 4 *In any equilibrium, for each $i^s \in \{B', N'\}$ there exists a single cutoff type $\widehat{\theta}_{i^s} \in (0, \bar{\theta}]$ such that all sender types with $\theta \leq \widehat{\theta}_{i^s}$ send a message leading to investment and all types with $\theta > \widehat{\theta}_{i^s}$ (if any) send a message leading to abstaining.*

The rationale behind this result is the following. First, there are always types in any information state who are sufficiently insensitive to guilt to prefer inducing investment (which they can do by Lemma 2). Second, if some type prefers to induce investment over getting zero from abstaining, then all less guilt-sensitive types would also prefer at least the same message over zero, hence, also inducing investment. Analogously, once some type prefers to induce abstaining over any possible investment-inducing message, all higher types would prefer to induce abstaining as well. This corroborates the cutoff structure described in Lemma 4.

Another important distinction between different information states (besides the cutoff structure) is that types in state G' have different preferences over receiver beliefs relative to types in other states: while the former are indifferent to the persuasiveness of the message (they do not expect to let down the receiver anyway), sender types in states B' and N' always strictly prefer a less persuasive message. In the next subsections, it is shown that such asymmetry gives rise to the possibility of separation between types in state G' on the one side, and types in states B' and N' on the other. As a result, all equilibria can be classified by the degree of this separation (complete separation (or "evasion" equilibrium), complete pooling (or "lying" equilibrium) and partial separation (or "hybrid" equilibria)). I begin the equilibrium characterization with two extreme cases (either zero or complete separation), and discuss later in Subsection 3.4 why only these two equilibria are robust to slight perturbations in sender preferences.

3.2 Lying equilibrium

3.2.1 Characterization

The first type of equilibrium is termed the lying equilibrium. In this equilibrium, the sender uses either the least or the most persuasive message (depending on his guilt sensitivity and information state), so that lying, whenever it occurs, has the largest extent. In terms of equilibrium structure, a distinctive feature of this equilibrium is that there is

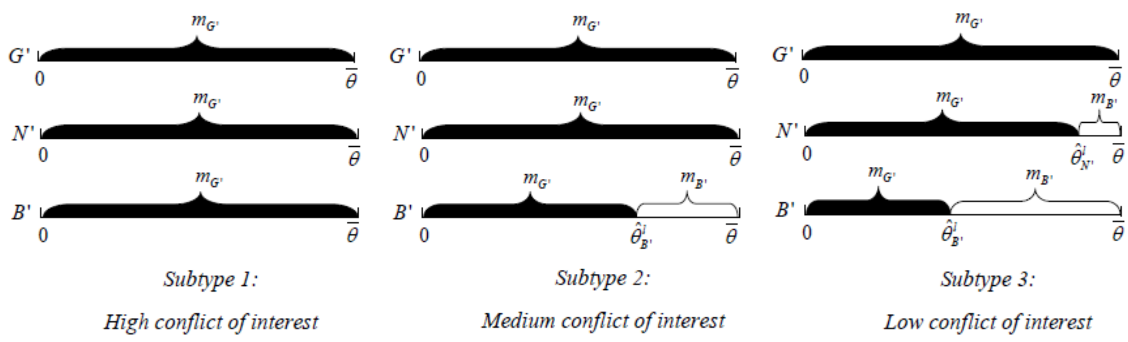

Figure 2: Subtypes of the lying equilibrium.

a complete pooling of investment-inducing types in states B' and N' with types in state G'.

Definition 1 *The lying equilibrium is defined as an equilibrium in which*

- *The message $m_{G'}$ is sent if $\theta \in (0, \bar{\theta}]$ in state G', $\theta \in (0, \widehat{\theta}^l_{B'}]$ in state B', and $\theta \in (0, \widehat{\theta}^l_{N'}]$ in state N';*
- *All other types in states B' and N' (if any) send the message $m_{B'}$;*
- *The receiver invests after $m_{G'}$, but abstains after $m_{B'}$ (if used);*
- *The beliefs after $m_{B'}$ (if used) and $m_{G'}$ are determined by Bayes rule, while for any out-of-equilibrium message $\widetilde{m}$ it holds that $\eta(\widetilde{m}) \in [0, \underline{\eta}) \cup [\eta(m_{G'}), 1]$. The receiver invests after an out-of-equilibrium message if and only if $\eta(\widetilde{m}) \geq \underline{\eta}$.*

Proposition 1 *There exists a unique lying equilibrium if either of the following holds:*

1. *$\kappa \leq \frac{P+c}{P}$ and $F > \widetilde{F}^l(\kappa)$, where $\widetilde{F}^l(\kappa) > 0$ is some threshold value;*
2. *$\kappa > \frac{P+c}{P}$.*

Otherwise, the lying equilibrium does not exist.

Figure 2 shows possible subtypes of this equilibrium. Here, each horizontal line represents the set of sender types for a given information state. The black bracket indicates types who use message $m_{G'}$, while the white bracket indicates types who use message $m_{B'}$. The figure shows three possible subtypes of this equilibrium depending on whether $\widehat{\theta}^l_{N'}$ and $\widehat{\theta}^l_{B'}$ are equal to $\bar{\theta}$.

The basic mechanism behind this equilibrium is described as follows: In the first subtype of the equilibrium, the monetary sender's incentive F (i.e., the degree of his conflict of interest) is high enough such that all sender types in all states want to pool on the message $m_{G'}$, which induces investment. If the monetary incentive decreases (Subtypes 2 and 3), then the most guilt-sensitive types in states N' and B' prefer to deviate to the message $m_{B'}$, which yields abstaining and hence zero guilt (see (7)). Clearly, the fraction of such types is larger in state B', where the expected guilt is higher than in state N' for a given sensitivity θ. Finally, no type has an incentive to deviate to out-of-equilibrium messages, which either lead to abstaining (if $\eta(\widetilde{m}) < \underline{\eta}$), or to higher guilt (if $\eta(\widetilde{m}) \geq \eta(m_{G'})$).

The equilibrium beliefs conditional on the messages $m_{G'}$ and $m_{B'}$ are determined by the cutoff types $\widehat{\theta}^l_{B'}$ and $\widehat{\theta}^l_{N'}$ (in particular, by substituting ($\Pr[m_{G'}|G'] = 1$, $\Pr[m_{G'}|N'] = \frac{\widehat{\theta}^l_{N'}}{\overline{\theta}}$ and $\Pr[m_{G'}|B'] = \frac{\widehat{\theta}^l_{B'}}{\overline{\theta}}$ into (15)). One can show that lower cutoff types correspond to a higher persuasiveness of the message $m_{G'}$ (i.e., a higher probability of the good state of the world conditional on the message). As Proposition 1 implies, the cutoffs types, which lead to the persuasiveness of $m_{G'}$ making these types indifferent between $m_{G'}$ and $m_{N'}$ (i.e., ensuring the equilibrium), are always unique for given parameter values.

To complete the equilibrium characterization, let us consider the receiver's incentives. In equilibrium, two receiver's incentive constraints must be satisfied: $\eta(m_{G'}) \geq \underline{\eta}$ (so that the receiver prefers to invest after $m_{G'}$, see (8)), while $\eta(m_{B'}) < \underline{\eta}$ (so that she prefers to abstain after $m_{B'}$). The first requirement is always satisfied, due to the fact that all types in state G' pool on $m_{G'}$, which ensures $\eta(m_{G'}) \geq 0.5$ (while $\underline{\eta} < 0.5$ by Assumption 1). Yet, the second incentive constraint $\eta(m_{B'}) < \underline{\eta}$ can be violated if the share of uninformed types sending $m_{B'}$ becomes sufficiently high, which raises the persuasiveness of $m_{B'}$ above $\underline{\eta}$. This occurs when the unconditional share of uninformed types is sufficiently high (i.e., $\kappa \leq \frac{P+c}{P}$), while the monetary incentive, and hence the cutoff in state N', is sufficiently low (i.e., $F \leq \widetilde{F}^l(\kappa)$, see Proposition 1). In this case, the lying equilibrium does not exist.

3.2.2 Discussion

The main feature of the lying equilibrium is that the restrictions on out-of-equilibrium beliefs imply that only one message $m_{G'}$ can be used to induce investment in equilibrium. Since this message is sent by all types in state G', it gets relatively persuasive, thus leading to relatively high psychological costs if it is used for deception. As it is shown later in Subsections 3.5 and 4.1, this can work as a disciplining device for the sender, hence making the lying equilibrium relatively attractive for the receiver.

Let us consider the intuitive interpretation of the messages used in the lying equilibrium (since I impose no structure on the message space, the meaning of each equilibrium message arises endogenously). The message $m_{G'}$ is always used by types who indeed observe the best state of the world and, thus, have no incentive to downgrade information or to pool with types in other information states. Hence, this can be interpreted as an explicit claim that the observed state is good with certainty (in terms of exogenous formulation of the message). Thus, the pooling of types in other information states with this message can be interpreted as explicit lying. On the other hand, the message $m_{B'}$ is used by the most guilt-sensitive types in information state B', who would like to induce the receiver to abstain from investment. This message can thus be interpreted as an explicit claim that the observed state was bad with certainty. Out-of-equilibrium messages $\widetilde{m}$ can be interpreted as being either insufficiently persuasive to induce investment (if $\eta(\widetilde{m}) \in [0, \underline{\eta})$) or overly explicit (if $\eta(\widetilde{m}) \in [\eta(m_{G'}), 1]$).

Intuitively, the former group of messages can be thought to include "evasive" messages, i.e., claims that the sender has not observed information, since this represents one of the actually possible information states. The fact that the receiver does not find these messages sufficiently persuasive can be understood as that she believes that such messages are rather used by types who have observed the bad state of the world and want to hide the inconvenient truth. This reasoning is supported by the fact that types in state B' have greater disutility from guilt than types in state N' (conditional on the same sensitivity to guilt); consequently, they indeed have a greater incentive to send less explicit messages like evasive ones.

Note, finally, that in this equilibrium there are two types of deception that lead to a loss to the receiver from the ex ante perspective. The first is sending the message $m_{G'}$ in state B' (deception driven by the monetary bias in sender incentives, termed as bias-driven deception). The second type of deception is sending the message $m_{B'}$ in state N' (or guilt-driven deception), i.e., inducing the receiver to abstain from investment by pretending to observe the bad state of the world, while in fact not having observed any state. As discussed in Subsection 2.1, the sender does not feel guilt in this case, as he avoids any risk of letting down the receiver. At the same time, since investment is ex ante profitable by Assumption 1, the receiver would strictly prefer to invest had she known that the sender is actually uninformed. Such a situation can be interpreted as an inefficient reluctance of the sender to recommend products that are risky though profitable from an ex ante perspective. In terms of the medical example, a doctor who is too afraid of appearing incompetent (or being prosecuted for bad treatment) might advise his patient to undertake only the most conservative traditional treatments with predictable but low efficiency, instead of trying out more innovative (and hence, more risky), but more promising treatment methods. Analogously, a financial advisor might be reluctant to recommend reasonably risky but profitable financial products. In such situations, promoting additional monetary incentives for the sender may mitigate these adverse effects; this is considered later in Subsection 4.2.

3.3 Evasion equilibrium

3.3.1 Characterization

The second type of equilibrium in this game is the evasion equilibrium, where, besides the explicit messages $m_{G'}$ and $m_{B'}$ as in the lying equilibrium, an additional evasive message $m_{N'}$ is used. This equilibrium implies complete separation of types in state G'.

Definition 2 *Evasion equilibrium is defined as an equilibrium in which*

- *The message $m_{G'}$ is sent by all types in state G';*
- *The message $m_{N'}$ is sent if $\theta \in (0, \widehat{\theta}^e_{B'}]$ in state B', and $\theta \in (0, \widehat{\theta}^e_{N'}]$ in state N';*

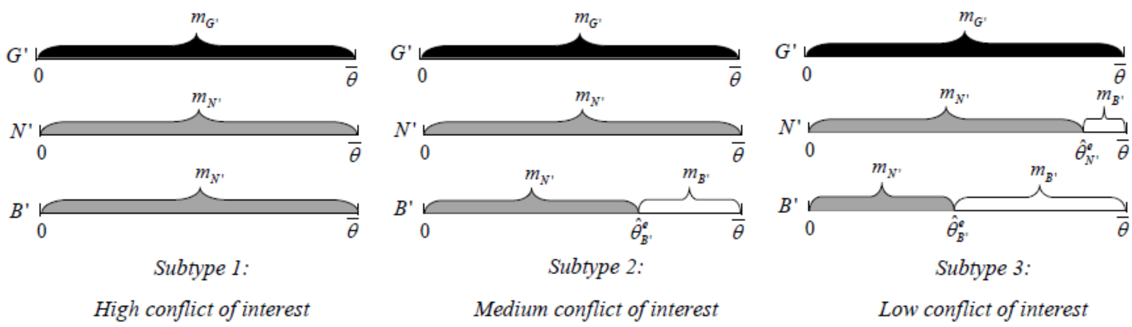

Figure 3: Subtypes of the evasion equilibrium.

- *All other types in states B' and N' (if any) send the message $m_{B'}$;*
- *The receiver invests after $m_{G'}$ and $m_{N'}$, but abstains after $m_{B'}$ (if used);*
- *The beliefs after $m_{G'}$, $m_{N'}$ and $m_{B'}$ (if used) are determined by Bayes rule, while for any out-of-equilibrium message $\widetilde{m}$ it holds that $\eta(\widetilde{m}) \in [0, \underline{\eta}) \cup [\eta(m_{N'}), 1]$. The receiver invests after an out-of-equilibrium message if and only if $\eta(\widetilde{m}) \geq \underline{\eta}$.*

Proposition 2 *There exists a unique evasion equilibrium if either of the following holds:*

1. *$\kappa \in \left(0, \frac{P+c}{P}\right]$ and $F > \widetilde{F}^e(\kappa)$, where $\widetilde{F}^e(\kappa) > 0$ is some threshold value;*
2. *$\kappa \in \left(\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right]$ and $F \leq \bar{\theta}\,(P + c)\,\frac{(1-\kappa)}{\kappa}$.*

Otherwise, the evasion equilibrium does not exist.

The scheme of this equilibrium is given in Figure 3. Besides types sending the highest message $m_{G'}$ (black figure bracket) and types sending the lowest message $m_{B'}$ (white figure bracket) as in the lying equilibrium, there is a set of types using the message $m_{N'}$ (grey figure bracket), which, as discussed below in Subsection 3.3.2, can be interpreted as an evasive message (i.e., a claim of being uninformed).

The basic mechanism is as follows: Sender types in state G', facing no guilt, have the same utility from both $m_{G'}$ and $m_{N'}$, which is equal to F insofar as the receiver invests after both messages. Hence, they do not have a strict incentive to deviate to $m_{N'}$. At the same time, types in states N' and B', who induce investment, face a strictly positive expected guilt, and hence strictly prefer the evasive message $m_{N'}$ since $\eta(m_{N'}) < \eta(m_{G'})$ (while the monetary payoff is the same). Hence, the evasive message provides a way to mitigate guilt by inducing less optimistic payoff expectations on the part of the receiver, while still keeping her investing after receiving advice. Finally, no type has an incentive to deviate to any out-of-equilibrium message $\widetilde{m}$, which yields either abstaining and a utility of 0 (if $\eta(\widetilde{m}) \in [0, \underline{\eta})$) or the same monetary outcome with a higher guilt (if $\eta(\widetilde{m}) \in [\eta(m_{N'}), 1]$).

The receiver's incentive constraints are $\eta(m_{G'}) \geq \underline{\eta}$ (investment after $m_{G'}$), $\eta(m_{N'}) \geq \underline{\eta}$ (investment after $m_{N'}$) and $\eta(m_{B'}) < \underline{\eta}$ (abstaining after $m_{B'}$). The first constraint is trivially satisfied. The second constraint (meaning that the evasive message is sufficiently credible) is satisfied whenever the share of truly uninformed types is sufficiently high (i.e., κ is sufficiently low):

$$\kappa \leq \frac{2(P+c)}{2P+c}. \tag{10}$$

In addition, in Subtype 2 of the equilibrium, the only interior cutoff $\widehat{\theta}^{e}_{B'}$ should be sufficiently distant from the boundary $\bar{\theta}$ (so that there is no excessive pooling of types in state B' pretending to be uniformed), which places additional restriction on the monetary bias F (see case 2 of Proposition 2). Finally, the constraint $\eta(m_{B'}) < \underline{\eta}$ requires that, under sufficiently low values of κ (i.e., a high prior share of uninformed types), F should not be too low. In this case, sufficiently many uninformed types in state N' prefer to induce abstaining with $m_{B'}$ (since they are not sufficiently incentivized to induce investment), which eventually increases the persuasiveness of $m_{B'}$ above $\underline{\eta}$.

An important structural distinction of the evasion equilibrium relative to the lying equilibrium is that the cutoffs in the evasion equilibrium are higher.

Lemma 5 *Whenever both lying and evasion equilibria exist, for any $i^s \in \{B', N'\}$ it holds that $\widehat{\theta}^{e}_{i^s} \geq \widehat{\theta}^{l}_{i^s}$ with a strict inequality whenever $\widehat{\theta}^{l}_{i^s} < \bar{\theta}$.*

The reason for this is that the evasive message is strictly less persuasive, so that a larger fraction of sender types in a given state prefer to induce investment with this message. This property has important implications for receiver welfare, considered later in Subsection 3.5.

3.3.2 Discussion

The main feature of the evasion equilibrium is that deception of types in state B' takes the form of downgrading the quality of obtained information (i.e., pooling with actually uninformed types), rather than misrepresentation of information (i.e., pooling with types in state G'). The emergence of this form of deception is based on the sender's incentives to reduce the receiver's ex-ante expectations as much as possible, while still keeping her choosing the most preferred sender's action (investment). Evasion (mimicking the uninformed types) naturally fits these sender's incentives by being less persuasive than lying, yet sufficiently persuasive to induce investment (recall that the receiver is assumed to have sufficiently high prior beliefs by Assumption 1, so that investing in uncertainty is ex-ante efficient). Importantly, evasion relaxes the costs of deception for the sender (relative to the explicit lying in the lying equilibrium), and thus can lead to the contraction of truth-telling in state B' (the cutoff in this state increases), as asserted in Lemma 5.

In the evasion equilibrium, while messages $m_{G'}$ and $m_{B'}$ can be interpreted in the same way as in the lying equilibrium (claims to have observed with certainty the good and the bad states, respectively), the evasive message $m_{N'}$ can be interpreted as a claim to have not observed information about the state of the world (in terms of the exogenous formulation of the message). Indeed, this message is primarily sent by types who actually have not observed information (state N') and who have no incentive to hide this from the receiver (since the receiver prefers to invest in uncertainty). At the same time, types $(0, \widehat{\theta}^{e}_{B'}]$ in state B' have a strict incentive to pool with uninformed types to hide the inconvenient truth. Hence, their strategy can be interpreted as (literally) claiming to be uninformed.

Note that in the evasion equilibrium types in state G' are indifferent between messages $m_{G'}$ and $m_{N'}$: both of these messages induce investment, while the expected guilt is zero in either case since in state G' the sender knows that the receiver is never going to be let down after investment. One can make the model more robust to possible varioations in underlying assumptions by introducing a small fixed lying cost in the sense of Kartik (2009) in the sender's preferences so that the latter would *strictly* prefer message $m_{G'}$ over $m_{N'}$ all else equal (which is also empirically plausible). In this case, the evasion equilibrium would be retained even if the sender would be assumed to be not perfectly certain about the state of the world in either information state: a sufficiently small risk of letting down the receiver would then be still insufficent to make the sender prefer $m_{N'}$ over $m_{G'}$ in state G'. At the same time, no other qualitative results would be affected if such fixed lying cost is added.

3.4 Other equilibria and equilibrium selection

Besides the two considered types of equilibria, where there is either complete pooling or complete separation relative to state G', there exist a continuum of "hybrid" equilibria with partial pooling (see Figure 4) .

Definition 3 *Hybrid equilibrium is defined as an equilibrium in which*

- *The message $m_{G'}$ is sent by a fraction of types $\gamma \in [0,1]$ in state G';*
- *The message $m_{N'}$ is sent by a fraction of types $1-\gamma$ in state G', and types with $\theta \in (0, \widehat{\theta}^{h}_{B'}]$ in state B' and $\theta \in (0, \widehat{\theta}^{h}_{N'}]$ in state N';*
- *All other types in states B' and N' (if any) send the message $m_{B'}$;*
- *The receiver invests after $m_{G'}$ and $m_{N'}$, but abstains after $m_{B'}$ (if used);*
- *The beliefs after $m_{G'}$, $m_{N'}$ and $m_{B'}$ (if used) are determined by Bayes rule, while for any out-of-equilibrium message $\widetilde{m}$ it holds that $\eta(\widetilde{m}) \in [0, \underline{\eta}) \cup [\eta(m_{N'}), 1]$. The receiver invests after an out-of-equilibrium message if and only if $\eta(\widetilde{m}) \geq \underline{\eta}$.*

Thus, the lying and evasion equilibria are the special cases of the hybrid equilibrium for $\gamma = 1$ and $\gamma = 0$, respectively. Moreover, one can show that all possible cutoffs of existing hybrid equilibria range between the cutoff in the lying equilibrium and the cutoff in the evasion equilibrium.

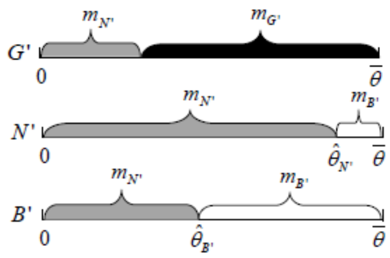

Figure 4: The hybrid equilibrium.

Proposition 3 *If both lying and evasion equilibria exist, then for any $z \in [\widehat{\theta}^l_{i^{s'}}, \widehat{\theta}^e_{i^s}]$ with $i^s \in \{B', N'\}$ there exists γ such that the hybrid equilibrium exists with $\widehat{\theta}^h_{i^s} = z$. There exists no hybrid equilibrium with $\widehat{\theta}^h_{i^s} \notin [\widehat{\theta}^l_{i^{s'}}, \widehat{\theta}^e_{i^s}]$.*

The mechanism here is that for given γ, there exist unique cutoffs in each state supporting the equilibrium. At the same time, since a higher fraction of types in state G' pooling on the message $m_{N'}$ (i.e., higher γ) increases its persuasiveness and hence the associated guilt (all else equal), this also pushes the equilibrium cutoffs down. Since the lying equilibrium is the limit of the hybrid equilibrium if $\gamma \to 1$ (complete pooling), while the evasion equilibrium is the limit if $\gamma \to 0$ (complete separation), all possible cutoffs of hybrid equilibria lie between these two cases.

Although generally there also exist equilibria in this game with a different messaging structure than in the hybrid equilibria (e.g., there can be multiple investment-inducing messages in states N' and B'), the following proposition provides the reasoning for why it is sufficient to consider only these equilibria (including the lying and evasion equilibria as special cases).

Proposition 4 *All existing equilibria are outcome-equivalent to the hybrid equilibria.*

Inter alia, this (together with Proposition 3) implies that in any equilibrium, the sender cannot have higher expected guilt than in the lying equilibrium, or lower expected guilt than in the evasion equilibrium. *Thus, the lying equilibrium and the evasion equilibrium represent two limit cases of the whole continuum of possible equilibria in this game.* Moreover, one can show that a simple and intuitive assumption on the sender's preferences immediately rules out all equilibria besides the lying and evasion equilibria.

Assumption 3 *Conditional on equal utility, the sender has a strict lexicographic preference over messages in each information state.*

Proposition 5 *Under Assumption 3, there exist no equilibria, which are not outcome equivalent to either lying or evasion equilibria.*

Assumption 3 implies that the sender strictly ranks any two messages even if they yield the same expected utility. This can be interpreted as if the exogenously given formulation of messages also has some value, though of infinitely small order. For example, the sender can have an infinitely small cost of lying in the sense of Kartik (2009), i.e., he can (on top of belief-dependent cost of lying) have disutility from a literal lie, which is determined by the deviation of the exogenously given formulation of the message from the truth. Assumption 3 eliminates multiplicity of messages used in state G', since the sender is then never indifferent between any two messages in this state. This, in turn, rules out all hybrid equilibria with $0 < \gamma < 1$, because they are based on the *partial* pooling of other information states with state G' (on one out of several messages used in G').

This result justifies the focus of the subsequent analysis on only two robust equilibria, the lying and the evasion equilibria, although the main qualitative results also remain under consideration of the hybrid equilibria.

3.5 Welfare comparison

One of the key questions that can be studied with this model is whether evasion is eventually detrimental from the point of view of the receiver's welfare. In fact, from the ex ante perspective, the receiver's utility can be both higher and lower in the evasion equilibrium than in the lying equilibrium, depending on the monetary conflict of interest.

Proposition 6 *Whenever both lying and evasion equilibria exist, there exists* $F^* \in \left[(1-\kappa)\,\bar{\theta}\frac{P-c}{4-3\kappa}, \min[\frac{\bar{\theta}(P+c)(1-\kappa)}{\kappa}, \bar{\theta}\frac{P-c}{4-\kappa}]\right]$ *such that the lying equilibrium yields higher ex ante utility for the receiver than the evasion equilibrium if* $F \geq F^*$*, and a lower utility if* $F < F^*$*.*

This result is based on the fact that the cutoff types in the evasion equilibrium are higher (see Lemma 5), since $m_{N'}$ in the evasion equilibrium is less persuasive, and hence less costly in terms of guilt than $m_{G'}$ in the lying equilibrium. This relation of the cutoffs implies that the rate of bias-driven deception (sending an investment-inducing message in state B') is higher in the evasion equilibrium. Such type of deception is clearly detrimental to the receiver's welfare. On the other hand, a lower cutoff in state N' in the lying equilibrium implies that the rate of guilt-driven deception (sending the message $m_{B'}$ in state N', see Figure 2) is higher there. Such kind of deception is also detrimental to the receiver's welfare, since she prefers investment over abstaining *ex-ante* (by Assumption 1). Thus, the sender's option to use the evasive message in equilibrium has two effects on the receiver's welfare. The negative effect stems from providing psychologically cheap opportunities for the sender to induce investment after observing the bad state of the world (by credibly pretending to be uninformed, which incurs less guilt than explicit lying). The positive effect stems from raising the efficiency of communication of uninformed types, whose expected guilt in case of inducing investment is reduced. The total effect depends on which of these two effects dominates.

In particular, if F is sufficiently large (above $\bar{\theta}(P-c)/(4-\kappa)$), then both lying and evasion equilibria are of either Subtype 1 or Subtype 2, where there is no guilt-driven deception (see Figures 2 and 3). Consequently, the total effect of evasion is limited to enhancing bias-driven deception in state B', leading to a welfare loss for the receiver (when both equilibria are of Subtype 1, welfare does not change). If, to the contrary, F is sufficiently small (below $(1-\kappa)\,\bar{\theta}(P-c)/(4-3\kappa)$), both lying and evasion equilibria are of Subtype 3. Then, besides the negative effect, there is an additional positive effect of evasion due to a reduction in guilt-driven deception. The clear-cut result here is that this positive effect in this case is always larger than the negative effect related to bias-driven deception.[13]

If F is between the aforementioned thresholds, then the overall effect of evasion depends on how large is the scope of guilt-driven deception in the lying equilibrium. If F is sufficiently close to the lower threshold, guilt-driven deception is large enough to induce the overall positive effect of evasion. On the other hand, if F is closer to the upper bound, then the increase in bias-driven deception becomes the dominant effect, making the receiver worse off in this equilibrium. As stated in Proposition 6 there is a unique threshold F^* separating the two cases.

4 Effects of policy intervention

4.1 Prohibition of lying

In this subsection, I consider possible effects of a policy that restricts lying. The results suggest that under some conditions, such a policy can be detrimental to the receiver's welfare.

4.1.1 Verifiability structure

First, one needs a definition of lying that reflects its legal sense. Normally, lying is understood as a misrepresentation of private information (Kartik 2009), i.e., when a stated meaning of the message deviates from the truth. As discussed in Subsections 3.2 and 3.3, the meaning of equilibrium messages in our model arises endogenously in both lying and evasion equilibria: $m_{G'}$ ($m_{B'}$) can be interpreted as a claim to have observed the good (bad) state of the world with certainty, while $m_{N'}$ can be interpreted as a claim not to have observed any information. To avoid considering out-of-equilibrium messages (which can potentially have arbitrarily many possible formulations), I further assume that the message space is restricted to $m_{G'}$, $m_{B'}$ and $m_{N'}$. Given these endogenous message meanings, it is reasonable to define lying in our setting as follows:

[13] The reason for this is as follows. First, note that a switch from the evasion to the lying equilibrium leads to an overall reduction of investment in states N' and B' (due to the decrease in the cutoffs). At the same time, the expected receiver's payoff *conditional* on obtaining an investment-inducing message in these information states remains the same in both equilibria. This is ensured by the fact that the ratio of the cutoffs in states N' and B' is the same (see Lemmas 11 and 14 in the Appendix). Finally, this conditional expected payoff is positive, because the receiver invests after $m_{N'}$ in the evasion equilibrium. Hence, the switch from the evasion to the lying equilibrium in this case effectively results (merely) in contraction of ex ante efficient investment, yielding a loss in welfare.

Definition 4 *Lying is sending a message m_{i^s} in an information state other than i^s, for $i^s \in \{G', N', B'\}$.*[14]

Note that according to this definition both $m_{N'}$ and $m_{G'}$ are considered as lies if reported in states other than N' and G', respectively, and hence both can be subject to regulation. Yet, intuitively, these two types of lies may differ in the degree of their verifiability for an external regulator. This is further formalized by the following assumptions:

Assumption 4 *The sender's message is (ex post) verifiable, while the sender's information state is not.*

Assumption 5 *The state of the world is (ex post) verifiable if and only if the receiver invests.*

Assumption 4 can be justified by the fact that the sender's message is clearly observable to the receiver, and hence can be fixed by some communication protocol. At the same time, in real-life applications, it is normally difficult to verify what information the sender actually possesses (especially, *whether* the sender has obtained certain information), as far as the sender obtains his information privately. Assumption 5 is justified if one thinks of the state of the world in terms of fitness of some advised product to the needs of a specific receiver, which can only be verified if the receiver actually tries the product, like in the case of medical treatment (see also footnote 9). That is, the assumption reflects the fact that it is much easier to make the sender liable for already realized losses, than for foregone potential profits (e.g., in the case of sending $m_{B'}$ in the good state of the world).[15]

Assumptions 4 and 5 lead to the following result.

Lemma 6 *Lying is not (ex post) verifiable in the evasion equilibrium. It is verifiable in the lying equilibrium if and only if $m_{G'}$ is sent and the bad state of the world is realized.*

Indeed, consider the case of sending the message $m_{N'}$ in the evasion equilibrium. Then, by Definition 4, lying can be verified only if one could prove that the sender was in an information state other than N'. At the same time, $i^s \neq N'$ can be verified neither directly (by Assumption 4) nor by the realized state of the world, since all states of the world are consistent with N'. In case the sender sends the lowest message $m_{B'}$, then in both lying and evasion equilibria, the receiver always abstains, so that the state of the world is not verifiable by Assumption 5. Finally, if $m_{G'}$ is sent and the good state of the world is realized, the only possible case of lying in this case — sending $m_{G'}$ while being uninformed in the lying equilibrium — cannot be distinguished from truth-telling either. The only possible case where lying is verifiable is when the message $m_{G'}$ is sent and the bad state of the world is realized (this is possible only in the lying equilibrium). Then, the information state G' can be excluded with certainty since $\Pr[G'|\sigma = B] = 0$.

[14] The messages $m_{G'}$, $m_{N'}$ and $m_{B'}$ themselves are implicitly defined within Definitions 1 and 2.

[15] Note that the main qualitative result of this subsection (that a lying prohibition policy reduces ex ante welfare in certain cases) also holds without Assumption 5.

4.1.2 Post-policy equilibrium characterization

Consider a policy intervention such that a sufficiently high fine is introduced for any verifiable lying. Lemma 6 then implies that incentives in the evasion equilibrium are not affected by the policy. Yet, if the fine is sufficiently high, then the lying equilibrium is no longer possible: no types in information states N' and B' would then prefer to send $m_{G'}$, which could lead to the fine (if B is realized) with a positive probability. Finally, the policy intervention renders a new possible equilibrium, termed the 'evasive babbling equilibrium.' This is defined as an equilibrium where all types in state B' pool with all types in state N' by sending the evasive message $m_{N'}$, which effectively induces the receiver to abstain conditional on this message.

Definition 5 *Evasive babbling equilibrium is defined as an equilibrium in which:*

- *The message $m_{G'}$ is sent by all types in state G';*
- *The message $m_{N'}$ is sent by all types in state B' and all types in state N';*
- *The receiver invests after $m_{G'}$, but abstains after $m_{N'}$;*
- *The beliefs after $m_{G'}$, and $m_{N'}$ are determined by Bayes rule. For the out-of-equilibrium message $m_{B'}$ it holds that $\eta(m_{B'}) \in [0, \underline{\eta})$ while the receiver abstains.*

Under a sufficiently high fine for lying, this indeed constitutes an equilibrium. As before, no types in state G' would like to deviate to a message that leads to abstaining. Besides, no types in states N' and B' would like to deviate to $m_{G'}$, while this would lead to a positive probability of being fined ex post. Deviation to the out-of-equilibrium message $m_{B'}$ does not make the sender better off either, as the receiver still abstains. Finally, the receiver invests after $m_{G'}$ and abstains after $m_{N'}$ insofar as the persuasiveness of $m_{N'}$ is sufficiently low, i.e., $\eta(m_{N'}) < \underline{\eta}$. One can show that this holds whenever $\kappa > \frac{P+c}{P}$ (i.e., the share of truly uninformed types is sufficiently low). The following proposition summarizes the equilibrium characterization under the policy intervention.

Proposition 7 *If lying is prohibited, then:*

- *There exists an evasion equilibrium under the same parameter restrictions as before;*
- *There exists an evasive babbling equilibrium if and only if $\kappa > \frac{P+c}{P}$;*
- *No other equilibria exist.*

4.1.3 Welfare implications

Let us now consider how the policy intervention (i.e., imposition of a sufficiently high fine for verifiable lying) changes the receiver welfare relative to the pre-policy status quo. Since the policy does not distort incentives in the evasion equilibrium, I assume that the pre-policy equilibrium is the lying equilibrium (otherwise the policy-maker would not have

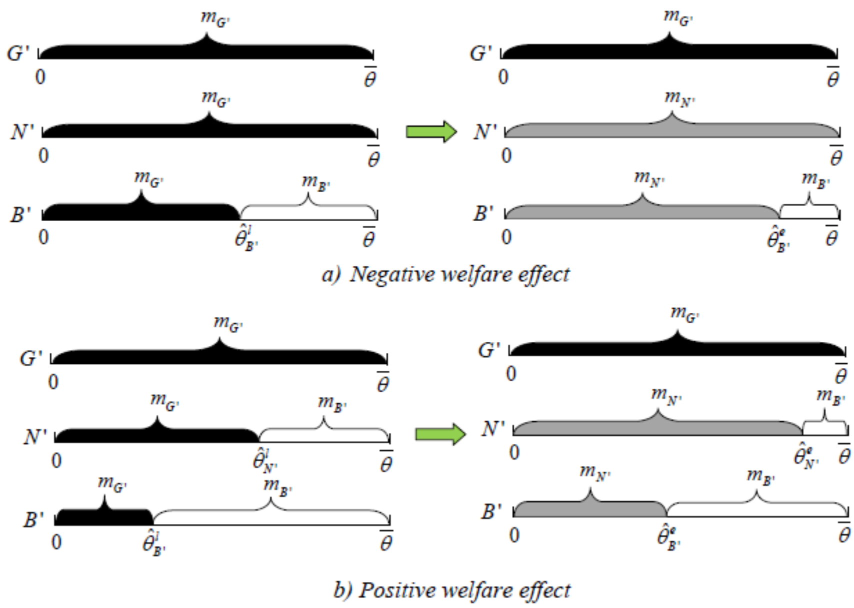

Figure 5: The effect of lying prohibition in the case of constructive evasion.

strict incentives to implement the policy). Besides, to make the results easier to represent, I assume that the evasive babbling equilibrium does not emerge if the (more informative) evasion equilibrium is possible instead.

The following proposition summarizes the findings regarding welfare implications of the policy intervention.

Proposition 8 *If the initial equilibrium is the lying equilibrium, then the prohibition of lying results in either of the following:*

- *Switch to the evasion equilibrium if* $\kappa \leq \frac{P+c}{P}$, *or* $\kappa \in \left(\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right]$ *and* $F \leq \bar{\theta}\,(P+c)\,\frac{(1-\kappa)}{\kappa}$ *(case of constructive evasion):*
  - *Decrease in the receiver's welfare if* $F \geq F^*$, *and increase otherwise.*
- *Switch to the evasive babbling equilibrium if* $\kappa \in \left(\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right]$ *and* $F > \bar{\theta}\,(P+c)\,\frac{(1-\kappa)}{\kappa}$, *or* $\kappa > \frac{2(P+c)}{2P+c}$ *(case of destructive evasion):*
  - *Decrease in the receiver's welfare if* $\kappa \leq \frac{2(P+c)}{2P+c}$ *and* $F \leq \frac{(P^2-c^2)(1-\kappa)\bar{\theta}}{(P(1-\kappa)-c)\kappa}$, *increase otherwise.*

Thus, the consequence of the policy is the destruction of the lying equilibrium, and the selection of either evasive or evasive babbling equilibrium, which are the only equilibria existing under the policy (by Proposition 7). If there is a switch from the lying to the evasion equilibrium (constructive evasion case), then all welfare implications given by Proposition 6 are in place. That is, if the monetary conflict of interest is high enough so that guilt-driven deception is negligible, the policy has a net negative effect due to

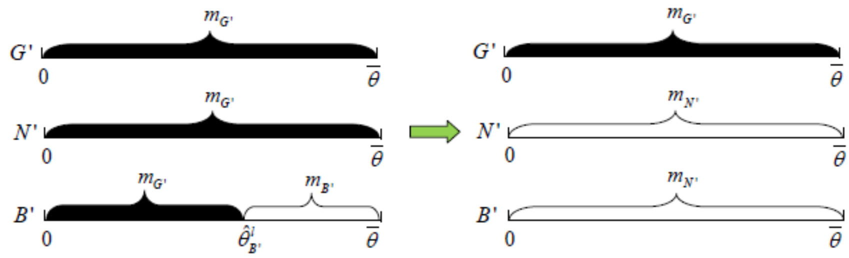

Figure 6: The effect of lying prohibition in the case of destructive evasion.

an increase in bias-driven deception (Figure 5, a). In other words, the pre-policy lying equilibrium disciplines the sender by polarizing the communication alternatives (being either highly persuasive $m_{G'}$ or highly pessimistic $m_{B'}$, so that inducing investment with $m_{G'}$ remains unattractive for a relatively large share of sender types). In contrast, the evasion equilibrium relaxes the cost of guilt for the sender by enabling a credible separation from the highly persuasive types in state G'.

At the same time, under a low conflict of interest (or, equivalently, relatively high guilt aversion), the policy has a positive effect due to reduction in guilt-driven deception (Figure 5, b; see Subsection 3.5). Note that lying prohibition policy is naturally motivated by high conflict of interest between experts and consumers on the market, so that the first case (characterized by decrease in welfare) may be considered more economically relevant.

The other scenario is when the policy induces a switch from the lying to the evasive babbling equilibrium (the case of destructive evasion, Figure 6). This effectively destroys all investment, which takes place in the lying equilibrium conditional on $i^s \neq G'$. Whether this has a net positive effect for the receiver depends on the quality of such investment, i.e., on the relative shares of types in states N' and B', inducing investment in the lying equilibrium. The relative share of types in state N' is sufficiently high if, first, the prior probability of being uninformed is relatively high ($\kappa \leq \frac{2(P+c)}{2P+c}$), and second, the conflict of interest is relatively low (so that sufficiently many types in state B' do not induce investment by sending $m_{B'}$). In this case, investment in the lying equilibrium conditional on $i^s \neq G'$ yields an ex ante positive payoff for the receiver. Abandoning such investment due to destruction of any persuasive communication in states N' and B', which happens in the evasive babbling equilibrium, leads to a net welfare loss for the receiver. The main intuition behind this result is that the lying prohibition policy may lead to excessive evasion of types in state B', resulting in destruction of any efficient communication from uncertain senders.

In summary, the policy aimed at lying prohibition can be welfare improving in several cases. First, if most of sender types are informed about the true state of the world, then lying prohibition basically eliminates deception of types who have observed the bad state (Scenario 2, positive case). Second, if there is too much of inefficient precaution of uninformed senders on the market (guilt-driven deception), then enhancing credible evasive communication as a result of the policy can serve as a way to reduce such precaution (Scenario 1, positive case). At the same time, if the precaution is not an issue and, besides, there is a sufficiently high share of uninformed senders on the market, then a prohibition of lying can lead to a spread of deception in the form of evasive communication (Scenario 1, negative case). Finally, pooling with types observing the good state of the world can be the only way to induce investment for uncertain senders. Eliminating this possibility due to lying prohibition can also lead to a net welfare loss (Scenario 2, negative case).

Thus, although the receiver does not like lying per se, she might prefer to have lying on the equilibrium path, as it then serves as a disciplining device for guilt-averse senders. In this sense, lying, as a population phenomenon, can be interpreted to have a "hidden value," which can be destroyed by an overly interventionist policy.

4.2 Regulation of commission payments

Let us consider the case where the regulator mitigates the conflict of interest between the sender and the receiver by capping the sender's contingent payment F. One can show that such policy can backfire by making the sender's guilt motivation too prominent, in particular, for uninformed sender types. This can lead to overly precautionary advice from uninformed types (guilt-driven deception), and hence foregone investment opportunities. At the same time, the policy also works positively by inducing more truth-telling among types observing the bad state of the world, i.e., by reducing bias-driven deception. The following proposition summarizes the total effect.

Proposition 9 *Welfare strictly increases with F if $\kappa < \frac{2(P+c)}{2P+c}$ and:*

- $F < \frac{(P-c)}{4-\kappa}\bar{\theta}$ *in the lying equilibrium;*
- $F < \frac{(P-c)(1-\kappa)}{4-3\kappa}\bar{\theta}$ *in the evasion equilibrium.*

Otherwise, welfare decreases with F.

Thus, a higher bias in the sender's monetary incentives has a net positive effect on the receiver's welfare if κ and F are sufficiently small. In this case, there is a sufficiently large share of uninformed senders (due to small κ), who are involved in guilt-driven deception (due to small F). Hence, higher bias works positively due to reducing this kind of deception. In fact, whenever there is *any* guilt-driven deception while the evasion equilibrium exists, higher F has a net positive effect on welfare (see Subsection 3.5). To sum up, biased incentives of the sender may be eventually beneficial to the receiver, insofar as they help to mitigate excessive precaution of the sender.

One can demonstrate another notable result regarding the effect of complete elimination of the monetary conflict of interest, i.e., setting $F = 0$. For example, this can correspond to banning of commissions for financial advisors, which has been recently undertaken in the UK, as noted in the Introduction. It follows from Proposition 9 that $F = 0$ is always suboptimal as far as either lying or evasion equilibrium exists. At the same time, Propositions 1 and 2 suggest that these equilibria exist for $F = 0$ only if

$\kappa > \frac{P+c}{P}$. To obtain an equilibrium prediction for the case if $\kappa \leq \frac{P+c}{P}$ (i.e., the share of uninformed sender types is sufficiently high), let us lift Assumption 2.[16] Then, one can show that setting $F = 0$ may lead to a welfare loss also through a reversal to a completely uninformative equilibrium.

Proposition 10 *If $\kappa \leq \frac{P+c}{P}$ then setting $F = 0$ leads to either the evasion or the lying equilibrium of Subtype 1 and to a lowest possible ex ante receiver welfare.*

Indeed, if $\kappa \leq \frac{P+c}{P}$, i.e., the unconditional share of uninformed sender types is sufficiently high, no informative equilibria are incentive-compatible as far as $F = 0$. The reason for this is that the receiver prefers to invest even after the lowest message $m_{B'}$ since too many truly uninformed sender types try to induce abstaining with $m_{B'}$. In this case, the only possible equilibria are when the receiver invests after any message, i.e., either the lying or the evasion equilibrium of Subtype 1.[17] Since the advice is completely uninformative in this case, the receiver has the lowest possible ex ante payoff in comparison to any other equilibria, which are possible with higher F.

Thus, eliminating the sender's monetary interest in recommending risky options for the receiver may sometimes lead to excessive downgrading of information by senders who are not certain of the eventual outcome, inducing a reversal to an uninformative equilibrium and, consequently, a welfare loss.

4.3 Regulation of the quality of the sender's information

Another conventional way of regulating advice is to ensure minimum standards of conduct for experts towards their clients (for example, a minimum time spent with the client for financial advisors). Common intuition suggests that once the receiver is rational, and thus cannot be made worse off by communication with the sender, she should benefit from the sender being more informed. However, as shown below, guilt aversion may cause a negative externality of such policy for the consumer's welfare under certain conditions. The reason is that for an ex-ante less informed expert it is sometimes easier to commit to providing truthful advice once he is guilt-averse.

The standard of conduct in our setting can be thought as corresponding to the parameter κ, denoting the probability that the sender knows the consequences of the receiver's investment decision (i.e., knows sufficiently well both the receiver's needs and the product he offers). The following proposition summarizes the comparative statics of the receiver's welfare with respect to κ.

[16] The assumption is not critical for all the previous results and helps only to eliminate multiplicity of equilibria: Without this assumption, the uninformative lying and evasion equilibria of Subtype 1 are also possible whenever informative equilibria exist.

[17] Since Assumption 2 is lifted, one could then allow for out-of-equilibrium messages that do not lead to abstaining (or to a lower guilt), thus making the sender's strategy (which implies a negative utility due to guilt not compensated by F) consistent with these equilibria.

Proposition 11 *i) In the lying equilibrium, $U^{r,l}$ is U-shaped in κ if*

$$\bar{\theta}\frac{P-c}{4-\kappa} > F > \bar{\theta}\frac{(P-c)^2}{2(3P+c)}$$

and increasing in κ otherwise.

ii) In the evasion equilibrium, $U^{r,e}$ is strictly decreasing in κ if

$$\frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa} > F > \max[\frac{\bar{\theta}(1-\kappa)(P-c)}{1+\sqrt{1+\kappa^2}}, \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}]$$

and increasing in κ otherwise.

Consider the lying equilibrium. There, an increase in κ has two effects: a direct positive and an indirect negative. The positive effect relates to the fact that once the sender is informed (i.e., is either in state G' or B') the receiver is more likely to invest in the good than in the bad state of the world. The reason for this is that the share of types inducing investment in state B' is lower than such share in state G' (unless the equilibrium is of Subtype 1 where these are equal). On the other hand, if the sender is uninformed, then the receiver is equally likely to invest in both states of the world. Moreover, there can be foregone investment in the good state of the world (due to guilt-driven deception), which never happens with an informed sender. Altogether, this implies that

$$E[U^{r,l}|G' \vee B'] \geq E[U^{r,l}|N'], \tag{11}$$

so that an increase in κ and, hence, the probability of facing an informed sender has a positive effect on welfare (all else equal).

However, an increase in κ also has an indirect negative effect on welfare through guilt aversion. In particular, higher κ implies that the message $m_{G'}$ is more likely to be sent by informed types, which by (11) should lead to a higher expected receiver's utility conditional on the message (all else equal). This results in higher expected guilt of the sender from sending $m_{G'}$, pushing the cutoff types down. This can result in an increase in guilt-driven deception , which under some parameter values can overweigh the positive effect described in the preceding paragraph.

In the evasion equilibrium, an increase in κ again has two effects. The first (positive) effect is the same as in the previous case: the receiver prefers (from the ex ante perspective) to deal with an informed rather than an uninformed sender. The second (negative) effect is driven by guilt aversion: higher κ *decreases* the persuasiveness of the evasive message $m_{N'}$. In particular, higher κ implies that the share of truly uninformed types is lower, so that the evasive message $m_{N'}$ becomes less credible and rather signals types in state B' who want to conceal their bad news. By being less persuasive, the message $m_{N'}$ induces less guilt on the part of the sender so that the equilibrium cutoffs increase. If the equilibrium is of Subtype 2, this leads to a decrease in welfare due to the spread of bias-driven deception. Under certain parameter values this can render the total effect of higher κ to be negative.

Thus, the ex ante likelihood of obtaining the informative signal affects the sender's anticipation of guilt, and hence the rate of truth-telling conditional on a given information state. This mechanism leads to seemingly paradoxical cases where even a completely rational receiver prefers to deal with the sender, who is less likely to have the information she needs.[18]

5 Comparison to an outcome-based model

One may ask whether the obtained results are attributable to belief-dependent preferences, and cannot be derived with a simpler model based only on outcome-based preferences. To address this question, let us assume that the sender cares solely about his and the receiver's monetary outcomes. In particular, the sender has some fixed cost ψ of incurring the receiver's losses, e.g., arising from inequity aversion (Bolton and Ockenfels 2000 and Fehr and Schmidt 1999), which he bears if the receiver gets a negative payoff of c. For simplicity, I assume that the sender has no change in utility if the receiver gets 0 or P, although the subsequent results are easily generalizable if one assumes the opposite.[19] Then, the sender's utility conditional on investment, denoted by $U_{i^s}^{s,0}$, can be represented in the following form:

$$U_{i^s}^{s,0}(\theta, I) = F - \Pr[B|i^s] \cdot \theta\psi, \tag{12}$$

where θ is the sensitivity parameter, distributed uniformly as above on $(0, \bar{\theta}]$. Conditional on abstaining, we have

$$U_{i^s}^{s,0}(\theta, A) = 0. \tag{13}$$

It is then straightforward to show that the equilibrium in this model also has a cutoff structure. The cutoff type, once interior, must be indifferent between investment and abstaining:

$$\begin{aligned} F - \Pr[B|i^s] \cdot \widehat{\theta}_{i^s}^0 \psi &= 0, \\ \widehat{\theta}_{i^s}^0 &= \frac{F}{\Pr[B|i^s]\psi}. \end{aligned} \tag{14}$$

Importantly, the cutoffs are the same independently of the equilibrium message structure (which, as before, comprises the pooling case (the lying equilibrium), and the separating case (the evasion equilibrium)).[20]

Let us consider whether the lying prohibition policy in the outcome-based model has the same welfare effects as those considered in Subsection 4.1. As in the case of the belief-dependent preferences, the prohibition of verifiable lying (i.e., pooling with types in

[18] A positive effect of noise in the sender's information (through a different mechanism than considered here) is also found by Blume et al. (2007) within the benchmark cheap-talk framework of Crawford and Sobel (1982).

[19] The only assumption that matters here is that the sender's utility does not directly depend on beliefs.

[20] The proof is available upon request.

state G') results in a switch from the lying to either the evasion or the evasive babbling equilibrium. However, the welfare effect of this switch is very different from the one in the main model.

Proposition 12 *Under the outcome-based preferences, prohibition of lying never results in a decrease in welfare.*

Intuitively, welfare does not change if there is a switch from the lying to the evasion equilibrium. The reason for this is that the cutoffs (and hence the ex ante investment efficiency) are the same in both equilibria, since they do not depend on beliefs induced by the messages, but only on their monetary consequences (see (14)). In the other possible scenario, if there is a switch to the evasive babbling equilibrium (whenever the evasion equilibrium cannot emerge), welfare always increases. This is because, if investment conditional on $i^s \neq G'$ is ex-ante inefficient in the (candidate) evasion equilibrium (while it does not exist), such investment must be also inefficient in the lying equilibrium with the same cutoffs. Since a switch from the lying to the evasive babbling equilibrium eliminates this case of investment, welfare increases.

Regarding other results, one can show that, in the outcome-based model, reducing or banning commissions may also have negative welfare effects due to the resulting inefficient precaution of uninformed sender types. There can also be a switch to an uninformative equilibrium if commissions are eliminated, similar to Proposition 10.[21] However, the result that the receiver's welfare can decline with the quality of the sender's private signal cannot be explained by outcome-based preferences.

Proposition 13 *Under outcome-based preferences, the welfare never decreases with κ.*

The reason for this is that now the increase in κ does not affect the sender's utility from inducing investment, and hence the equilibrium cutoffs and the rate of deception in each state. Consequently, the only result of such increase is that the receiver is more likely to obtain more ex ante efficient advice from an informed sender.

Thus, outcome-based preferences alone cannot explain the previous results that welfare can decrease due to lying prohibition or making the sender's signal more precise.

6 Conclusion

This paper studies a model of strategic communication where the sender's cost of lying is endogenously determined by the receiver's beliefs. One of the main results is that such preferences can lead to the endogenous emergence of evasive communication in equilibrium (i.e., pretending to be uninformed), which helps to avoid guilt from biasing the receiver's expectations. At the same time, there also exists an alternative equilibrium where only explicit communication is used, precluding any meaningful evasive messages and, thus, disciplining guilt-sensitive senders. A policy aimed at lying prohibition may induce a

[21] The proof is available upon request.

switch from explicit to evasive communication, which can have both negative and positive welfare consequences. A negative effect occurs since deception is less psychologically costly once it takes the form of evasive messages, while a positive effect is related to enhancing communication for truly uncertain senders. The net effect can be both positive and negative, depending on the precision of the sender's information and his monetary preference bias.

I also find that monetary commissions for the advisor can enhance efficiency of communication by serving as a reasonable compensation for the risk, which the caring advisor faces while being uncertain about the suitability of his advice. Hence, banning such commissions can result in an inefficient degree of precaution by providing advice. Finally, a policy imposing a higher (ex-ante) precision of the sender's private information can harm the receiver in certain cases. This occurs due to altered credibility of equilibrium messages, and hence the associated guilt, which can lead to a situation when an ex-ante more informed sender is less prone to truth-telling.

The analysis generally suggests that policies aimed at restriction of experts' lying, their conflict of interest with consumers or standards of care should be applied with caution. In certain cases, such policies can backfire by interfering with otherwise efficient self-regulatory mechanisms, which are based on the intrinsic and natural motivation of experts to provide good advice to their clients.

Appendix: Omitted Proofs

Proof of Lemma 1. We have

$$\begin{aligned}
\eta(m) &\equiv \Pr[G|m] = \frac{\Pr[m|G]\Pr[G]}{\Pr[m]} \\
&= \frac{(\Pr[m|G' \cap G]\Pr[G'|G] + \Pr[m|N' \cap G]\Pr[N'|G])\Pr[G]}{\Pr[m|G']\Pr[G'] + \Pr[m|N']\Pr[N'] + \Pr[m|B']\Pr[B']} \\
&= \frac{(\Pr[m|G']\Pr[G'|G] + \Pr[m|N']\Pr[N'|G])\Pr[G]}{\Pr[m|G']\Pr[G'] + \Pr[m|N']\Pr[N'] + \Pr[m|B']\Pr[B']} \\
&= \frac{\Pr[m|G']\kappa + \Pr[m|N'](1-\kappa)}{(\Pr[m|G'] + \Pr[m|B'])\kappa + 2\Pr[m|N'](1-\kappa)},
\end{aligned} \tag{15}$$

where the second equality is by Bayes rule, the third equality is by the law of total probability, and the fourth equality is by the fact that equilibrium messages of any sender type θ are contingent on his information state i^s, but not on the true state of the world (by the definition of equilibrium in Subsection 2.2), i.e., $\Pr[m|i^s \cap G] = \Pr[m|i^s]$ for any i^s. ■

Proof of Lemma 2. Assume by contradiction that all messages used on the equilibrium path induce conditional beliefs lower than $\underline{\eta}$ (so that the receiver abstains after all of them, see (8)), i.e.,

$$\forall m \in \Upsilon_E, \Pr[G|m] < \underline{\eta}, \tag{16}$$

where Υ_E is the union set of all messages used in equilibrium. By Bayes rule,

$$\Pr[G|\Upsilon_E] = \frac{\Pr[\Upsilon_E|G]\Pr[G]}{\Pr[\Upsilon_E]} = \Pr[G] = 0.5, \tag{17}$$

since $\Pr[\neg\Upsilon_E|G] = \Pr[\neg\Upsilon_E] = 0$ (by definition of Υ_E) so that $\Pr[\Upsilon_E|G] = \Pr[\Upsilon_E] = 1$. At the same time, since messages are mutually exclusive events,

$$\Pr[\Upsilon_E|G] = \sum_{m\in\Upsilon_E}\Pr[m|G] = \sum_{m\in\Upsilon_E}\frac{\Pr[G|m]\Pr[m]}{\Pr[G]}, \tag{18}$$

where the last equality is by Bayes rule. Substituting (18) and $\Pr[\Upsilon_E] = 1$ into (17) we obtain

$$\begin{aligned}\Pr[G|\Upsilon_E] &= \frac{\Pr[\Upsilon_E|G]\Pr[G]}{\Pr[\Upsilon_E]}\\ &= \frac{\sum_{m\in\Upsilon_E}\frac{\Pr[G|m]\Pr[m]}{\Pr[G]}\Pr[G]}{1} = \sum_{m\in\Upsilon_E}\Pr[G|m]\Pr[m]\\ &< \sum_{m\in\Upsilon_E}\underline{\eta}\Pr[m] = \underline{\eta}\sum_{m\in\Upsilon_E}\Pr[m] = \underline{\eta} = \frac{-c}{P-c} < 0.5,\end{aligned} \tag{19}$$

where the first inequality is by (16) and the second one is by Assumption 1. Since (19) contradicts (17), the claim follows. ■

Lemma 7 *A necessary condition for the existence of equilibrium is that the persuasiveness of any investment-inducing message used in states $i^s \in \{B', N'\}$ is the same.*

Proof. Assume by contradiction that there exist two investment-inducing messages m_1 and m_2 sent in at least some of the information states $i^s \in \{B', N'\}$ such that $\eta(m_1) > \eta(m_2)$. Then, any type θ would strictly prefer m_2 over m_1 in any information state $i^s \in \{B', N'\}$ since as far as $i^s \neq G'$ (so that $\lambda_{i^s} > 0$) we have

$$U^s_{i^s}(\theta, \eta(m_1), I) = F - \theta\lambda_{i^s}\eta(m_1)(P-c) < F - \theta\lambda_{i^s}\eta(m_2)(P-c) = U^s_{i^s}(\theta, \eta(m_2), I). \tag{20}$$

Thus, m_1 cannot be an equilibrium message. ■

Proof of Lemma 3. Assume by contradiction that some sender type θ in state G' sends a message leading to abstaining, so that his expected utility from sending this message is 0 (by (7)). At the same time, by Lemma 2 there exists at least one message m' leading to investment. By (9), the expected utility from sending this message for the considered sender type is

$$U^s_{G'}(\theta, \eta(m'), I) = F - \theta\lambda_{G'}\eta(m)(P-c) = F > 0, \tag{21}$$

where the second equality follows from the fact that $\lambda_{G'} = \Pr[B|G'] = 0$. Consequently, the sender would have a strict incentive to deviate to m', which yields a contradiction. ■

Proof of Lemma 4. Assume that equilibrium exists. By Lemma 2 there exists at least one message m' leading to investment. Then, there always exists sufficiently small $\widetilde{\theta}$ such that

$$U^s_{i^s}(\widetilde{\theta}, \eta(m'), I) = F - \widetilde{\theta}\lambda_{i^s}\eta(m')(P - c) \geq 0 = U^s_{i^s}(\widetilde{\theta}, \cdot, A), \tag{22}$$

where the first equality is by (9), and the second by (7). Thus, in any equilibrium at least some types in states $i^s \neq G'$ must prefer an investment-inducing message.

Next, by Lemma 7 all messages used by types in states B' and N' have the same persuasiveness, which I denote by η'. Consider further the following possible cases depending on the utility of the most guilt sensitive type in some state $i^s \in \{B', N'\}$.

Case 1: $U^s_{i^s}(\bar{\theta}, \eta', I) \geq 0$. In this case, type $\bar{\theta}$ prefers to induce investment over abstaining (since the latter action yields zero utility). Then, since $U^s_{i^s}(\theta, \eta, I)$ is continuously decreasing in θ for given η, it follows that for all $\theta < \bar{\theta}$ it holds

$$U^s_{i^s}(\theta, \eta', I) > 0, \tag{23}$$

that is all these types should induce investment in equilibrium as well.

Case 2: $U^s_{i^s}(\bar{\theta}, \eta', I) < 0$. In this case, since at the same time $U^s_{i^s}(0, \eta', I) = F > 0$, by the intermediate value theorem there must exist type $\widehat{\theta}_{i^s} \in (0, \bar{\theta})$ such that

$$U^s_{i^s}(\widehat{\theta}_{i^s}, \eta', I) = 0. \tag{24}$$

Then, for all $\theta \leq \widehat{\theta}_{i^s}$ it holds $U^s_{i^s}(\theta, \eta', I) \geq 0$ (so that these types prefer to induce investment), while for all $\theta > \widehat{\theta}_{i^s}$ it holds $U^s_{i^s}(\theta, \eta', I) < 0$ (so that these types prefer to induce abstaining). Finally, in the considered case there must be at least one message leading to abstaining in equilibrium so that the types $\theta > \widehat{\theta}_{i^s}$ can choose this message. Otherwise, these types would have a strict incentive to deviate to an out-of-equilibrium message leading to abstaining and get 0 instead of a negative utility (such out-of-equilibrium messages exist by Assumption 2).

Hence, in all possible cases, if there exists an equilibrium, then it must have the cutoff structure described in the lemma. ■

Lemma 8 *A necessary condition for the existence of equilibrium is* $\widehat{\theta}_{N'} \geq \widehat{\theta}_{B'}$.

Proof. Assume by contradiction that there exists an equilibrium such that

$$\widehat{\theta}_{N'} < \widehat{\theta}_{B'}. \tag{25}$$

Then, all types in the interval $(\widehat{\theta}_{N'}, \widehat{\theta}_{B'})$ induce abstaining in state N' and induce investment in state B'. Consequently, these types have a negative utility from any investment-inducing message in state N' (otherwise they would deviate to such message from inducing abstaining, which yields the utility of 0), and a positive utility from inducing investment in state B' (otherwise, they would deviate to inducing abstaining). Hence, for any $\theta' \in (\widehat{\theta}_{N'}, \widehat{\theta}_{B'})$ it holds

$$U^s_{N'}(\theta', \widehat{\eta}, I) < 0 \leq U^s_{B'}(\theta', \widehat{\eta}, I), \tag{26}$$

where $\widehat{\eta}$ is the persuasiveness of investment-inducing messages in states $i^s \in \{B', N'\}$ (unique by Lemma 7). At the same time, for any $\theta \in (0, \bar{\theta}]$ it holds (given (9))

$$U^s_{B'}(\theta, \widehat{\eta}, I) = F - \theta\widehat{\eta}(P - c) < F - 0.5\theta\widehat{\eta}(P - c) = U^s_{N'}(\theta, \widehat{\eta}, I), \tag{27}$$

which yields a contradiction to (26). ■

Lemma 9 *A necessary condition for the existence of equilibrium is that either $\widehat{\theta}_{i^s} < \bar{\theta}$ and $U^s_{i^s}(\widehat{\theta}_{i^s}, \eta(m(\widehat{\theta}_{i^s})), I) = 0$ or $\widehat{\theta}_{i^s} = \bar{\theta}$ and $U^s_{i^s}(\widehat{\theta}_{i^s}, \eta(m(\widehat{\theta}_{i^s})), I) \geq 0$.*

Proof. Denote by $\widehat{\eta}$ the persuasiveness of investment-inducing messages in states $i^s \in \{B', N'\}$ (unique by Lemma 7). Assume by contradiction that equilibrium exists and either $\widehat{\theta}_{i^s} < \bar{\theta}$ and $U^s_{i^s}(\widehat{\theta}_{i^s}, \widehat{\eta}, I) \neq 0$ or $\widehat{\theta}_{i^s} = \bar{\theta}$ and $U^s_{i^s}(\widehat{\theta}_{i^s}, \widehat{\eta}, I) < 0$ for some $i^s \in \{B', N'\}$. Let us demonstrate a contradiction in each of these cases separately.

Case 1: $\widehat{\theta}_{i^s} < \bar{\theta}$ and $U^s_{i^s}(\widehat{\theta}_{i^s}, \widehat{\eta}, I) \neq 0$.

First, let us consider the case $U^s_{i^s}(\widehat{\theta}_{i^s}, \widehat{\eta}, I) > 0$. Then, since $U^s_{i^s}(\widehat{\theta}_{i^s}, \widehat{\eta}, I)$ is continuously decreasing in the first argument (see (9)), there would exist types sufficiently close to $\widehat{\theta}_{i^s}$ such that $\theta > \widehat{\theta}_{i^s}$ and $U^s_{i^s}(\theta, \widehat{\eta}, I) > 0$. At the same time, according to Lemma 4 such types induce abstaining and get the utility of 0, so that they would have a strict incentive to deviate to an investment-inducing message yielding the positive utility $U^s_{i^s}(\theta, \widehat{\eta}, I)$. Analogously, if $U^s_{i^s}(\widehat{\theta}_{i^s}, \widehat{\eta}, I) < 0$, then types sufficiently close to $\widehat{\theta}_{i^s}$ on the left have a negative utility from inducing investment and hence a strict incentive to deviate to inducing abstaining. We have come to contradiction in all possible cases.

Case 2: $\widehat{\theta}_{i^s} = \bar{\theta}$ and $U^s_{i^s}(\widehat{\theta}_{i^s}, \widehat{\eta}, I) < 0$.

Then, analogously to the previous case, there would exist types θ such that $\theta < \widehat{\theta}_{i^s}$ and $U^s_{i^s}(\theta, \widehat{\eta}, I) < 0$, which then would like to deviate at least to an out-of-equilibrium message leading to abstaining (existing by Assumption 2) that yields a contradiction. ■

Lemma 10 *A necessary condition for the existence of equilibrium is* $\widehat{\theta}_{N'} = \min[\bar{\theta}, 2\widehat{\theta}_{B'}]$.

Proof. Denote by $\widehat{\eta}$ the persuasiveness of investment-inducing messages in states $i^s \in \{B', N'\}$ (unique by Lemma 7). Assume by contradiction that equilibrium exists and $\widehat{\theta}_{N'} \neq \min[\bar{\theta}, 2\widehat{\theta}_{B'}]$. Let us consider two possible cases $\widehat{\theta}_{N'} < \bar{\theta}$ and $\widehat{\theta}_{N'} = \bar{\theta}$.

Case 1: $\widehat{\theta}_{N'} < \bar{\theta}$. By Lemma 8 it then follows $\widehat{\theta}_{B'} < \bar{\theta}$. Consequently, by Lemma 9 and equation (9) we have

$$\begin{aligned} U^s_{B'}(\widehat{\theta}_{B'}, \widehat{\eta}, I) &= F - \widehat{\theta}_{B'}\widehat{\eta}(P - c) = 0, &(28)\\ U^s_{N'}(\widehat{\theta}_{N'}, \widehat{\eta}, I) &= F - 0.5\widehat{\theta}_{N'}\widehat{\eta}(P - c) = 0. &(29)\end{aligned}$$

This implies

$$\begin{aligned}\widehat{\theta}_{B'} &= \frac{F}{\widehat{\eta}(P-c)}, &(30)\\ \widehat{\theta}_{N'} &= \frac{F}{0.5\widehat{\eta}(P-c)}, &(31)\end{aligned}$$

and, consequently,

$$\widehat{\theta}_{N'} = 2\widehat{\theta}_{B'}. \tag{32}$$

This, together with $\widehat{\theta}_{N'} < \bar{\theta}$, implies $\min[\bar{\theta}, 2\widehat{\theta}_{B'}] = 2\widehat{\theta}_{B'}$, which finally yields a contradiction to $\widehat{\theta}_{N'} \neq \min[\bar{\theta}, 2\widehat{\theta}_{B'}]$.

Case 2: $\widehat{\theta}_{N'} = \bar{\theta}$. Since we have also assumed $\widehat{\theta}_{N'} \neq \min[\bar{\theta}, 2\widehat{\theta}_{B'}]$, this implies

$$\begin{aligned}\bar{\theta} &\neq \min[\bar{\theta}, 2\widehat{\theta}_{B'}] &(33)\\ &\Rightarrow \widehat{\theta}_{B'} < 0.5\bar{\theta}. &(34)\end{aligned}$$

Next, we have

$$\begin{aligned}U^s_{B'}(0.5\bar{\theta}, \widehat{\eta}, I) &= F - 0.5\bar{\theta}\widehat{\eta}(P-c)\\ &= U^s_{N'}(\bar{\theta}, \widehat{\eta}, I) \geq 0, &(35)\end{aligned}$$

where the last inequality holds due to $\widehat{\theta}_{N'} = \bar{\theta}$ and Lemma 9. Consequently, for any $\theta < 0.5\bar{\theta}$ it holds $U^s_{B'}(\theta, \widehat{\eta}, I) > 0$ so that all types with $\theta < 0.5\bar{\theta}$ should strictly prefer to induce investment, contradicting to (34). ■

Lemma 11 *If the strategies and beliefs are specified according to Definition 1, then the sender has no incentives to deviate if and only if the following two conditions hold:*

1) either $\widehat{\theta}^l_{B'} < \bar{\theta}$ *and* $U^s_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{G'}), I) = 0$ *or* $\widehat{\theta}^l_{B'} = \bar{\theta}$ *and* $U^s_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{G'}), I) \geq 0$.

2) $\widehat{\theta}^l_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^l_{B'}]$.

Proof. The necessity of the both conditions follows by Lemmas 9 and 10. Let us consider their sufficiency and assume that both conditions hold. First, no type in state G' has an incentive to deviate to $m_{B'}$, while his utility from investment is always strictly positive (see (21)). Second, it is straightforward to show that, once the first condition holds, no type in state B' has a strict incentive to deviate to another equilibrium message given that $U^s_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{G'}), I)$ is continuously decreasing in its first argument.

Let us show that all types in state N' also do not have incentives to deviate to another equilibrium message. Consider the following possible cases.

Case 1: $\widehat{\theta}^l_{B'} < 0.5\bar{\theta}$. Then, by assumption, $\widehat{\theta}^l_{N'} = 2\widehat{\theta}^l_{B'} < \bar{\theta}$. Hence,

$$\begin{aligned}
U^s_{N'}(\widehat{\theta}^l_{N'}, \eta(m_{G'}), I) &= U^s_{N'}(2\widehat{\theta}^l_{B'}, \eta(m_{G'}), I) \\
&= F - 0.5 \cdot 2\widehat{\theta}^l_{B'}\eta(m_{G'})(P-c) \\
&= F - \widehat{\theta}^l_{B'}\eta(m_{G'})(P-c) \\
&= U^s_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{G'}), I) = 0, \qquad (36)
\end{aligned}$$

where the last equality follows from $\widehat{\theta}^l_{B'} < 0.5\bar{\theta}$ and the first condition of the lemma. Then, by the same arguments as in the case of state B', no sender types in state N' have an incentive to deviate.

Case 2: $\widehat{\theta}^l_{B'} \geq 0.5\bar{\theta}$. Then, by assumption, $\widehat{\theta}^l_{N'} = \bar{\theta}$. This yields

$$\begin{aligned}
U^s_{N'}(\widehat{\theta}^l_{N'}, \eta(m_{G'}), I) &= U^s_{N'}(\bar{\theta}, \eta(m_{G'}), I) \\
&= F - 0.5\bar{\theta}\eta(m_{G'})(P-c) \\
&\geq F - \widehat{\theta}^l_{B'}\eta(m_{G'})(P-c) \\
&= U^s_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{G'}), I) \geq 0, \qquad (37)
\end{aligned}$$

where the first inequality follows from $\widehat{\theta}^l_{B'} \geq 0.5\bar{\theta}$ and the second one from the first condition of the lemma. Then, for all $\theta < \widehat{\theta}^l_{N'}$ it holds $U^s_{N'}(\theta, \eta(m_{G'}), I) > 0$, so that no sender type in state N' have an incentive to deviate.

Finally, no type in either state has an incentive to deviate to out-of-equilibrium messages, which either induce abstaining (while all types inducing investment have a positive utility) or are more persuasive (see Definition 1). ■

Lemma 12 *If the strategies and beliefs are specified as in Definition 1 and* $\widehat{\theta}^l_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^l_{B'}]$, *then* $U^s_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{G'}|\widehat{\theta}^l_{B'}), I)$ *is continuous and strictly decreasing in* $\widehat{\theta}^l_{B'}$ *on* $(0, \bar{\theta}]$.

Proof. We have

$$\frac{dU^s_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{G'}|\widehat{\theta}^l_{B'}), I)}{d\widehat{\theta}^l_{B'}} = \frac{\partial U^s_{B'}}{\partial \widehat{\theta}^l_{B'}} + \frac{\partial U^s_{B'}}{\partial \eta} \frac{d\eta(m_{G'}|\widehat{\theta}^l_{B'})}{d\widehat{\theta}^l_{B'}}. \qquad (38)$$

Substituting for $U^s_{B'}$ from (9) we get

$$\frac{dU^s_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{G'}|\widehat{\theta}^l_{B'}), I)}{d\widehat{\theta}^l_{B'}} = -(P-c)\widehat{\theta}^l_{B'} \left(\frac{d\eta(m_{G'}|\widehat{\theta}^l_{B'})}{d\widehat{\theta}^l_{B'}} + \frac{\eta(m_{G'}|\widehat{\theta}^l_{B'})}{\widehat{\theta}^l_{B'}} \right). \qquad (39)$$

Further, consider the following possible cases given that $\widehat{\theta}^l_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^l_{B'}]$.

Case 1: $\widehat{\theta}^l_{B'} \in (0, 0.5\bar{\theta}]$, $\widehat{\theta}^l_{N'} = 2\widehat{\theta}^l_{B'}$.

Then, by (15) (substituting $\Pr[m_{G'}|G'] = 1$, $\Pr[m_{G'}|N'] = 2\widehat{\theta}^l_{B'}/\bar{\theta}$ and $\Pr[m_{G'}|B'] = \widehat{\theta}^l_{B'}/\bar{\theta}$)

$$\eta(m_{G'}|\widehat{\theta}^l_{B'}) = \frac{2\widehat{\theta}^l_{B'}(1-\kappa)+\kappa\bar{\theta}}{\widehat{\theta}^l_{B'}(4-3\kappa)+\kappa\bar{\theta}}. \tag{40}$$

This function is convex in $\widehat{\theta}^l_{B'}$:

$$\frac{d^2\eta(m_{G'}|\widehat{\theta}^l_{B'})}{d(\widehat{\theta}^l_{B'})^2} = \frac{2(2-\kappa)(4-3\kappa)\kappa\bar{\theta}}{((4-3\kappa)\widehat{\theta}^l_{B'}+\kappa\bar{\theta})^3} > 0. \tag{41}$$

The convexity implies

$$\frac{d\eta(m_{G'}|\widehat{\theta}^l_{B'})}{d\widehat{\theta}^l_{B'}} > \frac{\eta(m_{G'}|\widehat{\theta}^l_{B'})-\eta(m_{G'}|0)}{\widehat{\theta}^l_{B'}} = \frac{\eta(m_{G'}|\widehat{\theta}^l_{B'})-1}{\widehat{\theta}^l_{B'}}, \tag{42}$$

where the last equality is by (40). Then, coming back to (39) we have

$$\begin{aligned} & \frac{d\eta(m_{G'}|\widehat{\theta}^l_{B'})}{d\widehat{\theta}^l_{B'}} + \frac{\eta(m_{G'}|\widehat{\theta}^l_{B'})}{\widehat{\theta}^l_{B'}} \\ > \; & \frac{\eta(m_{G'}|\widehat{\theta}^l_{B'})-1}{\widehat{\theta}^l_{B'}} + \frac{\eta(m_{G'}|\widehat{\theta}^l_{B'})}{\widehat{\theta}^l_{B'}} \\ = \; & \frac{2\eta(m_{G'}|\widehat{\theta}^l_{B'})-1}{\widehat{\theta}^l_{B'}} = \frac{1}{\widehat{\theta}^l_{B'}}\frac{\kappa(\bar{\theta}-\widehat{\theta}^l_{B'})}{\widehat{\theta}^l_{B'}(4-3\kappa)+\bar{\theta}\kappa} > 0, \end{aligned} \tag{43}$$

where the first inequality is by (42) and the last equality by (40). Taken together, (43) and (39) lead to the claim for $\widehat{\theta}^l_{B'} \in (0, 0.5\bar{\theta}]$.

Case 2: $\widehat{\theta}^l_{B'} \in (0.5\bar{\theta}, \bar{\theta}]$, $\widehat{\theta}^l_{N'} = \bar{\theta}$.

In this case, $\Pr[m_{G'}|G'] = 1$, $\Pr[m_{G'}|N'] = 1$ and $\Pr[m_{G'}|B'] = \widehat{\theta}^l_{B'}/\bar{\theta}$ so that (15) implies

$$\eta(m_{G'}|\widehat{\theta}^l_{B'}) = \frac{\bar{\theta}}{\kappa\widehat{\theta}^l_{B'}+(2-\kappa)\bar{\theta}}. \tag{44}$$

Then, it is possible to obtain a simple closed-form solution for the RHS of (39):

$$\begin{aligned} & \frac{d\eta(m_{G'}|\widehat{\theta}^l_{B'})}{d\widehat{\theta}^l_{B'}} + \frac{\eta(m_{G'}|\widehat{\theta}^l_{B'})}{\widehat{\theta}^l_{B'}} \\ = \; & -\frac{\kappa\bar{\theta}}{(\kappa\widehat{\theta}^l_{B'}+(2-\kappa)\bar{\theta})^2} + \frac{\bar{\theta}}{\widehat{\theta}^l_{B'}(\kappa\widehat{\theta}^l_{B'}+(2-\kappa)\bar{\theta})} \\ = \; & \frac{(2-\kappa)\bar{\theta}^2}{\widehat{\theta}^l_{B'}(\kappa\widehat{\theta}^l_{B'}+(2-\kappa)\bar{\theta})^2} > 0. \end{aligned} \tag{45}$$

Taken together, (45) and (39) lead to the claim for for $\widehat{\theta}^l_{B'} \in (0.5\bar{\theta}, \bar{\theta}]$.

Finally, since $\eta(m_{G'}|\widehat{\theta}^l_{B'})$, and hence $U^s_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{G'}|\widehat{\theta}^l_{B'}), I)$, is continuous at $\widehat{\theta}^l_{B'} = 0.5\bar{\theta}$ the claim holds for the whole interval $(0, \bar{\theta}]$. ■

Lemma 13 *If the strategies and beliefs are specified as in Definition 1, then for any given parameter values there always exist unique cutoffs $\widehat{\theta}^l_{B'}$ and $\widehat{\theta}^l_{N'}$ such that the sender does not have incentives to deviate. Moreover:*

1) If $F \geq 0.5\bar{\theta}(P-c)$ then $\widehat{\theta}^l_{B'} = \widehat{\theta}^l_{N'} = \bar{\theta}$.

2) If $F \in [\bar{\theta}\frac{P-c}{4-\kappa}, 0.5\bar{\theta}(P-c))$ then $\widehat{\theta}^l_{B'} \in [0.5\bar{\theta}, \bar{\theta})$ and $\widehat{\theta}^l_{N'} = \bar{\theta}$.

3) If $F < \bar{\theta}\frac{P-c}{4-\kappa}$ then $\widehat{\theta}^l_{B'} \in (0, 0.5\bar{\theta})$ and $\widehat{\theta}^l_{N'} = 2\widehat{\theta}^l_{B'}$.

Proof. For notational simplicity denote

$$\varpi(\widehat{\theta}^l_{B'}) = U^s_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{G'}|\widehat{\theta}^l_{B'}), I) \tag{46}$$

with $\widehat{\theta}^l_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^l_{B'}]$. Consider the cases listed in the lemma.

Case 1: $F \geq 0.5\bar{\theta}(P-c)$. Consider the behavior of $\varpi(\cdot)$ at the exterior point $\bar{\theta}$. First, note that

$$\eta(m_{G'}|\widehat{\theta}^l_{B'} = \bar{\theta}) = 0.5, \tag{47}$$

which results from substituting $\Pr[m_{G'}|G'] = \Pr[m_{G'}|N'] = \Pr[m_{G'}|B'] = 1$ into (15). Then,

$$\varpi(\bar{\theta}) = F - \bar{\theta}\eta(m_{G'}|\widehat{\theta}^l_{B'} = \bar{\theta})(P-c) = F - 0.5\bar{\theta}(P-c) \geq 0, \tag{48}$$

where the second equality follows from (47) and the inequality by the assumption of the case. Then, the sender has no incentives to deviate if $\widehat{\theta}^l_{B'} = \widehat{\theta}^l_{N'} = \bar{\theta}$ by Lemma 11. At the same time, it follows from Lemma 12 and (48) that

$$\forall \widehat{\theta}^l_{B'} < \bar{\theta}, \varpi(\widehat{\theta}^l_{B'}) > 0. \tag{49}$$

Consequently, by Lemma 11 there are no possible cutoffs except for $\widehat{\theta}^l_{B'} = \widehat{\theta}^l_{N'} = \bar{\theta}$ such that the sender does not have an incentive to deviate.

Case 2: $F \in [\bar{\theta}\frac{P-c}{4-\kappa}, 0.5\bar{\theta}(P-c))$. We have

$$\begin{aligned} \varpi(0) &= F - 0 \cdot \eta(m_{G'}|\widehat{\theta}^l_{B'} = 0)(P-c) = F > 0, && (50) \\ \varpi(\bar{\theta}) &= F - \bar{\theta}\eta(m_{G'}|\widehat{\theta}^l_{B'} = \bar{\theta})(P-c) = F - 0.5\bar{\theta}(P-c) < 0, && (51) \end{aligned}$$

where the last inequality follows by the assumption of the case. Then, from Lemma 12 and the intermediate value theorem it follows that there exists a unique cutoff value $0 < \widehat{\theta}^l_{B'} < \bar{\theta}$ such that $\varpi(\widehat{\theta}^l_{B'}) = 0$ (the necessary and sufficient condition for an interior cutoff by Lemma 11). At the same time, the cutoff $\widehat{\theta}^l_{B'} = \bar{\theta}$ is impossible due to (51) and Lemma

11, so that the existing interior cutoff is the only possible cutoff. Lemma 11 also implies that the corresponding unique cutoff in state N' is then given by $\widehat{\theta}^l_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^l_{B'}]$.

Let us show that in this case $\widehat{\theta}^l_{B'} \geq 0.5\bar{\theta}$. By Lemma 12 it holds

$$\varpi(\widehat{\theta}^l_{B'}) = 0 \wedge \widehat{\theta}^l_{B'} \geq 0.5\bar{\theta} \Leftrightarrow \varpi(0.5\bar{\theta}) \geq 0. \tag{52}$$

From (15) we get

$$\eta(m_{G'}|\widehat{\theta}^l_{B'} = 0.5\bar{\theta}) = \frac{2}{4-\kappa} \tag{53}$$

so that

$$\varpi(0.5\bar{\theta}) = F - 0.5\bar{\theta}\frac{2}{4-\kappa}(P-c) \geq 0, \tag{54}$$

where the inequality follows from the assumption $F \in [\bar{\theta}\frac{P-c}{4-\kappa}, 0.5\bar{\theta}(P-c))$. By (52) and (54) it then follows that $\widehat{\theta}^l_{B'} \geq 0.5\bar{\theta}$.

Case 3: $F < \bar{\theta}\frac{P-c}{4-\kappa}$. From $\varpi(0) > 0$, $\varpi(\bar{\theta}) < 0$ and Lemma 12 it follows that there is a unique interior cutoff $\widehat{\theta}^l_{B'}$ with $\widehat{\theta}^l_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^l_{B'}]$. Finally, (52), the left equality in (54) and $F < \bar{\theta}\frac{P-c}{4-\kappa}$ result in $\widehat{\theta}^l_{B'} < 0.5\bar{\theta}$. ■

Proof of Proposition 1. Lemma 13 shows that for any parameter values there exist unique cutoffs $\widehat{\theta}^l_{B'}$ and $\widehat{\theta}^l_{N'}$ such that the sender does not have an incentive to deviate once the receiver plays according to the prescribed equilibrium strategy. Thus, to show the claim of the proposition we need to find the range of parameters such that the receiver's incentive constraints are satisfied given the unique possible equilibrium cutoffs $\widehat{\theta}^l_{B'}$ and $\widehat{\theta}^l_{N'}$. The receiver's incentive constraints should ensure investment after $m_{G'}$ and abstaining after $m_{B'}$ if the latter is sent in equilibrium:

$$\begin{aligned} \eta(m_{G'}|\widehat{\theta}^l_{B'}) &\geq \underline{\eta}, & (55)\\ \eta(m_{B'}|\widehat{\theta}^l_{B'}) &< \underline{\eta}. & (56) \end{aligned}$$

(see (8)). We consider these constraints in three possible parameter cases according to Lemma 13.

Case 1: $F \geq 0.5\bar{\theta}(P-c)$ and $\widehat{\theta}^l_{B'} = \widehat{\theta}^l_{N'} = \bar{\theta}$. Then, the receiver gets only one message $m_{G'}$ on the equilibrium path, so that her only incentive constraint is

$$\eta(m_{G'}|\widehat{\theta}^l_{B'}) \geq \underline{\eta} \equiv -\frac{c}{P-c}. \tag{57}$$

By (47)

$$\eta(m_{G'}|\widehat{\theta}^l_{B'} = \bar{\theta}) = 0.5 > -\frac{c}{P-c} \tag{58}$$

as far as $P > -c$ by Assumption 1. Hence, the incentive constraint never binds and the lying equilibrium always exists in this case.

Case 2: $F \in [\bar{\theta}\frac{P-c}{4-\kappa}, 0.5\bar{\theta}(P-c))$, $\widehat{\theta}^l_{B'} \in [0.5\bar{\theta}, \bar{\theta})$ and $\widehat{\theta}^l_{N'} = \bar{\theta}$.

Consider the incentive constraint $\eta(m_{G'}|\widehat{\theta}^l_{B'}) \geq \underline{\eta}$. Substituting for $\eta(m_{G'}|\widehat{\theta}^l_{B'})$ we get

$$\eta(m_{G'}|\widehat{\theta}^l_{B'}) = \frac{\bar{\theta}}{(2-\kappa)\bar{\theta} + \kappa\widehat{\theta}^l_{B'}} > \frac{\bar{\theta}}{(2-\kappa)\bar{\theta} + \kappa\bar{\theta}} = 0.5 > -\frac{c}{P-c}.$$

Hence, the incentive constraint never binds.

Besides, since $m_{B'}$ is also used on equilibrium path, the receiver should find it optimal to abstain conditional on $m_{B'}$, that is

$$\eta(m_{B'}|\widehat{\theta}^l_{B'}) < -\frac{c}{P-c}. \tag{59}$$

Since in this subtype of the equilibrium $m_{B'}$ is sent only by those types who have observed the bad state for sure (since $\widehat{\theta}^l_{N'} = \bar{\theta}$), it holds $\eta(m_{B'}|\widehat{\theta}^l_{B'}) = 0$ so that (59) also does not bind. Hence, in Case 2 the unique lying equilibrium always exists.

Case 3: $F < \bar{\theta}\frac{P-c}{4-\kappa}$, $\widehat{\theta}^l_{B'} \in (0, 0.5\bar{\theta})$ and $\widehat{\theta}^l_{N'} = 2\widehat{\theta}^l_{B'}$.

Consider the first incentive constraint $\eta(m_{G'}|\widehat{\theta}^l_{B'}) \geq \underline{\eta}$. We have

$$\eta(m_{G'}|\widehat{\theta}^l_{B'}) = \frac{\kappa\bar{\theta} + 2\widehat{\theta}^l_{B'}(1-\kappa)}{\kappa\bar{\theta} + \widehat{\theta}^l_{B'}(4-3\kappa)}. \tag{60}$$

This implies

$$\frac{d\eta(m_{G'}|\widehat{\theta}^l_{B'})}{d\widehat{\theta}^l_{B'}} = -\frac{(2-\kappa)\kappa\bar{\theta}}{(\kappa(\bar{\theta} - 3\widehat{\theta}^l_{B'}) + 4\widehat{\theta}^l_{B'})^2} < 0. \tag{61}$$

At the same time,

$$\eta(m_{G'}|\widehat{\theta}^l_{B'} = 0.5\bar{\theta}) = \frac{2}{4-\kappa} > 0.5. \tag{62}$$

(61) and (62) imply that for any $\widehat{\theta}^l_{B'} < 0.5\bar{\theta}$ it holds $\eta(m_{G'}|\widehat{\theta}^l_{B'}) > 0.5$ and, hence, $\eta(m_{G'}|\widehat{\theta}^l_{B'}) > \underline{\eta}$.

Consider the second incentive constraint, $\eta(m_{B'}|\widehat{\theta}^l_{B'}) < \underline{\eta}$. We have

$$\eta(m_{B'}|\widehat{\theta}^l_{B'}) = \frac{(1-\kappa)(\bar{\theta} - 2\widehat{\theta}^l_{B'})}{(2-\kappa)\bar{\theta} - (4-3\kappa)\widehat{\theta}^l_{B'}}. \tag{63}$$

Let us first show that

$$\frac{d\eta(m_{B'}|\widehat{\theta}^l_{B'})}{dF} < 0. \tag{64}$$

Since $\eta(m_{B'}|\widehat{\theta}^l_{B'})$ depends on F only through $\widehat{\theta}^l_{B'}$, we have

$$\frac{d\eta(m_{B'}|\widehat{\theta}^l_{B'})}{dF} = \frac{\partial\eta(m_{B'}|\widehat{\theta}^l_{B'})}{\partial\widehat{\theta}^l_{B'}}\frac{d\widehat{\theta}^l_{B'}}{dF}. \tag{65}$$

The first term in the RHS is (differentiating (63))

$$\frac{\partial \eta(m_{B'}|\widehat{\theta}^l_{B'})}{\partial \widehat{\theta}^l_{B'}} = -\frac{\bar{\theta}\kappa(1-\kappa)}{\left(\bar{\theta}(2-\kappa) - \widehat{\theta}^l_{B'}(4-3\kappa)\right)^2} < 0. \tag{66}$$

Consider the second term. By the implicit function theorem and the fact that $\varpi(\widehat{\theta}^l_{B'}) = 0$ (since $\widehat{\theta}^l_{B'}$ is interior by assumption) we have

$$\frac{d\widehat{\theta}^l_{B'}}{dF} = -\frac{\partial\varpi/\partial F}{\partial\varpi/\partial\widehat{\theta}^l_{B'}} = -\frac{1}{\partial\varpi/\partial\widehat{\theta}^l_{B'}} > 0, \tag{67}$$

where the last inequality follows by Lemma 12 and the fact that $\varpi(\cdot)$ is differentiable on $(0, 0.5\bar{\theta})$. Finally, (65)-(67) lead to (64).

Let us consider the limit of $\eta(m_{B'}|\widehat{\theta}^l_{B'})$ as F goes to 0. First, we have

$$\lim_{F\to 0}\widehat{\theta}^l_{B'} = \lim_{F\to 0}\frac{F}{\eta(m_{G'}|\widehat{\theta}^l_{B'})(P-c)} = 0, \tag{68}$$

where the first equality follows from $\varpi(\widehat{\theta}^l_{B'}) = 0$ and the last equality follows from the fact that $\eta(m_{G'}|\widehat{\theta}^l_{B'}) > 0.5$ (as it was shown above for all possible cases), and hence bounded from 0. Consequently,

$$\lim_{F\to 0}\eta(m_{B'}|\widehat{\theta}^l_{B'}) = \lim_{\widehat{\theta}^l_{B'}\to 0}\frac{(1-\kappa)(\bar{\theta} - 2\widehat{\theta}^l_{B'})}{(2-\kappa)\bar{\theta} - (4-3\kappa)\widehat{\theta}^l_{B'}} = \frac{1-\kappa}{2-\kappa}. \tag{69}$$

Expressions (64) and (69) yield that $\frac{1-\kappa}{2-\kappa}$ is the upper bound of $\eta(m_{B'}|\widehat{\theta}^l_{B'})$. At the same time, we have (given that $\kappa \in (0,1]$)

$$\frac{1-\kappa}{2-\kappa} < -\frac{c}{P-c} \Leftrightarrow \kappa > \frac{P+c}{P}. \tag{70}$$

Consequently, for any possible $\kappa > \frac{P+c}{P}$ we have $\eta(m_{B'}|\widehat{\theta}^l_{B'}) < \underline{\eta}$ for any F, so that the receiver's incentive constraints do not bind and the equilibrium always exists.

Consider $\kappa \le \frac{P+c}{P}$. Then,

$$\lim_{F\to 0}\eta(m_{B'}|\widehat{\theta}^l_{B'}) = \frac{1-\kappa}{2-\kappa} \ge \underline{\eta}, \tag{71}$$

where the inequality is by (70). Besides, given that $\widehat{\theta}^l_{B'}$ converges to $0.5\bar{\theta}$ as $F \to \bar{\theta}\frac{P-c}{4-\kappa}$ (see the equality in (54)), it holds

$$\lim_{F\to\bar{\theta}\frac{P-c}{4-\kappa}}\eta(m_{B'}|\widehat{\theta}^l_{B'}) = \lim_{\widehat{\theta}^l_{B'}\to 0.5\bar{\theta}}\frac{(1-\kappa)(\bar{\theta} - 2\widehat{\theta}^l_{B'})}{(2-\kappa)\bar{\theta} - (4-3\kappa)\widehat{\theta}^l_{B'}} = 0 < \underline{\eta}. \tag{72}$$

By (64), (71), (72) and the intermediate value theorem for any $\kappa \leq \frac{P+c}{P}$ there must exist a threshold value $\widetilde{F}^l(\kappa) > 0$ such that for any $\widetilde{F}^l(\kappa) < F < \bar{\theta}\frac{P-c}{4-\kappa}$ it should hold $\eta(m_{B'}|\widehat{\theta}^l_{B'}) < \underline{\eta}$, and $\eta(m_{B'}|\widehat{\theta}^l_{B'}) \geq \underline{\eta}$ otherwise. Hence, the receiver's incentive constraints are violated if and only if $\kappa \leq \frac{P+c}{P}$ and $F \leq \widetilde{F}^l(\kappa)$. ■

Lemma 14 *If the strategies and beliefs are specified according to Definition 2, then the sender has no incentives to deviate if and only if the following two conditions hold:*

1) either $\widehat{\theta}^e_{B'} < \bar{\theta}$ *and* $U^s_{B'}(\widehat{\theta}^e_{B'}, \eta(m_{N'}), I) = 0$ *or* $\widehat{\theta}^e_{B'} = \bar{\theta}$ *and* $U^s_{B'}(\widehat{\theta}^e_{B'}, \eta(m_{N'}), I) \geq 0$.

2) $\widehat{\theta}^e_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^e_{B'}]$.

Proof. The proof is based on exactly the same arguments as the proof of Lemma 11 for the case of the lying equilibrium and, hence, is omitted. ■

Lemma 15 *If the strategies and beliefs are specified as in Definition 2 and* $\widehat{\theta}^e_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^e_{B'}]$*, then* $U^s_{B'}(\widehat{\theta}^e_{B'}, \eta(m_{N'}|\widehat{\theta}^e_{B'}), I)$ *is continuous and strictly decreasing in* $\widehat{\theta}^e_{B'}$ *on* $(0, \bar{\theta}]$.

Proof. As in the case of the lying equilibrium (see (38)-(39)) we have

$$\frac{dU^s_{B'}(\widehat{\theta}^e_{B'}, \eta(m_{N'}|\widehat{\theta}^e_{B'}), I)}{d\widehat{\theta}^e_{B'}} = -(P-c)\widehat{\theta}^e_{B'}\left(\frac{d\eta(m_{N'}|\widehat{\theta}^e_{B'})}{d\widehat{\theta}^e_{B'}} + \frac{\eta(m_{N'}|\widehat{\theta}^e_{B'})}{\widehat{\theta}^e_{B'}}\right). \tag{73}$$

for any $\widehat{\theta}^e_{B'} \in (0, \bar{\theta}]$. Consider the following possible cases given that $\widehat{\theta}^e_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^e_{B'}]$.

Case 1: $\widehat{\theta}^e_{B'} \in (0, 0.5\bar{\theta}]$, $\widehat{\theta}^e_{N'} = 2\widehat{\theta}^e_{B'}$.

Then, by (15) (substituting $\Pr[m_{N'}|G'] = 0$, $\Pr[m_{N'}|N'] = 2\widehat{\theta}^e_{B'}/\bar{\theta}$ and $\Pr[m_{N'}|B'] = \widehat{\theta}^e_{B'}/\bar{\theta}$)

$$\eta(m_{N'}|\widehat{\theta}^e_{B'}) = \frac{2(1-\kappa)}{4-3\kappa}, \tag{74}$$

i.e., is constant. Thus,

$$\frac{d\eta(m_{N'}|\widehat{\theta}^e_{B'})}{d\widehat{\theta}^e_{B'}} = 0 \tag{75}$$

that together with (73) leads to the claim for $\widehat{\theta}^e_{B'} \in (0, 0.5\bar{\theta}]$.

Case 2: $\widehat{\theta}^e_{B'} \in (0.5\bar{\theta}, \bar{\theta}]$, $\widehat{\theta}^e_{N'} = \bar{\theta}$.

In this case, $\Pr[m_{N'}|G'] = 0$, $\Pr[m_{N'}|N'] = 1$ and $\Pr[m_{N'}|B'] = \widehat{\theta}^e_{B'}/\bar{\theta}$ so that (15) implies

$$\eta(m_{N'}|\widehat{\theta}^e_{B'}) = \frac{(1-\kappa)\bar{\theta}}{2(1-\kappa)\bar{\theta} + \kappa\widehat{\theta}^e_{B'}}. \tag{76}$$

Then, it is possible to obtain a simple closed-form solution for the RHS of (73):

$$\begin{aligned}
& \frac{d\eta(m_{N'}|\widehat{\theta}^e_{B'})}{d\widehat{\theta}^e_{B'}} + \frac{\eta(m_{N'}|\widehat{\theta}^e_{B'})}{\widehat{\theta}^e_{B'}} \\
= & -\frac{(1-\kappa)\kappa\bar{\theta}}{(\kappa\widehat{\theta}^e_{B'} + 2(1-\kappa)\bar{\theta})^2} + \frac{(1-\kappa)\bar{\theta}}{\widehat{\theta}^e_{B'}(\kappa\widehat{\theta}^e_{B'} + 2(1-\kappa)\bar{\theta})} \\
= & \frac{2(1-\kappa)^2\bar{\theta}^2}{\widehat{\theta}^e_{B'}(\kappa\widehat{\theta}^e_{B'} + 2(1-\kappa)\bar{\theta})^2} > 0. \qquad (77)
\end{aligned}$$

By (77) and (73) the claim follows (for $\widehat{\theta}^e_{B'} \in (0.5\bar{\theta}, \bar{\theta}]$).

Finally, since $\eta(m_{N'}|\widehat{\theta}^e_{B'})$, and hence $U^s_{B'}(\widehat{\theta}^e_{B'}, \eta(m_{N'}|\widehat{\theta}^e_{B'}), I)$, is continuous at $\widehat{\theta}^e_{B'} = 0.5\bar{\theta}$ the claim holds for the whole interval $(0, \bar{\theta}]$. ■

Lemma 16 *If the strategies and beliefs are specified as in Definition 2, then for any given parameter values there always exist unique cutoffs $\widehat{\theta}^e_{B'}$ and $\widehat{\theta}^e_{N'}$ such that the sender does not have incentives to deviate. Moreover:*

1) If $F \geq \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}$ then $\widehat{\theta}^e_{B'} = \widehat{\theta}^e_{N'} = \bar{\theta}$.

2) If $F \in [\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa})$ then $\widehat{\theta}^e_{B'} \in [0.5\bar{\theta}, \bar{\theta})$ and $\widehat{\theta}^e_{N'} = \bar{\theta}$.

3) If $F < \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$ then $\widehat{\theta}^e_{B'} \in (0, 0.5\bar{\theta})$ and $\widehat{\theta}^e_{N'} = 2\widehat{\theta}^e_{B'}$.

Proof. The proof proceeds analogously to the case of the lying equilibrium. For notational simplicity denote

$$\phi(\widehat{\theta}^e_{B'}) = U^s_{B'}(\widehat{\theta}^e_{B'}, \eta(m_{N'}|\widehat{\theta}^e_{B'}), I) \qquad (78)$$

with $\widehat{\theta}^e_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^l_{B'}]$. Consider the cases listed in the lemma.

Case 1: $F \geq \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}$. Consider the behavior of $\phi(\cdot)$ at the exterior point $\bar{\theta}$. We have

$$\eta(m_{N'}|\widehat{\theta}^e_{B'} = \bar{\theta}) = \frac{1-\kappa}{2-\kappa}, \qquad (79)$$

which results from substituting $\Pr[m_{N'}|G'] = 0$ and $\Pr[m_{G'}|N'] = \Pr[m_{G'}|B'] = 1$ into (15). Then,

$$\phi(\bar{\theta}) = F - \bar{\theta}\eta(m_{N'}|\widehat{\theta}^e_{B'} = \bar{\theta})(P-c) = F - \bar{\theta}\frac{1-\kappa}{2-\kappa}(P-c) \geq 0, \qquad (80)$$

where the second equality follows from (79) and the inequality by the assumption of the case. Then, the sender has no incentives to deviate if $\widehat{\theta}^e_{B'} = \widehat{\theta}^e_{N'} = \bar{\theta}$ by Lemma 14. At the same time, it follows from Lemma 15 and (80) that

$$\forall \widehat{\theta}^e_{B'} < \bar{\theta}, \phi(\widehat{\theta}^e_{B'}) > 0. \qquad (81)$$

Consequently, by Lemma 14 there are no possible cutoffs except for $\widehat{\theta}^e_{B'} = \widehat{\theta}^e_{N'} = \bar{\theta}$ such that the sender does not have incentives to deviate.

Case 2: $F \in [\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa})$. We have

$$\begin{aligned} \phi(0) &= F - 0 \cdot \eta(m_{N'}|\widehat{\theta}^e_{B'} = 0)(P-c) = F > 0, & (82) \\ \phi(\bar{\theta}) &= F - \bar{\theta}\eta(m_{N'}|\widehat{\theta}^e_{B'} = \bar{\theta})(P-c) = F - \bar{\theta}\frac{1-\kappa}{2-\kappa}(P-c) < 0, & (83) \end{aligned}$$

where the second equality follows by (79), and the inequality follows by the assumption of the case. Then, from Lemma 15 and the intermediate value theorem it follows that there exists a unique cutoff value $0 < \widehat{\theta}^e_{B'} < \bar{\theta}$ such that $\phi(\widehat{\theta}^e_{B'}) = 0$ (the necessary and sufficient condition for an interior cutoff by Lemma 14). At the same time, the cutoff $\widehat{\theta}^e_{B'} = \bar{\theta}$ is impossible due to (83) and Lemma 14, so that the existing interior cutoff is the only possible cutoff. Lemma 14 also implies that the corresponding unique cutoff in state N' is then given by $\widehat{\theta}^e_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^l_{B'}]$.

Let us show that in this case $\widehat{\theta}^e_{B'} \geq 0.5\bar{\theta}$. By Lemma 15 it holds

$$\phi(\widehat{\theta}^e_{B'}) = 0 \wedge \widehat{\theta}^e_{B'} \geq 0.5\bar{\theta} \Leftrightarrow \phi(0.5\bar{\theta}) \geq 0. \tag{84}$$

From (15) we get

$$\eta(m_{N'}|\widehat{\theta}^e_{B'} = 0.5\bar{\theta}) = \frac{2(1-\kappa)}{4-3\kappa} \tag{85}$$

so that

$$\phi(0.5\bar{\theta}) = F - 0.5\bar{\theta}\frac{2(1-\kappa)}{4-3\kappa}(P-c) \geq 0, \tag{86}$$

where the inequality follows from the assumption $F \in [\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa})$. By (84) and (86) it then follows that $\widehat{\theta}^e_{B'} \geq 0.5\bar{\theta}$.

Case 3: $F < \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$. From $\phi(0) > 0$, $\phi(\bar{\theta}) < 0$ and Lemma 15 it follows that there is a unique interior cutoff $\widehat{\theta}^e_{B'}$ with $\widehat{\theta}^e_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^l_{B'}]$. Finally, (84), the left equality in (86) and $F < \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$ result in $\widehat{\theta}^e_{B'} < 0.5\bar{\theta}$. ∎

Proof of Proposition 2. As in the case of the lying equilibrium, to show the claim of the proposition we need to find the range of parameters such that the receiver's incentive constraints are satisfied given the unique equilibrium cutoffs $\widehat{\theta}^e_{B'}$ and $\widehat{\theta}^e_{N'}$, which always exist by Lemma 16. Clearly, since message $m_{G'}$ is sent by the sender only if he has indeed observed the good state of the world, it holds $\eta(m_{G'}) = 1 > \underline{\eta}$, so that the receiver always prefers to invest after $m_{G'}$. The remaining incentive constraints ensure investment after $m_{N'}$ and abstaining after $m_{B'}$ (if $m_{B'}$ is sent in equilibrium):

$$\begin{aligned} \eta(m_{N'}|\widehat{\theta}^e_{B'}) &\geq \underline{\eta}, & (87) \\ \eta(m_{B'}|\widehat{\theta}^e_{B'}) &< \underline{\eta}. & (88) \end{aligned}$$

We consider these constraints in three possible parameter cases according to Lemma 16.

Case 1: $F \geq \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}$ and $\widehat{\theta}^e_{B'} = \widehat{\theta}^e_{N'} = \bar{\theta}$.

The only relevant incentive constraint for the receiver is then

$$\eta(m_{N'}|\widehat{\theta}^e_{B'}) \geq -\frac{c}{P-c}, \tag{89}$$

which ensures investment after $m_{N'}$. By (79) we have

$$\eta(m_{N'}|\widehat{\theta}^e_{B'} = \bar{\theta}) = \frac{1-\kappa}{2-\kappa} \tag{90}$$

so that

$$\eta(m_{N'}|\widehat{\theta}^e_{B'} = \bar{\theta}) \geq -\frac{c}{P-c} \Leftrightarrow \kappa \leq \frac{P+c}{P}. \tag{91}$$

Case 2: $F \in [\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa})$, $\widehat{\theta}^e_{B'} \in [0.5\bar{\theta}, \bar{\theta})$ and $\widehat{\theta}^e_{N'} = \bar{\theta}$.

In this case the message $m_{B'}$ is sent only by types in state B' so that $\eta(m_{B'}|\widehat{\theta}^e_{B'}) = 0 < \underline{\eta}$, i.e., the incentive constraint for abstaining after $m_{B'}$ is always satisfied. Consider the remaining incentive constraint for investment after $m_{N'}$

$$\eta(m_{N'}|\widehat{\theta}^e_{B'}) \geq \underline{\eta}. \tag{92}$$

Substituting for $\eta(m_{N'}|\widehat{\theta}^e_{B'})$ given that $\widehat{\theta}^e_{N'} = \bar{\theta}$ we get

$$\eta(m_{N'}|\widehat{\theta}^e_{B'}) = \frac{\bar{\theta}(1-\kappa)}{(1-\kappa)2\bar{\theta} + \kappa\widehat{\theta}^e_{B'}}. \tag{93}$$

We have

$$\frac{d\eta(m_{N'}|\widehat{\theta}^e_{B'})}{dF} = \frac{\partial\eta(m_{N'}|\widehat{\theta}^e_{B'})}{\partial\widehat{\theta}^e_{B'}} \frac{d\widehat{\theta}^e_{B'}}{dF}. \tag{94}$$

The first term in the RHS is

$$\frac{\partial\eta(m_{N'}|\widehat{\theta}^e_{B'})}{\partial\widehat{\theta}^e_{B'}} = -\frac{\bar{\theta}\kappa(1-\kappa)}{\left(2\bar{\theta}(1-\kappa) + \kappa\widehat{\theta}^e_{B'}\right)^2} < 0. \tag{95}$$

Consider the second term. By the implicit function theorem and the fact that $\phi(\widehat{\theta}^e_{B'}) = 0$ (since $\widehat{\theta}^e_{B'}$ is interior by assumption) we have

$$\frac{d\widehat{\theta}^e_{B'}}{dF} = -\frac{\partial\phi/\partial F}{\partial\phi/\partial\widehat{\theta}^e_{B'}} = -\frac{1}{\partial\phi/\partial\widehat{\theta}^e_{B'}} > 0, \tag{96}$$

where the last inequality follows by Lemma 15.[22] Finally, (94)-(96) lead to

$$\frac{d\eta(m_{N'}|\widehat{\theta}^e_{B'})}{dF} < 0. \tag{97}$$

[22] Note that $\phi(\cdot)$ is only right-differentiable at $\widehat{\theta}^e_{B'} = 1/2\bar{\theta}$, which is still sufficient for arguments of the proof to go through.

Then, under considered range of parameters, $\eta(m_{N'}|\widehat{\theta}^e_{B'})$ obtains its highest value if $F = \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$ so that, correspondingly, $\widehat{\theta}^e_{B'} = 0.5\bar{\theta}$ (see (86)). In this case

$$\eta(m_{N'}|\widehat{\theta}^e_{B'} = 0.5\bar{\theta}) = \frac{2(1-\kappa)}{4-3\kappa}. \tag{98}$$

Consequently, if

$$\begin{aligned} \frac{2(1-\kappa)}{4-3\kappa} &< \underline{\eta} \\ &\Leftrightarrow \kappa > \frac{2(P+c)}{2P+c}, \end{aligned} \tag{99}$$

then for any F in the considered case the incentive constraint $\eta(m_{N'}|\widehat{\theta}^e_{B'}) \geq \underline{\eta}$ is violated.

At the same time, $\eta(m_{N'}|\widehat{\theta}^e_{B'})$ obtains its lowest value if F converges to $\frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}$ so that $\widehat{\theta}^e_{B'}$ converges to $\bar{\theta}$ (see (80)). Then,

$$\eta(m_{N'}|\widehat{\theta}^e_{B'} = \bar{\theta}) = \frac{1-\kappa}{2-\kappa}. \tag{100}$$

Consequently, if

$$\begin{aligned} \frac{1-\kappa}{2-\kappa} &\geq \underline{\eta} \\ &\Leftrightarrow \kappa \leq \frac{P+c}{P}, \end{aligned} \tag{101}$$

then for any F in the considered case the incentive constraint $\eta(m_{N'}|\widehat{\theta}^e_{B'}) \geq \underline{\eta}$ is satisfied.

Next, consider the case when $\kappa \in (\frac{P+c}{P}, \frac{2(P+c)}{2P+c}]$. Then,

$$\eta(m_{N'}|\widehat{\theta}^e_{B'} = 0.5\bar{\theta}) = \frac{2(1-\kappa)}{4-3\kappa} \geq \underline{\eta}, \tag{102}$$

$$\eta(m_{N'}|\widehat{\theta}^e_{B'} = \bar{\theta}) = \frac{1-\kappa}{2-\kappa} < \underline{\eta}. \tag{103}$$

This, together with (95) and the intermediate value theorem, implies that there exists a threshold value of $\widehat{\theta}^e_{B'}$, and hence F (by (96)), such that the incentive constraint binds, i.e., $\eta(m_{N'}|\widehat{\theta}^e_{B'}) = \underline{\eta}$. This equality yields

$$\begin{aligned} \eta(m_{N'}|\widehat{\theta}^e_{B'}) &= \frac{\bar{\theta}(1-\kappa)}{(1-\kappa)2\bar{\theta} + \kappa\widehat{\theta}^e_{B'}} = -\frac{c}{P-c}, \\ \widehat{\theta}^e_{B'} &= \frac{(P+c)(1-\kappa)\bar{\theta}}{-c\kappa}. \end{aligned} \tag{104}$$

Substituting this into $\phi(\widehat{\theta}^e_{B'})$ we get

$$\begin{aligned}\phi(\widehat{\theta}^e_{B'}) &= F - \frac{(P+c)(1-\kappa)\bar{\theta}}{-c\kappa}\left(-\frac{c}{P-c}\right)(P-c) \\ &= F - \frac{(P+c)(1-\kappa)\bar{\theta}}{\kappa} = 0, \end{aligned} \tag{105}$$

where the last equality is by the fact that $\phi(\widehat{\theta}^e_{B'}) = 0$ by Lemma 14. This implies that the value of F which leads to $\eta(m_{N'}|\widehat{\theta}^e_{B'}) = \underline{\eta}$ is

$$F_{|\eta(m_{N'}|\widehat{\theta}^e_{B'})=\underline{\eta}} = \frac{(P+c)(1-\kappa)\bar{\theta}}{\kappa}. \tag{106}$$

Thus, in Case 2 the receiver's incentive constraints are satisfied whenever $\kappa \leq \frac{P+c}{P}$ for any F or $\kappa \in (\frac{P+c}{P}, \frac{2(P+c)}{2P+c}]$ and $F \leq \frac{(P+c)(1-\kappa)\bar{\theta}}{\kappa}$.

Case 3: $F < \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$, $\widehat{\theta}^e_{B'} \in (0, 0.5\bar{\theta})$ and $\widehat{\theta}^e_{N'} = 2\widehat{\theta}^e_{B'}$.

Substituting $\Pr[m_{N'}|G'] = 0$, $\Pr[m_{N'}|N'] = 2\widehat{\theta}^e_{B'}/\bar{\theta}$ and $\Pr[m_{N'}|B'] = \widehat{\theta}^e_{B'}/\bar{\theta}$ into (15) we get

$$\eta(m_{N'}|\widehat{\theta}^e_{B'}) = \frac{2(1-\kappa)}{4-3\kappa}. \tag{107}$$

Then,

$$\begin{aligned}\frac{2(1-\kappa)}{4-3\kappa} &\geq \underline{\eta} = -\frac{c}{P-c} \\ &\Leftrightarrow 0 < \kappa \leq \frac{2(P+c)}{2P+c}, \end{aligned} \tag{108}$$

determining the parameter range where the incentive constraint (87) holds.

Consider the second incentive constraint (88). We have

$$\eta(m_{B'}|\widehat{\theta}^e_{B'}) = \frac{(1-\kappa)(\bar{\theta} - 2\widehat{\theta}^e_{B'})}{(2-\kappa)\bar{\theta} - (4-3\kappa)\widehat{\theta}^e_{B'}}, \tag{109}$$

which is equivalent to the previously analyzed expression (63). Hence, by the same arguments as before we have

$$\frac{d\eta(m_{B'}|\widehat{\theta}^e_{B'})}{dF} < 0. \tag{110}$$

Besides, from $\phi(\widehat{\theta}^e_{B'}) = 0$ it follows

$$\lim_{F\to 0} \widehat{\theta}^e_{B'} = \lim_{F\to 0} \frac{F}{\eta(m_{N'}|\widehat{\theta}^e_{B'})(P-c)} = 0, \tag{111}$$

where the last equality holds since $\eta(m_{N'}|\widehat{\theta}^e_{B'})$ is constant by (107). Then, by the analogous arguments as in the analysis of Case 3 in the proof of Proposition 1, $\eta(m_{B'}) < \underline{\eta}$ is satisfied whenever $\kappa > \frac{P+c}{P}$, or $\kappa \leq \frac{P+c}{P}$ and $\widetilde{F}^e(\kappa) < F < \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$ where $\widetilde{F}^e(\kappa) > 0$ is some threshold value.

Merging Case 1, Case 2 and Case 3 together, we obtain that the evasion equilibrium exists only in the following cases:

Case 1:

$$F \geq \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa} \wedge \kappa \leq \frac{P+c}{P};$$

Case 2:

$$\left(F \in \left[\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}\right) \wedge \kappa \leq \frac{P+c}{P}\right) \text{ or}$$
$$\left(F \in \left[\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}, \frac{(P+c)(1-\kappa)\bar{\theta}}{\kappa}\right] \wedge \kappa \in \left(\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right]\right);$$

Case 3:

$$\left(F \in \left(\widetilde{F}^e(\kappa), \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}\right) \wedge \kappa \leq \frac{P+c}{P}\right) \text{ or}$$
$$\left(F \in \left(0, \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}\right) \wedge \kappa \in \left(\frac{P+c}{P}, \frac{2(P+c)}{2P+c}\right]\right).$$

This is equivalent to the statement of the proposition. ■

Proof of Lemma 5. Given that $\widehat{\theta}^l_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^l_{B'}]$ and $\widehat{\theta}^e_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^e_{B'}]$ (by Lemmas 11 and 14), it is sufficient to show the claim only for $i^s = B'$. We have that for any $\widehat{\theta}^l_{B'}$

$$\eta(m_{G'}|\widehat{\theta}^l_{B'}) \geq 0.5 \tag{112}$$

(see the proof of Proposition 1). At the same time,

$$\begin{aligned} \eta(m_{N'}|\widehat{\theta}^e_{B'}) &= \Pr[G|m_{N'}] = \Pr[G|N' \cap m_{N'}]\Pr[N'|m_{N'}] + \Pr[G|B' \cap m_{N'}]\Pr[B'|m_{N'}] \\ &= 0.5\Pr[N'|m_{N'}] < 0.5, \end{aligned} \tag{113}$$

because at least some types in state B' send $m_{N'}$ so that $\Pr[N'|m_{N'}] < 1$. (112) and (113) yield

$$\eta(m_{G'}|\widehat{\theta}^l_{B'}) > \eta(m_{N'}|\widehat{\theta}^e_{B'}) \tag{114}$$

so that

$$U^s_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{N'}|\widehat{\theta}^e_{B'}), I) > U^s_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{G'}|\widehat{\theta}^l_{B'}), I). \tag{115}$$

If $\widehat{\theta}^l_{B'} < \bar{\theta}$ so that $U^s_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{G'}|\widehat{\theta}^l_{B'}), I) = 0$ by Lemma 11, then (115) implies

$$U^s_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{N'}|\widehat{\theta}^e_{B'}), I) > 0. \tag{116}$$

Then, $\widehat{\theta}^e_{B'} > \widehat{\theta}^l_{B'}$ by Lemma 14 and the fact that $U^s_{B'}(\theta, \eta(m_{N'}|\widehat{\theta}^e_{B'}), I)$ is decreasing in the first argument.

In the other case, if $\widehat{\theta}^l_{B'} = \bar{\theta}$, by (115) and Lemma 11

$$U^s_{B'}(\bar{\theta}, \eta(m_{N'}|\widehat{\theta}^e_{B'}), I) > 0. \tag{117}$$

Consequently, all types in state B' prefer to send $m_{N'}$ in the evasion equilibrium and the only possible equilibrium cutoff is $\widehat{\theta}^e_{B'} = \bar{\theta}$. ■

Lemma 17 *If the strategies and beliefs are specified according to Definition 3, then the sender has no incentives to deviate if and only if the following two conditions hold:*

1) either $\widehat{\theta}^h_{B'} < \bar{\theta}$ and $U^s_{B'}(\widehat{\theta}^h_{B'}, \eta(m_{N'}), I) = 0$ or $\widehat{\theta}^h_{B'} = \bar{\theta}$ and $U^s_{B'}(\widehat{\theta}^h_{B'}, \eta(m_{N'}), I) \geq 0$.

2) $\widehat{\theta}^h_{N'} = \min[\bar{\theta}, 2\widehat{\theta}^h_{B'}]$.

Proof. The proof is based on exactly the same arguments as the proof of Lemma 11 for the case of the lying equilibrium and, hence, is omitted. ■

Lemma 18 *If the strategies and beliefs are specified according to Definition 3, then for any possible cutoff value $z \in [\widehat{\theta}^l_{B'}, \widehat{\theta}^e_{B'}]$ (where $\widehat{\theta}^l_{B'}$ and $\widehat{\theta}^e_{B'}$ are the unique cutoffs in the sense of Lemmas 13 and 16) there exists at least one γ such that the sender has no incentive to deviate given $\widehat{\theta}^h_{B'} = z$. There exists no hybrid equilibrium with $\widehat{\theta}^h_{B'} \notin [\widehat{\theta}^l_{B'}, \widehat{\theta}^e_{B'}]$.*

Proof. Denote the sender's utility and the persuasiveness of message m as $U^{s,l}$ and $\eta^l(m)$ in the lying equilibrium, as $U^{s,e}$ and $\eta^e(m)$ in the evasion equilibrium, and as $U^{s,h}$ and $\eta^h(m)$ in the hybrid equilibrium. For notational simplicity, let us further suppress the receiver's action $x \in \{I, A\}$ in the sender's utility function (it is I in all expressions in the proof below). Besides, let us denote the persuasiveness of the message $m_{N'}$ in the hybrid equilibrium given the cutoff $\widehat{\theta}^h_{B'}$ and γ as $\eta^h(m_{N'}|\widehat{\theta}^h_{B'}, \gamma)$.

First, the claim holds for $z = \widehat{\theta}^l_{B'}$ $(\widehat{\theta}^e_{B'})$, since then the sender has no incentive to deviate if $\widehat{\theta}^h_{B'} = z$ at least for $\gamma = 1(0)$ by Lemma 13 (16).

Consider the case when $z \in (\widehat{\theta}^l_{B'}, \widehat{\theta}^e_{B'})$. Let us first show that for given cutoff $\widehat{\theta}^h_{B'}$ the sender's utility in the hybrid equilibrium $U^{s,h}_{B'}(\widehat{\theta}^h_{B'}, \gamma)$ is continuous and strictly decreasing in γ. We have

$$\begin{aligned}
\frac{\partial U^{s,h}_{B'}(\widehat{\theta}^h_{B'}, \eta^h(m_{N'}|\widehat{\theta}^h_{B'}, \gamma))}{\partial \gamma} &= \frac{\partial U^{s,h}_{B'}}{\partial \eta^h(m_{N'}|\widehat{\theta}^h_{B'}, \gamma)} \frac{\partial \eta^h(m_{N'}|\widehat{\theta}^h_{B'}, \gamma)}{\partial \gamma} \\
&= -\widehat{\theta}^h_{B'}(B - c)\frac{\kappa\bar{\theta}((1-\kappa)\widehat{\theta}^l_{N'} + \kappa\widehat{\theta}^l_{B'})}{(2(1-\kappa)\widehat{\theta}^l_{N'} + \kappa(\gamma\bar{\theta} + \widehat{\theta}^l_{B'}))^2} < 0.
\end{aligned} \tag{118}$$

Further, consider two possible cases depending on whether $\widehat{\theta}^l_{B'} = \bar{\theta}$.

Case 1. $\widehat{\theta}^l_{B'} < \bar{\theta}$. Then, by Lemma 11

$$U^{s,l}_{B'}(\widehat{\theta}^l_{B'}, \eta^l(m_{G'}|\widehat{\theta}^l_{B'})) = 0. \tag{119}$$

At the same time, it is easy to check that for any z it holds $\eta^h(m_{N'}|z,1) = \eta^l(m_{G'}|z)$, and hence

$$U^{s,h}_{B'}(z, \eta^h(m_{N'}|z,1)) = U^{s,l}_{B'}(z, \eta(m_{G'}|z)). \tag{120}$$

Lemma 12, (119) and (120) imply that for any $z > \widehat{\theta}^l_{B'}$ it holds

$$U^{s,h}_{B'}(z, \eta^h(m_{N'}|z,1)) < 0. \tag{121}$$

Analogously,

$$U^{s,h}_{B'}(z, \eta^h(m_{N'}|z,0)) = U^{s,e}_{B'}(z, \eta(m_{N'}|z)). \tag{122}$$

From Lemma 14 it follows

$$U^{s,e}_{B'}(\widehat{\theta}^e_{B'}, \eta(m_{N'}|\widehat{\theta}^e_{B'})) \geq 0. \tag{123}$$

This, together with Lemma 15 and (122) yields that for any $z < \widehat{\theta}^e_{B'}$ it holds

$$U^{s,h}_{B'}(z, \eta^h(m_{N'}|z,0)) > 0. \tag{124}$$

Finally, from (118), (121), (124) and the intermediate value theorem it follows that for any value $z \in (\widehat{\theta}^l_{B'}, \widehat{\theta}^e_{B'})$ there exists a unique value of $\gamma \in (0,1)$ such that

$$U^{s,h}_{B'}(z, \eta^h(m_{N'}|z,\gamma)) = 0, \tag{125}$$

which is a necessary and sufficient condition for the sender not having incentive to deviate if $\widehat{\theta}^h_{B'} = z$ by Lemma 17 (given that $z < \widehat{\theta}^e_{B'} \leq \bar{\theta}$ by assumption).

At the same time, for any $z < \widehat{\theta}^l_{B'}$ it holds (due to Lemma 12, (119) and (120))

$$U^{s,h}_{B'}(z, \eta^h(m_{N'}|z,1)) > 0. \tag{126}$$

This together with (118) implies that for all $\gamma < 1$ this inequality holds as well. Consequently, since the necessary condition for the equilibrium cutoff, if $\widehat{\theta}^h_{B'} < \widehat{\theta}^l_{B'}$, is

$$U^s_{B'}(\widehat{\theta}^h_{B'}, \eta^h(m_{N'}|\widehat{\theta}^h_{B'}, \gamma)) = 0$$

(by Lemma 17), we have that no hybrid equilibrium is characterized by a cutoff below $\widehat{\theta}^l_{B'}$. By analogous arguments, for any $z > \widehat{\theta}^e_{B'}$ it holds

$$U^{s,h}_{B'}(z, \eta^h(m_{N'}|z,\gamma)) < 0 \tag{127}$$

for any γ, hence no hybrid equilibrium is characterized by a cutoff above $\widehat{\theta}^e_{B'}$.

This completes the proof for Case 1.

Case 2. $\widehat{\theta}^l_{B'} = \bar{\theta}$. Here, (since $(\widehat{\theta}^l_{B'}, \widehat{\theta}^e_{B'})$ is empty) we only need to show that there exists no hybrid equilibria with $\widehat{\theta}^h_{B'} < \bar{\theta}$. By Lemma 11

$$U^{s,l}_{B'}(\widehat{\theta}^l_{B'}, \eta(m_{G'}|\widehat{\theta}^l_{B'})) \geq 0. \tag{128}$$

This together with Lemma 12 implies that $U^{s,l}_{B'}(z,\eta(m_{G'}|z)) > 0$ for any $z < \widehat{\theta}^l_{B'}$. Then, (120) yields that for any $z < \widehat{\theta}^l_{B'}$

$$U^{s,h}_{B'}(z,\eta^h(m_{N'}|z,1)) > 0. \tag{129}$$

Finally, by (118) and (129) there is no $\gamma < 1$ such that $U^{s,h}_{B'}(z,\eta^h(m_{N'}|z,\gamma)) = 0$ (for any $z < \widehat{\theta}^l_{B'}$). Consequently (by Lemma 17), there is no hybrid equilibrium with $\widehat{\theta}^h_{B'} < \bar{\theta}$ in Case 2. This completes the proof. ■

Proof of Proposition 3. While $\widehat{\theta}_{N'} = \min[\bar{\theta}, 2\widehat{\theta}_{B'}]$ in any equilibrium (by Lemma 10), it is sufficient to show the claim only for state B'. By Lemma 18 there exists no hybrid equilibrium if $\widehat{\theta}^h_{B'} \notin [\widehat{\theta}^l_{B'}, \widehat{\theta}^e_{B'}]$, which proves the second part of the proposition. By the same lemma for any possible $\widehat{\theta}^h_{B'} \in [\widehat{\theta}^l_{B'}, \widehat{\theta}^e_{B'}]$ there exists at least one γ such that the sender's incentives are consistent with his strategies. It is left to show that in this case the receiver's incentive constraints are satisfied as well (given that both lying and evasion equilibria exist).

Let us use the same notation as in the proof of Lemma 18. The receiver's incentive constraints are $\eta^h(m_{G'}) \geq \underline{\eta}$, $\eta^h(m_{N'}) \geq \underline{\eta}$ and $\eta^h(m_{B'}) < \underline{\eta}$. The constraint $\eta^h(m_{G'}) \geq \underline{\eta}$ is trivially satisfied, since $m_{G'}$ is sent only by types in state G'. Let us consider the second constraint

$$\eta^h(m_{N'}) \geq \underline{\eta}. \tag{130}$$

Let us rewrite (15) in the following way:

$$\eta(m) = \frac{p_{G'}(m)\kappa + (1-\kappa)}{(p_{G'}(m) + p_{B'}(m))\kappa + 2(1-\kappa)}, \tag{131}$$

where $p_{G'}(m) = \frac{\Pr[m|G']}{\Pr[m|N']}$ and $p_{B'}(m) = \frac{\Pr[m|B']}{\Pr[m|N']}$ (the transformation is valid as far as $\Pr[m|N'] > 0$). Then,

$$\frac{\partial \eta(m)}{\partial p_{G'}(m)} = \frac{\kappa(1-\kappa+\kappa p_{B'}(m)))}{(2(1-\kappa)+\kappa(p_{B'}(m)+p_{G'}(m)))^2} > 0, \tag{132}$$

$$\frac{\partial \eta(m)}{\partial p_{B'}(m)} = -\frac{\kappa(1-\kappa+\kappa p_{G'}(m)))}{(2(1-\kappa)+\kappa(p_{B'}(m)+p_{G'}(m)))^2} < 0. \tag{133}$$

Let us now compare $\eta^h(m_{N'})$ and $\eta^e(m_{N'})$ based on (131), (132) and (133) (which we can apply due to $\Pr[m_{N'}|N'] > 0$ in both equilibria). Applying the same notation for upper indexes denoting equilibrium type, we have

$$p^h_{G'}(m_{N'}) = \frac{\gamma}{\widehat{\theta}^e_{N'}/\bar{\theta}} \geq 0 = p^e_{G'}(m_{N'}). \tag{134}$$

Besides,

$$p^h_{B'}(m_{N'}) = \frac{\widehat{\theta}^h_{B'}/\bar{\theta}}{\widehat{\theta}^h_{N'}/\bar{\theta}} = \frac{\widehat{\theta}^h_{B'}}{\min[\bar{\theta}, 2\widehat{\theta}^h_{B'}]} \leq \frac{\widehat{\theta}^e_{B'}}{\min[\bar{\theta}, 2\widehat{\theta}^e_{B'}]} = p^e_{B'}(m_{N'}), \tag{135}$$

where the second equality follows by Lemma 17, the inequality follows by assumption, and the last equality by Lemma 14. Altogether, (131)-(135) yield

$$\eta^h(m_{N'}) \geq \eta^e(m_{N'}), \tag{136}$$

which together with the fact that $\eta^e(m_{N'}) \geq \underline{\eta}$ (since the evasion equilibrium exists by initial conditions) implies (130).

Consider the last incentive constraint

$$\eta^h(m_{B'}) < \underline{\eta}. \tag{137}$$

If $\widehat{\theta}^h_{N'} = \bar{\theta}$, then $m_{B'}$ is sent only by types in state B' and the constraint is clearly satisfied. Consider $\widehat{\theta}^h_{N'} < \bar{\theta}$. Then, $\widehat{\theta}^l_{N'} < \bar{\theta}$ as well since $\widehat{\theta}^h_{N'} \geq \widehat{\theta}^l_{N'}$ by assumption. Hence, $\Pr[m_{B'}|N'] > 0$ in both equilibria and we can apply transformation (131). We have

$$p^h_{G'}(m_{B'}) = 0 = p^l_{G'}(m_{B'}). \tag{138}$$

At the same time,

$$p^h_{B'}(m_{B'}) = \frac{\bar{\theta} - \widehat{\theta}^h_{B'}}{\bar{\theta} - 2\widehat{\theta}^h_{B'}} \geq \frac{\bar{\theta} - \widehat{\theta}^l_{B'}}{\bar{\theta} - 2\widehat{\theta}^l_{B'}} = p^l_{B'}(m_{B'}), \tag{139}$$

where the equalities follow from Lemmas 17 and 11, while the inequality follows from $\widehat{\theta}^h_{B'} > \widehat{\theta}^l_{B'}$ (by initial conditions). Since $\eta(m_{B'})$ is decreasing in $p^h_{B'}(m_{B'})$ by (133), it follows that

$$\eta^h(m_{B'}) \leq \eta^l(m_{B'}). \tag{140}$$

This, together with $\eta^l(m_{B'}) < \underline{\eta}$ (since the lying equilibrium exists by initial conditions), leads to (137).

Thus, all receiver's constraints in any hybrid equilibrium are satisfied whenever both lying and evasion equilibria exist, which completes the proof. ■

Lemma 19 *If any two equilibria are characterized by the same persuasiveness in states $i^s \in \{B', N'\}$ (in the sense of Lemma 7), then these equilibria are outcome-equivalent.*

Proof. Denote by $\widehat{\eta}$ the persuasiveness of messages used in states $i^s \in \{B', N'\}$ (unique by Lemma 7). We need to show that, in the case named in the lemma, all sender types in all states and the receiver have the same utility for a given realization of parameters in both equilibria.

First, once $\widehat{\eta}$ coincide in both equilibria, the cutoffs must also be the same. Indeed, assume the opposite by contradiction. Then, given that $\widehat{\theta}_{N'}$ is uniquely determined by $\widehat{\theta}_{B'}$ by Lemma 10, at least in one equilibrium it must hold $\widehat{\theta}_{B'} < \bar{\theta}$ (denote this cutoff as $\widehat{\theta}_{B'1}$). Then, by Lemma 9 it should hold

$$U^s_{B'}(\widehat{\theta}_{B'1}, \widehat{\eta}, I) = 0,$$

which gives

$$\forall\theta < \widehat{\theta}_{B'1}, U^s_{B'}(\theta,\widehat{\eta},I) > 0,$$
$$\forall\theta > \widehat{\theta}_{B'1}, U^s_{B'}(\theta,\widehat{\eta},I) < 0.$$

Then, by Lemma 9 no equilibrium can have a cutoff in state B' different from $\widehat{\theta}_{B'1}$ (given $\widehat{\eta}$), which yields a contradiction.

Next, consider the outcome-equivalence for the receiver. We have that the receiver's utility for given sender type θ and information state i^s depends only on the receiver's action and the state of the world (being 0 after abstaining, and $P(c)$ after investment in the good (bad) state of the world). Since the correspondence of receiver's actions to sender types in each information state i^s is uniquely determined by the cutoff in this state, while the distribution of states of the world is uniquely determined by i^s, it holds that any two equilibria with the same $\widehat{\eta}$ (and hence the same cutoffs) are outcome-equivalent in terms of the receiver's utility.

By the same argument, the monetary payoff of the sender for given sender type θ and information state i^s is also the same in both equilibria (since this payoff depends only on the receiver's action conditional on the sender's type and state i^s).

Finally, the equivalence of guilt (i.e., the term $-\theta D^r(m)$, see (4)) for types in states B' and N' follows from the equality of $\widehat{\eta}$. The types in state G' have zero guilt in any equilibrium. This completes the proof. ■

Proof of Proposition 4. By Lemma 4 every existing equilibrium must have a cutoff structure. Let us assume by contradiction that there exists an equilibrium ζ with some cutoffs $\widehat{\theta}'_{B'}$ and $\widehat{\theta}'_{N'}$, which is not outcome-equivalent to any existing hybrid equilibrium.

Let us denote by Υ'_σ the set of messages which are sent only by types in state G' in equilibrium ζ. Denote

$$\gamma' = \Pr[m \notin \Upsilon'_\sigma | G'],$$

i.e., the probability that types in state G' pool with types in other states.

Let us consider a hybrid equilibrium with $\widehat{\theta}^h_{B'} = \widehat{\theta}'_{B'}$, $\widehat{\theta}^h_{N'} = \widehat{\theta}'_{N'}$ and $\gamma = \gamma'$. Let us show that such equilibrium exists. Denote by $\widehat{\eta}^h$ the persuasiveness of $m_{N'}$ in the constructed hybrid equilibrium, and by $\widehat{\eta}'$ the persuasiveness of investment-inducing messages used in states B' and N' in equilibrium ζ (which does not differ between the messages by Lemma 7).

Claim 1. $\widehat{\eta}^h = \widehat{\eta}'$.

Proof. Let us denote the set of investment-inducing messages in equilibrium ζ except for separating messages in set Υ'_σ as Υ'_π. By Bayes rule and the law of total probability we obtain

$$
\begin{aligned}
\Pr[G|m \in \Upsilon'_\pi] &= \frac{\Pr[m \in \Upsilon'_\pi|G]\Pr[G]}{\Pr[m]} \\
&= \frac{(\Pr[m \in \Upsilon'_\pi|G' \cap G]\Pr[G'|G] + \Pr[m \in \Upsilon'_\pi|N' \cap G]\Pr[N'|G])\Pr[G]}{\Pr[m \in \Upsilon'_\pi|G']\Pr[G'] + \Pr[m \in \Upsilon'_\pi|N']\Pr[N'] + \Pr[m \in \Upsilon'_\pi|B']\Pr[B']} \\
&= \frac{\Pr[m \in \Upsilon'_\pi|G']\kappa + \Pr[m \in \Upsilon'_\pi|N'](1-\kappa)}{(\Pr[m \in \Upsilon'_\pi|G'] + \Pr[m \in \Upsilon'_\pi|B'])\kappa + 2\Pr[m \in \Upsilon'_\pi|N'](1-\kappa)} \\
&= \frac{\gamma'\kappa + \widehat{\theta}'_{N'}/\bar{\theta}(1-\kappa)}{(\gamma' + \widehat{\theta}'_{B'}/\bar{\theta})\kappa + 2\widehat{\theta}'_{N'}/\bar{\theta}(1-\kappa)}, \qquad (141)
\end{aligned}
$$

where the third equality follows from the fact that the sender cannot condition his messaging strategy on the true state of the world directly (but only on his information state). At the same time,

$$
\begin{aligned}
\Pr[G|m \in \Upsilon'_\pi] &= \frac{\Pr[m \in \Upsilon'_\pi|G]\Pr[G]}{\Pr[m \in \Upsilon'_\pi]} \\
&= \frac{\left(\sum_{m\in\Upsilon'_\pi}\Pr[m|G]\right)\Pr[G]}{\sum_{m\in\Upsilon'_\pi}\Pr[m]} \\
&= \frac{\sum_{m\in\Upsilon'_\pi}\frac{\Pr[G|m]\Pr[m]}{\Pr[G]}\Pr[G]}{\sum_{m\in\Upsilon'_\pi}\Pr[m]} \\
&= \frac{\sum_{m\in\Upsilon'_\pi}\widehat{\eta}'\Pr[m]}{\sum_{m\in\Upsilon'_\pi}\Pr[m]} = \widehat{\eta}', \qquad (142)
\end{aligned}
$$

where the second equality follows from the fact that messages are disjoint events, the third equality follows by Bayes rule, and the fourth equality follows from the fact that all messages in Υ'_π have the same persuasiveness $\widehat{\eta}'$. Equations (141) and (142) yield

$$
\widehat{\eta}' = \frac{\gamma'\kappa + \widehat{\theta}'_{N'}/\bar{\theta}(1-\kappa)}{(\gamma' + \widehat{\theta}'_{B'}/\bar{\theta})\kappa + 2\widehat{\theta}'_{N'}/\bar{\theta}(1-\kappa)}. \qquad (143)
$$

At the same time, by (15)

$$
\widehat{\eta}^h = \frac{\gamma\kappa + \widehat{\theta}^h_{N'}/\bar{\theta}(1-\kappa)}{(\gamma + \widehat{\theta}^h_{B'}/\bar{\theta})\kappa + 2\widehat{\theta}^h_{N'}/\bar{\theta}(1-\kappa)}. \qquad (144)
$$

Given that $\widehat{\theta}^h_{i^s} = \widehat{\theta}'_{i^s}$ for $i^s \in \{B', N'\}$ and $\gamma = \gamma'$ by construction, (143) and (144) imply

$$
\widehat{\eta}^h = \widehat{\eta}'. \qquad (145)
$$

Claim 2. *The constructed hybrid equilibrium is indeed an equilibrium.*

Proof. First, let us show that in the constructed hybrid equilibrium no sender type has a strict incentive to deviate (as far as the receiver plays as prescribed by the equilibrium). Since equilibrium ζ exists by assumption, by Lemma 9 for any $i^s \in \{B', N'\}$ it must hold $U^s_{i^s}(\widehat{\theta}'_{i^s}, \widehat{\eta}', I) = 0$ if $\widehat{\theta}'_{i^s} < \bar{\theta}$ and $U^s_{i^s}(\widehat{\theta}'_{i^s}, \widehat{\eta}', I) \geq 0$ if $\widehat{\theta}'_{i^s} = \bar{\theta}$. Given that $\widehat{\theta}^h_{i^s} = \widehat{\theta}'_{i^s}$ and $\gamma = \gamma'$ by construction and $\widehat{\eta}^h = \widehat{\eta}'$ by Claim 1, the same must holds for the constructed hybrid equilibrium: $U^s_{i^s}(\widehat{\theta}^h_{i^s}, \widehat{\eta}^h, I) = 0$ if $\widehat{\theta}^h_{i^s} < \bar{\theta}$ and $U^s_{i^s}(\widehat{\theta}^h_{i^s}, \widehat{\eta}^h, I) \geq 0$ if $\widehat{\theta}^h_{i^s} = \bar{\theta}$. It then follows by Lemma 17 that no sender type has an incentive to deviate.

Next, consider the receiver's incentives in the constructed hybrid equilibrium. The receiver does not have incentives to deviate if and only if $\eta^h(m_{G'}) \geq \underline{\eta}$, $\eta^h(m_{N'}) \geq \underline{\eta}$ and $\eta^h(m_{B'}) < \underline{\eta}$. The first constraint is trivially satisfied. For the second constraint we have

$$\eta^h(m_{N'}) \equiv \widehat{\eta}^h = \widehat{\eta}' \geq \underline{\eta}, \tag{146}$$

where the second equality is by Claim 1 and the inequality is by the fact that equilibrium ζ exists. Finally, let us prove that $\eta^h(m_{B'}) < \underline{\eta}$. Assume by contradiction that $\eta^h(m_{B'}) \geq \underline{\eta}$. By (15) we have

$$\eta^h(m_{B'}) = \frac{(\bar{\theta} - \widehat{\theta}^h_{N'})(1-\kappa)}{(\bar{\theta} - \widehat{\theta}^h_{B'})\kappa + 2(\bar{\theta} - \widehat{\theta}^h_{N'})(1-\kappa)}. \tag{147}$$

Consider equilibrium ζ. Denote the set of all messages leading to abstaining as Υ'_α. Applying the same transformations as in (141) we obtain

$$\Pr[G|m \in \Upsilon'_\alpha] = \frac{(\bar{\theta} - \widehat{\theta}'_{N'})(1-\kappa)}{(\bar{\theta} - \widehat{\theta}'_{B'})\kappa + 2(\bar{\theta} - \widehat{\theta}'_{N'})(1-\kappa)}, \tag{148}$$

which together with (147), $\widehat{\theta}^h_{i^s} = \widehat{\theta}'_{i^s}$ and $\gamma = \gamma'$ yields

$$\Pr[G|m \in \Upsilon'_\alpha] = \eta^h(m_{B'}) \geq \underline{\eta}, \tag{149}$$

where the inequality is by assumption. At the same time, applying analogous steps as in (142) we get

$$\begin{aligned}
\Pr[G|m \in \Upsilon'_\alpha] &= \frac{\Pr[m \in \Upsilon'_\alpha|G]\Pr[G]}{\Pr[m \in \Upsilon'_\alpha]} \\
&= \frac{\left(\sum_{m\in\Upsilon'_\alpha} \Pr[m|G]\right)\Pr[G]}{\sum_{m\in\Upsilon'_\alpha} \Pr[m]} \\
&= \frac{\sum_{m\in\Upsilon'_\alpha} \eta(m)\Pr[m]}{\sum_{m\in\Upsilon'_\alpha} \Pr[m]} < \frac{\sum_{m\in\Upsilon'_\alpha} \underline{\eta}\Pr[m]}{\sum_{m\in\Upsilon'_\alpha} \Pr[m]} = \underline{\eta},
\end{aligned} \tag{150}$$

where the inequality follows from the fact that ζ is an equilibrium so that the receiver's incentive constraints are satisfied for each m (in particular, $\eta(m) < \underline{\eta}$ for any $m \in \Upsilon'_\alpha$). Since (149) and (150) yield a contradiction, it follows that $\eta^h(m_{B'}) < \underline{\eta}$, so that all receiver's incentive constraints in the constructed hybrid equilibrium are satisfied. Thus, Claim 2 holds.

Claim 3. *The constructed hybrid equilibrium is outcome-equivalent to equilibrium ζ.*

Proof. The claim follows from Lemma 19 and Claim 1.

Claims 2 and 3 imply together that for any existing equilibrium there exists an outcome-equivalent hybrid equilibrium. ■

Proof of Proposition 5. Assumption 3 rules out equilibria where in some information state more than one investment-inducing message is used. Indeed, assume the opposite so that in some state i^s two message m_1 and m_2 induce investment in equilibrium. Then, by Lemma 7 $\eta(m_1) = \eta(m_2)$, and hence for any type θ

$$U_{i^s}(\theta, m_1) = U_{i^s}(\theta, m_2).$$

Consequently, by Assumption 3 all types in i^s should strictly prefer the same message out of $\{m_1, m_2\}$, which is a contradiction.

Next, denote the (unique) investment inducing message in state i^s as χ_{i^s}. We are left with four possible cases:

1. $\chi_{B'} = \chi_{N'} = \chi_{G'}$.
2. $\chi_{B'} = \chi_{N'} \neq \chi_{G'}$.
3. $\chi_{B'} = \chi_{G'} \neq \chi_{N'}$.
4. $\chi_{B'} \neq \chi_{N'} = \chi_{G'}$.
5. $\chi_{B'} \neq \chi_{N'} \neq \chi_{G'}$.

Cases 4 and 5 are obviously ruled out, since then the receiver would never invest after $\chi_{B'}$, which is either perfectly informative about B' or coincides with one of the messages leading to abstaining. Given Lemmas 3 and 4, Cases 1 and 2 correspond to the lying and the evasion equilibria, i.e. satisfy Definition 1 or 2, respectively. These equilibria do not contradict Assumption 3, since then no types (except for the atomic cutoff types) get the same utility from any two equilibrium messages, so that Assumption 3 does not restrict the sender's incentives in equilibrium.

Finally, Case 3 is only possible when $\eta(\chi_{G'}) = \eta(\chi_{N'}) = 0.5$, where the first equality is by Lemma 7 (noting that $\chi_{G'} = \chi_{B'}$ by assumption, i.e., $\chi_{G'}$ is sent also in B'), and the second equality is from the fact that $\chi_{N'}$ is then perfectly informative about N'. Given that all types in state G' send $\chi_{G'}$ by Lemma 3, it is possible that $\eta(\chi_{G'}) = 0.5$ only if all types in state B' also induce investment with this message (i.e., $\theta^*_{B'} = \bar{\theta}$). Then, by Lemma 8 all types in N' also induce investment. Consequently, this equilibrium is outcome-equivalent to the (lying or evasion) equilibrium of Subtype 1. ■

Proof of Proposition 6. Generally, the ex ante receiver's utility given any equilibrium cutoffs $\widehat{\theta}_{B'}$ and $\widehat{\theta}_{N'}$ is

$$\begin{aligned}
E[U^r] &= \Pr[I \cup G]P + \Pr[I \cup B]c \\
&= \left(\sum_{i^s \in \{G', N', B'\}} \Pr[I | i^s \cap G] \Pr[i^s | G] \right) \Pr[G] \cdot P \\
&\quad + \left(\sum_{i^s \in \{G', N', B'\}} \Pr[I | i^s \cap B] \Pr[i^s | B] \right) \Pr[B] \cdot c \\
&= \left(\sum_{i^s \in \{G', N', B'\}} \Pr[I | i^s] \Pr[i^s | G] \right) \Pr[G] \cdot P \\
&\quad + \left(\sum_{i^s \in \{G', N', B'\}} \Pr[I | i^s] \Pr[i^s | B] \right) \Pr[B] \cdot c \\
&= \kappa(0.5P + 0.5c\frac{\widehat{\theta}_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{\widehat{\theta}_{N'}}{\bar{\theta}}(0.5P + 0.5c), \qquad (151)
\end{aligned}$$

where the second equality is by the law of total probability, the third equality is by the fact that sender messages in equilibrium (and hence the receiver's actions) do not depend on the true state of the world once his information state i^s is conditioned upon, and the fourth equality is obtained by substituting $\Pr[I|G'] = 1$, $\Pr[I|N'] = \frac{\widehat{\theta}_{N'}}{\bar{\theta}}$, $\Pr[I|B'] = \frac{\widehat{\theta}_{B'}}{\bar{\theta}}$, $\Pr[G'|G] = \Pr[B'|B] = \kappa$, $\Pr[G'|B] = \Pr[B'|G] = 0$, and $\Pr[N'|G] = \Pr[N'|B] = 1 - \kappa$.

Denote further the ex ante receiver's utility in the lying equilibrium as $U^{r,l}$ and in the evasion equilibrium as $U^{r,e}$.

By Lemma 5 we have the following possible equilibrium cases: 1) $\widehat{\theta}^e_{N'} = \bar{\theta}$ and $\widehat{\theta}^l_{N'} = \bar{\theta}$; 2) $\widehat{\theta}^e_{N'} < \bar{\theta}$ and $\widehat{\theta}^l_{N'} < \bar{\theta}$; 3) $\widehat{\theta}^e_{N'} = \bar{\theta}$ and $\widehat{\theta}^l_{N'} < \bar{\theta}$. Let us consider these cases sequentially.

Case 1: $\widehat{\theta}^e_{N'} = \bar{\theta}$ and $\widehat{\theta}^l_{N'} = \bar{\theta}$.

By Propositions 1, 2 and Lemmas 13 and 16 such equilibria simultaneously exist whenever $F \in [\bar{\theta}\frac{P-c}{4-\kappa}, \frac{\bar{\theta}(P+c)(1-\kappa)}{\kappa}]$ if $\kappa \in (\frac{P+c}{P}, \frac{2(P+c)}{2P+c}]$ and $F \geq \bar{\theta}\frac{P-c}{4-\kappa}$ if $\kappa \leq \frac{P+c}{P}$.

By (151) we obtain

$$\begin{aligned}
U^{r,l} - U^{r,e} &= \kappa(0.5P + 0.5c\frac{\widehat{\theta}^l_{B'}}{\bar{\theta}}) - \kappa(0.5P + 0.5c\frac{\widehat{\theta}^e_{B'}}{\bar{\theta}}) \\
&= \kappa\frac{0.5c}{\bar{\theta}}(\widehat{\theta}^l_{B'} - \widehat{\theta}^e_{B'}) \geq 0, \qquad (152)
\end{aligned}$$

where the last inequality is by Lemma 5.

Case 2: $\widehat{\theta}^e_{N'} < \bar{\theta}$ and $\widehat{\theta}^l_{N'} < \bar{\theta}$.

By Propositions 1 and 2 and Lemmas 13 and 16 such equilibria simultaneously exist whenever $F < \frac{\bar{\theta}(P-c)(1-\kappa)}{4-3\kappa}$ and $\kappa \in (\frac{P+c}{P}, \frac{2(P+c)}{2P+c}]$, or $\kappa \leq \frac{P+c}{P}$ and $F \in (\max[\widetilde{F}^e, \widetilde{F}^l], \frac{\bar{\theta}(P-c)(1-\kappa)}{4-3\kappa})$.

By (151) and $\widehat{\theta}_{N'} = 2\widehat{\theta}_{B'}$ (by Lemmas 11 and 14) we obtain

$$\begin{aligned}
U^{r,l} - U^{r,e} &= \kappa(0.5P + 0.5c\frac{\widehat{\theta}^l_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{2\widehat{\theta}^l_{B'}}{\bar{\theta}}(0.5P + 0.5c) \\
&\quad -\kappa(0.5P + 0.5c\frac{\widehat{\theta}^e_{B'}}{\bar{\theta}}) - (1-\kappa)\frac{2\widehat{\theta}^e_{B'}}{\bar{\theta}}(0.5P + 0.5c) \\
&= \kappa\frac{0.5c}{\bar{\theta}}(\widehat{\theta}^l_{B'} - \widehat{\theta}^e_{B'}) + (1-\kappa)\frac{P+c}{\bar{\theta}}(\widehat{\theta}^l_{B'} - \widehat{\theta}^e_{B'}) \\
&= (\widehat{\theta}^l_{B'} - \widehat{\theta}^e_{B'})(\frac{0.5\kappa c + (1-\kappa)(P+c)}{\bar{\theta}}). \qquad (153)
\end{aligned}$$

By Lemma 5 the first term in the RHS is negative. At the same time, the second term is positive whenever $\kappa \leq \frac{2(P+c)}{2P+c}$, that is a necessary condition for the existence of the evasion equilibrium. Consequently,

$$U^{r,l} - U^{r,e} \leq 0 \qquad (154)$$

in this case. Note that the inequality is strict if $\kappa \neq \frac{2(P+c)}{2P+c}$.

Case 3: $\widehat{\theta}^e_{N'} = \bar{\theta}$ and $\widehat{\theta}^l_{N'} < \bar{\theta}$.

Propositions 1, 2 and Lemmas 13 and 16 imply that such equilibria simultaneously exist whenever $F \in [\frac{\bar{\theta}(P-c)(1-\kappa)}{4-3\kappa}, \min[\frac{\bar{\theta}(P+c)(1-\kappa)}{\kappa}, \bar{\theta}\frac{P-c}{4-\kappa}])$ if $\kappa \in (\frac{P+c}{P}, \frac{2(P+c)}{2P+c}]$ and $F \in [\frac{\bar{\theta}(P-c)(1-\kappa)}{4-3\kappa}, \bar{\theta}\frac{P-c}{4-\kappa})$ if $\kappa \leq \frac{P+c}{P}$.

Let us first show

$$\frac{d(U^{r,l} - U^{r,e})}{dF} > 0.$$

By (151), given that $\widehat{\theta}^l_{N'} = 2\widehat{\theta}^l_{B'}$ and $\widehat{\theta}^e_{N'} = \bar{\theta}$, we have

$$\begin{aligned}
U^{r,l} - U^{r,e} &= \kappa(0.5P + 0.5c\frac{\widehat{\theta}^l_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{2\widehat{\theta}^l_{B'}}{\bar{\theta}}(0.5P + 0.5c) \\
&\quad -\kappa(0.5P + 0.5c\frac{\widehat{\theta}^e_{B'}}{\bar{\theta}}) - (1-\kappa)\frac{1}{\bar{\bar{\theta}}}(0.5P + 0.5c) \\
&= -0.5\kappa c\frac{\widehat{\theta}^e_{B'}}{\bar{\theta}} + (P(1-\kappa) + c(1-0.5\kappa))\frac{\widehat{\theta}^l_{B'}}{\bar{\theta}} \\
&\quad -(1-\kappa)(0.5P + 0.5c). \qquad (155)
\end{aligned}$$

Then,

$$\frac{d(U^{r,l} - U^{r,e})}{dF} = \frac{-0.5\kappa c}{\bar{\theta}}\frac{d\widehat{\theta}^e_{B'}}{dF} + (P(1-\kappa) + c(1-0.5\kappa))\frac{1}{\bar{\bar{\theta}}}\frac{d\widehat{\theta}^l_{B'}}{dF}. \qquad (156)$$

By (67) and (96) we have that $\frac{d\widehat{\theta}^e_{B'}}{dF}$ and $\frac{d\widehat{\theta}^l_{B'}}{dF}$ are positive. Besides, the term $P(1-\kappa) + c(1-0.5\kappa)$ is positive since $\kappa \leq \frac{2(P+c)}{2P+c}$. Consequently,

$$\frac{d(U^{r,l} - U^{r,e})}{dF} > 0. \qquad (157)$$

Next, if F goes to $\frac{\bar{\theta}(P-c)(1-\kappa)}{4-3\kappa}$ (the lower bound of the case), then $\widehat{\theta}^e_{B'} \to 0.5\bar{\theta} = 0.5\widehat{\theta}^e_{N'}$ so that $U^{r,l} - U^{r,e}$ converges to the RHS of (153), which is negative. Hence,

$$\lim_{F \to \frac{\bar{\theta}(P-c)(1-\kappa)}{4-3\kappa}} U^{r,l} - U^{r,e} \leq 0 \tag{158}$$

with a strict inequality if $\kappa \neq \frac{2(P+c)}{2P+c}$. If F goes to $\bar{\theta}\frac{P-c}{4-\kappa}$ (one of the possible upper bounds of the case), then $\widehat{\theta}^l_{N'} \to \bar{\theta}$ corresponding to Case 1. Hence, by (152)

$$\lim_{F \to \bar{\theta}\frac{P-c}{4-\kappa}} U^{r,l} - U^{r,e} > 0 \tag{159}$$

(with a strict inequality since in this case $\widehat{\theta}^l_{B'}$ converges to $0.5\bar{\theta} < \bar{\theta}$ so that $\widehat{\theta}^l_{B'} < \widehat{\theta}^e_{B'}$ by Lemma 5). Finally, if F goes to $\frac{\bar{\theta}(P+c)(1-\kappa)}{\kappa}$ (the other possible upper bound of the case), then $\eta(m_{N'})$ converges to $\underline{\eta}$ (see (106)). Consequently, the receiver is indifferent between investing and abstaining after $m_{N'}$, i.e., $E[U^{r,e}|m_{N'}] = 0$. Then, the only ex ante profitable message in the evasion equilibrium is $m_{G'}$ (sent only by types in state G'), that is

$$U^{r,e} = E[U^{r,e}|m_{G'}]\Pr[m_{G'}] = 0.5P\kappa. \tag{160}$$

It follows

$$\begin{aligned} &\lim_{F \to \frac{\bar{\theta}(P+c)(1-\kappa)}{\kappa}} U^{r,l} - U^{r,e} \\ &= \kappa(0.5P + 0.5c\frac{\widehat{\theta}^l_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{2\widehat{\theta}^l_{B'}}{\bar{\theta}}(0.5P + 0.5c) - 0.5P\kappa \\ &= \frac{\widehat{\theta}^l_{B'}}{\bar{\theta}}(P(1-\kappa) + c(1-0.5\kappa)) \geq 0 \end{aligned} \tag{161}$$

as far as $\kappa \leq \frac{2(P+c)}{2P+c}$ (with a strict inequality if $\kappa < \frac{2(P+c)}{2P+c}$). From (157), (158), (159) and (161) it follows that in Case 3, if $\kappa < \frac{2(P+c)}{2P+c}$, then there exists a threshold value $\frac{\bar{\theta}(P-c)(1-\kappa)}{4-3\kappa} < F^* < \min[\frac{\bar{\theta}(P+c)(1-\kappa)}{\kappa}, \bar{\theta}\frac{P-c}{4-\kappa}]$ such that $U^{r,l} = U^{r,e}$ if $F = F^*$, $U^{r,l} > U^{r,e}$ if $F > F^*$, and $U^{r,l} < U^{r,e}$ if $F < F^*$. If $\kappa = \frac{2(P+c)}{2P+c}$ then $U^{r,l} = U^{r,e}$ in the considered case (since one can show that, under this value of κ, Case 3 is given by a single value of $F = \frac{\bar{\theta}(P-c)(1-\kappa)}{4-3\kappa} = \frac{\bar{\theta}(P+c)(1-\kappa)}{\kappa}$; then $U^{r,l} = U^{r,e}$ by (161)).

Combining the results of Cases 1-3 leads to the claim of the proposition. ■

Proof of Lemma 6. The argumentation is provided in the main text. ■

Proof of Proposition 7. By Lemma 6 the incentives in the evasion equilibrium are not affected by the policy, consequently, it exists under the same parameter restrictions as without the policy.

Let us consider the evasive babbling equilibrium. The consistency of the sender's incentives was considered in the text. The receiver's constraint $\eta(m_{G'}) \geq \underline{\eta}$ is trivially satisfied. Besides, the receiver abstains after $m_{N'}$ if and only if

$$\eta(m_{N'}) < \underline{\eta}. \tag{162}$$

If all types in states N' and B' send $m_{N'}$, then by (15)

$$\eta(m_{N'}) = \frac{1-\kappa}{2-\kappa}. \tag{163}$$

Consequently, the incentive constraint (162) is satisfied (so that the evasive babbling equilibrium exists) if and only if $\kappa > \frac{P+c}{P}$.

Let us consider whether other equilibria besides the evasion and the evasive babbling exist under lying prohibition. All equilibria except for the lying, evasion or evasive babbling equilibria are ruled out by Assumption 3 as before.[23] At the same time, the lying equilibrium does not exist by Lemma 6 and the assumption that the fine for lying is sufficiently high (since then all types sending $m_{G'}$ in states B' and N' prefer to deviate to $m_{B'}$). Consequently, the only possible equilibria under lying prohibition are the evasion and evasive babbling equilibria. ■

Proof of Proposition 8. Proposition 7 implies that the only possible results of the lying prohibition policy (given that the pre-policy equilibrium is the lying equilibrium) are switches to either evasion or evasive babbling equilibrium (cases of constructive and destructive evasion respectively). Let us consider each of these cases.

Case of constructive evasion: The parameter restrictions given in the proposition specify cases when the evasion equilibrium exists (see Proposition 2).[24] The welfare consequences of the switch from the lying to the evasion equilibrium are given by Proposition 6.

Case of destructive evasion: By Propositions 1 and 7 both lying and evasive babbling equilibria exist if and only if $\kappa > \frac{P+c}{P}$. To exclude the parallel existence of the evasion equilibrium (as the case requires), it must also hold that $F > \bar{\theta}\,(P+c)\,\frac{(1-\kappa)}{\kappa}$ whenever $\kappa \in (\frac{P+c}{P}, \frac{2(P+c)}{2P+c}]$.

Let us compare the receiver's ex ante utility in both equilibria (denoting by $U^{r,eb}$ the ex ante utility in the evasive babbling equilibrium). In the evasive babbling equilibrium the receiver invests if and only if the sender is in state G' which happens with probability

[23] Technically, lying is not defined for such equilibria (for simplicity of exposition), hence the incentive structure there remains the same as without the policy (so that the refinement of lexicographic preferences over messages still applies as before). However, one can show that this argument remains valid under *any* possible definition of lying (where lying is contingent on the information state and message formulation), as far as one assumes that the message used by types in state G' (once it is the same for all these types) is $m_{G'}$ in any equilibrium (in other words, if one generalizes the implicit definition of message $m_{G'}$ over other equilibria). The proof is available upon request.

[24] Note that $F > \widetilde{F}^e(\kappa)$ in case of $\kappa \leq \frac{P+c}{P}$ is implied by the existence of the lying equilibrium. Indeed, the latter yields $F \geq \widetilde{F}^l(\kappa) > \widetilde{F}^e(\kappa)$, where the last inequality follows from Lemma 5 and the fact that $\eta(m_{B'}|\widehat{\theta}_{B'})$ is decreasing in $\widehat{\theta}_{B'}$ (see the proofs of Propositions 1 and 2).

0.5κ, so that his ex ante utility is $U^{r,eb} = 0.5\kappa P$. Further, we have the following possible cases depending on whether $\widehat{\theta}^l_{N'} = 2\widehat{\theta}^l_{B'}$ (see Lemma 11).

Case 1: $\widehat{\theta}^l_{N'} = 2\widehat{\theta}^l_{B'}$ so that $F < \bar{\theta}\frac{P-c}{4-\kappa}$ (by Lemma 13). Then, by (151)

$$\begin{aligned} U^{r,l} - U^{r,eb} &= \kappa(0.5P + 0.5c\frac{\widehat{\theta}^l_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{2\widehat{\theta}^l_{B'}}{\bar{\theta}}(0.5P + 0.5c) - 0.5\kappa P \\ &= 0.5\frac{\widehat{\theta}^l_{B'}}{\bar{\theta}}(\kappa c + (1-\kappa)2(P+c)). \end{aligned} \tag{164}$$

Then,

$$\begin{aligned} U^{r,l} &\geq U^{r,eb} \text{ iff } \kappa \in (\frac{P+c}{P}, \frac{2(P+c)}{2P+c}], \\ U^{r,l} &< U^{r,eb} \text{ iff } \kappa \in (\frac{2(P+c)}{2P+c}, 1]. \end{aligned}$$

Case 2: $\widehat{\theta}^l_{N'} = \bar{\theta}$ so that $\widehat{\theta}^l_{B'} \geq 0.5\bar{\theta}$ (by Lemma 11) and $F \geq \bar{\theta}\frac{P-c}{4-\kappa}$ (by Lemma 13). Then,

$$\begin{aligned} U^{r,l} - U^{r,eb} &= \kappa(0.5P + 0.5c\frac{\widehat{\theta}^l_{B'}}{\bar{\theta}}) + (1-\kappa)(0.5P + 0.5c) - 0.5\kappa P \\ &= 0.5(\frac{\widehat{\theta}^l_{B'}}{\bar{\theta}}\kappa c + (1-\kappa)(P+c)). \end{aligned} \tag{165}$$

It follows that if $\kappa > \frac{2(P+c)}{2P+c}$ then for any $\widehat{\theta}^l_{B'} \geq 0.5\bar{\theta}$ we have $U^{r,l} - U^{r,eb} < 0$. Besides, if $\widehat{\theta}^l_{B'} = \bar{\theta}$ (i.e., $F \geq 0.5\bar{\theta}(P-c)$, see Lemma 13), then the RHS of (165) is always negative since $\kappa > \frac{P+c}{P}$ by assumption. Consider the remaining case $\widehat{\theta}^l_{B'} \in [0.5\bar{\theta}, \bar{\theta})$ (i.e., $F \in \big[\bar{\theta}\frac{P-c}{4-\kappa}, 0.5\bar{\theta}(P-c)\big)$) and $\kappa \in (\frac{P+c}{P}, \frac{2(P+c)}{2P+c}]$. Then, by Lemma 11 $\widehat{\theta}^l_{B'}$ must solve the indifference condition

$$F - \widehat{\theta}^l_{B'}\eta(m_{G'}|\widehat{\theta}^l_{B'})(P-c) = 0, \tag{166}$$

where by (15) (given $\widehat{\theta}^l_{N'} = \bar{\theta}$)

$$\eta(m_{G'}|\widehat{\theta}^l_{B'}) = \frac{\bar{\theta}}{2\bar{\theta} - \kappa(\bar{\theta} - \widehat{\theta}^l_{B'})}. \tag{167}$$

Solving (166) we get

$$\widehat{\theta}^l_{B'} = \frac{F\bar{\theta}(2-\kappa)}{(P-c)\bar{\theta} - F\kappa}. \tag{168}$$

Substituting this into (165) yields that if $F \in \big[\bar{\theta}\frac{P-c}{4-\kappa}, 0.5\bar{\theta}(P-c)\big)$ and $\kappa \in (\frac{P+c}{P}, \frac{2(P+c)}{2P+c}]$, then

$$U^{r,l} - U^{r,eb} \geq 0 \;\Leftrightarrow\; F \leq \frac{(P^2-c^2)(1-\kappa)\bar{\theta}}{(P(1-\kappa)-c)\kappa}. \tag{169}$$

Merging Case 1 and Case 2 together, given that

$$\frac{(P^2-c^2)(1-\kappa)\bar{\theta}}{(P(1-\kappa)-c)\kappa} \geq \max\left[\bar{\theta}\frac{P-c}{4-\kappa}, \bar{\theta}\left(P+c\right)\frac{(1-\kappa)}{\kappa}\right]$$

if $\kappa \in (\frac{P+c}{P}, \frac{2(P+c)}{2P+c}]$, we have that generally

$$U^{r,l} - U^{r,eb} \geq 0 \;\Leftrightarrow\; F \leq \frac{(P^2-c^2)(1-\kappa)\bar{\theta}}{(P(1-\kappa)-c)\kappa} \wedge \kappa \in (\frac{P+c}{P}, \frac{2(P+c)}{2P+c}]. \tag{170}$$

■

Proof of Proposition 9. Consider first the lying equilibrium. Since $U^{r,l}$ does not depend on F directly, but only through the cutoffs (and $\widehat{\theta}^l_{N'}$ is a function of $\widehat{\theta}^l_{B'}$ by Lemma 11), we have

$$\frac{dU^{r,l}}{dF} = \frac{\partial U^{r,l}}{\partial \widehat{\theta}^l_{B'}} \frac{d\widehat{\theta}^l_{B'}}{dF}. \tag{171}$$

Consider the RHS of (171). We have the following possible cases, as in the previous proofs.

Case 1: $\bar{\theta} > \widehat{\theta}^l_{N'} = 2\widehat{\theta}^l_{B'}$. Then, by (151)

$$\frac{\partial U^{r,l}}{\partial \widehat{\theta}^l_{B'}} = \frac{(c(2-\kappa)+2P(1-\kappa))}{2\bar{\theta}}. \tag{172}$$

At the same time, $\frac{d\widehat{\theta}^l_{B'}}{dF} > 0$ (see (67)). Then, (171) and (172) imply that $\frac{dU^{r,l}}{dF} > 0$ if $\kappa < \frac{2(P+c)}{2P+c}$ and $\frac{dU^{r,l}}{dF} \leq 0$ if $\kappa \geq \frac{2(P+c)}{2P+c}$.

Case 2: $\bar{\theta} = \widehat{\theta}^l_{N'} > \widehat{\theta}^l_{B'}$. Then, by (151)

$$\frac{\partial U^{r,l}}{\partial \widehat{\theta}^l_{B'}} = \frac{c\kappa}{2\bar{\theta}} < 0, \tag{173}$$

while $\frac{d\widehat{\theta}^l_{B'}}{dF} > 0$ as well. Hence, $\frac{dU^{r,l}}{dF} < 0$.

Case 3: $\bar{\theta} = \widehat{\theta}^l_{N'} = \widehat{\theta}^l_{B'}$. Then, $\frac{d\widehat{\theta}^l_{B'}}{dF} = \frac{dU^{r,l}}{dF} = 0$.

In total, Cases 1-3 imply that $\frac{dU^{r,l}}{dF} > 0$ if and only if $\kappa < \frac{2(P+c)}{2P+c}$ and $\widehat{\theta}^l_{N'} < \bar{\theta}$. The latter condition corresponds to $F < \frac{(P-c)}{4-\kappa}\bar{\theta}$ by Lemma 13.

In case of the evasion equilibrium, by analogous arguments (given that (171), (172) and (173) apply to this equilibrium as well, while $\frac{d\widehat{\theta}^e_{B'}}{dF} > 0$ as far as $\widehat{\theta}^e_{B'} < \bar{\theta}$, see (96)), we obtain that $\frac{dU^r_E}{dF} > 0$ if and only if $\kappa < \frac{2(P+c)}{2P+c}$ and $\widehat{\theta}^e_{N'} < \bar{\theta} \Leftrightarrow F < \frac{(P-c)(1-\kappa)}{4-3\kappa}\bar{\theta}$. ■

Proof of Proposition 10. If $\kappa \leq \frac{P+c}{P}$ and $F = 0$ then the incentive constraint $\eta(m_{B'}) < \underline{\eta}$ is violated in both lying and evasion equilibria (see Case 3 in the proofs of Propositions 1 and 2). Consequently, the only possible equilibria are of Subtype 1 where abstaining is not induced on equilibrium path (which exist if for any out-of-equilibrium message $\widetilde{m}$ it holds: $\eta(\widetilde{m}) \geq \eta(m_{N'})$ in the evasion equilibrium and $\eta(\widetilde{m}) \geq \eta(m_{G'})$ in the lying equilibrium. This is possible since we lift Assumption 2).

Let us show that the ex ante receiver's welfare in this case is lower than in any equilibrium of Subtype 2 or 3. Indeed, if either lying or evasion equilibrium is of Subtype 1, i.e., if $\bar{\theta} = \widehat{\theta}_{N'} = \widehat{\theta}_{B'}$, then by (151) we have

$$E[U^r|\bar{\theta} = \widehat{\theta}_{N'} = \widehat{\theta}_{B'}] = 0.5P + 0.5c. \tag{174}$$

At the same time, if at least $\widehat{\theta}_{B'} < \bar{\theta}$, that is the ex ante probability of getting message $m_{B'}$ is strictly positive, we have

$$\begin{aligned} E[U^r|\widehat{\theta}_{B'} &< \bar{\theta}] = E[U^r(I)|m \neq m_{B'}] \Pr[m \neq m_{B'}] + E[U^r(A)|m = m_{B'}] \Pr[m = m_{B'}] \\ &> E[U^r(I)|m \neq m_{B'}] \Pr[m \neq m_{B'}] + E[U^r(I)|m = m_{B'}] \Pr[m = m_{B'}] \\ &= 0.5P + 0.5c = E[U^r|\bar{\theta} = \widehat{\theta}_{N'} = \widehat{\theta}_{B'}], \end{aligned} \tag{175}$$

where the first equality is by the law of total probability, the inequality follows from the receiver's incentive compatibility (i.e., $E[U^r(A)|m = m_{B'}] > E[U^r(I)|m = m_{B'}]$ while $m_{B'}$ induces abstaining), the second equality is by the fact that if the receiver invests after all possible messages, then she has the same expected payoff as after investment without advice, and the last equality is by (174). ■

Proof of Proposition 11.

i) *The lying equilibrium.*

Let us consider the following possible cases depending on whether $\widehat{\theta}^l_{N'} = 2\widehat{\theta}^l_{B'}$ (see Lemma 11).

Case 1: $\bar{\theta} > \widehat{\theta}^l_{N'} = 2\widehat{\theta}^l_{B'}$ so that $F < \bar{\theta}\frac{P-c}{4-\kappa}$ (by Lemma 13). By (151) we have

$$U^{r,l} = \kappa(0.5P + 0.5c\frac{\widehat{\theta}^l_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{2\widehat{\theta}^l_{B'}}{\bar{\theta}}(0.5P + 0.5c). \tag{176}$$

Claim 1. $\frac{d^2U^{r,l}}{d\kappa^2} > 0$.

Proof. Denoting $\left(\widehat{\theta}^l_{B'}\right)' = \frac{d\widehat{\theta}^l_{B'}}{d\kappa}$ and $\left(\widehat{\theta}^l_{B'}\right)'' = \frac{d^2\widehat{\theta}^l_{B'}}{d\kappa^2}$ we have

$$\begin{aligned} \frac{dU^r_L}{d\kappa} &= \frac{d\left(\kappa(0.5P + 0.5c\frac{\widehat{\theta}^l_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{2\widehat{\theta}^l_{B'}}{\bar{\theta}}(0.5P + 0.5c)\right)}{d\kappa} \\ &= \frac{P\bar{\theta} - (2P+c)\widehat{\theta}^l_{B'} + (2(P+c) - (2P+c)\kappa)\left(\widehat{\theta}^l_{B'}\right)'}{2\bar{\theta}}, \end{aligned} \tag{177}$$

so that

$$\frac{d^2U^{r,l}}{d\kappa^2} = \frac{-2(2P+c)\left(\widehat{\theta}^l_{B'}\right)' + (2(P+c)-(2P+c)\kappa)\left(\widehat{\theta}^l_{B'}\right)''}{2\bar{\theta}}. \tag{178}$$

Consider $\left(\widehat{\theta}^l_{B'}\right)'$ and $\left(\widehat{\theta}^l_{B'}\right)''$. By Lemma 11, $\varpi(\widehat{\theta}^l_{B'}) = 0$. Consequently, by the implicit function theorem

$$\left(\widehat{\theta}^l_{B'}\right)' = -\frac{\partial\varpi(\widehat{\theta}^l_{B'})/\partial\kappa}{\partial\varpi(\widehat{\theta}^l_{B'})/\partial\widehat{\theta}^l_{B'}}. \tag{179}$$

Substituting and simplifying we obtain

$$\left(\widehat{\theta}^l_{B'}\right)' = -\frac{2\left(\widehat{\theta}^l_{B'}\right)^2(\bar{\theta}-\widehat{\theta}^l_{B'})}{\kappa^2\bar{\theta}^2 + 4(1-\kappa)\kappa\bar{\theta}\widehat{\theta}^l_{B'} + 2(1-\kappa)(4-3\kappa)\left(\widehat{\theta}^l_{B'}\right)^2}. \tag{180}$$

It follows, $\left(\widehat{\theta}^l_{B'}\right)' < 0$. Further, by the chain rule

$$\left(\widehat{\theta}^l_{B'}\right)'' = \frac{\partial\left(\widehat{\theta}^l_{B'}\right)'}{\partial\kappa} + \frac{\partial\left(\widehat{\theta}^l_{B'}\right)'}{\partial\widehat{\theta}^l_{B'}}\left(\widehat{\theta}^l_{B'}\right)'. \tag{181}$$

Calculating this expression by (180) and substituting it together with (180) into (178) we get

$$\frac{d^2U^{r,l}}{d\kappa^2} = \frac{4(\bar{\theta}-\widehat{\theta}^l_{B'})(\widehat{\theta}^l_{B'})^2}{\bar{\theta}(\kappa^2\bar{\theta}^2 + 4(1-\kappa)\kappa\bar{\theta}\widehat{\theta}^l_{B'} + 2(4-7\kappa+3\kappa^2)(\widehat{\theta}^l_{B'})^2)^3}Z, \tag{182}$$

where

$$\begin{aligned} Z &= (P+c)\kappa^3\bar{\theta}^4 + 8(P+c)(1-\kappa)\kappa^2\bar{\theta}^3\widehat{\theta}^l_{B'} - 2c(4-3\kappa)^2(1-\kappa)(\widehat{\theta}^l_{B'})^4 \\ &\quad +2\kappa\bar{\theta}^2(\widehat{\theta}^l_{B'})^2(10P(1-\kappa)^2 + c(10-23\kappa+12\kappa^2)) \\ &\quad +8(1-\kappa)(2P(1-\kappa)^2 + c(2-7\kappa+4\kappa^2))\bar{\theta}(\widehat{\theta}^l_{B'})^3. \end{aligned}$$

The multiple of Z in (182) is clearly positive, hence to show the claim we need to prove that $Z > 0$. The first three terms of Z are clearly positive. For the fourth term we have (given that $P > -c$ by Assumption 1)

$$\begin{aligned} 10P(1-\kappa)^2 + c(10-23\kappa+12\kappa^2) &> -10c(1-\kappa)^2 + c(10-23\kappa+12\kappa^2) \\ &= -c\kappa(3-2\kappa) > 0. \end{aligned} \tag{183}$$

Finally, for the fifth term we have

$$\begin{aligned} 2P(1-\kappa)^2 + c(2-7\kappa+4\kappa^2) &> -2c(1-\kappa)^2 + c(2-7\kappa+4\kappa^2) \\ &= -c\kappa(3-2\kappa) > 0. \end{aligned} \tag{184}$$

Consequently, $Z > 0$ so that by (182)

$$\frac{d^2U^{r,l}}{d\kappa^2} > 0. \tag{185}$$

Claim 2. $\frac{dU_L^r}{d\kappa}_{|\kappa=0} < 0$.

Proof. If $\kappa = 0$, then all senders who send $m_{G'}$ are uninformed so that $\eta(m_{G'}) = 0.5$. In this case $\varpi(\widehat{\theta}^l_{B'}) = 0$ implies

$$\begin{aligned} F - \widehat{\theta}^l_{B'} 0.5(P-c) &= 0, \\ \widehat{\theta}^l_{B'} &= \frac{2F}{P-c}. \end{aligned} \tag{186}$$

Substituting this together with $\kappa = 0$ into (180) we get

$$\left(\widehat{\theta}^l_{B'}\right)' = \frac{F}{2(P-c)} - \frac{\bar{\theta}}{4}. \tag{187}$$

Substituting (186) and (187) into (177) results in

$$\begin{aligned} \frac{dU_L^r}{d\kappa}_{|\kappa=0} &= \frac{P\bar{\theta} - (2P+c)\frac{2F}{P-c} + 2(P+c)(\frac{F}{2(P-c)} - \frac{\bar{\theta}}{4})}{2\bar{\theta}} \\ &= \frac{\bar{\theta}(P-c)^2 - 2F(3P+c)}{4\bar{\theta}(P-c)}. \end{aligned} \tag{188}$$

It follows that Claim 2 holds if and only if

$$F > \frac{\bar{\theta}(P-c)^2}{2(3P+c)}.$$

Given the initial constraint of Case 1 $F < \bar{\theta}\frac{P-c}{4-\kappa}$, the interval $(\frac{\bar{\theta}(P-c)^2}{2(3P+c)}, \bar{\theta}\frac{P-c}{4-\kappa})$ is nonempty if P is sufficiently high:

$$P > \frac{-c(6-\kappa)}{2+\kappa}.$$

Claim 3. $\lim_{\kappa\to 1} \frac{dU_L^r}{d\kappa} > 0$.

Proof. From (180) we have

$$\lim_{\kappa\to 1} \left(\widehat{\theta}^l_{B'}\right)' = -\frac{2(\bar{\theta} - \widehat{\theta}^l_{B'})(\widehat{\theta}^l_{B'})^2}{\bar{\theta}^2}. \tag{189}$$

Equations (177) and (189) lead to

$$\lim_{\kappa\to 1}\frac{dU_L^r}{d\kappa}=\frac{P\bar{\theta}^2(\bar{\theta}-2\widehat{\theta}^l_{B'})-c\widehat{\theta}^l_{B'}(\bar{\theta}+2\widehat{\theta}^l_{B'}(\bar{\theta}-\widehat{\theta}^l_{B'}))}{2\bar{\theta}^3}>0, \tag{190}$$

where the inequality is by $\bar{\theta}>2\widehat{\theta}^l_{B'}$ by assumption.

Claims 1-3 and the intermediate value theorem imply that whenever $F\in(\frac{\bar{\theta}(P-c)^2}{2(3P+c)},\bar{\theta}\frac{P-c}{4-\kappa})$ (which is nonempty if $P>\frac{-c(6-\kappa)}{2+\kappa}$), $U^{r,l}$ is U-shaped with respect to κ. Otherwise, it is always increasing with κ (in the considered case of $F<\bar{\theta}\frac{P-c}{4-\kappa}$) .

Case 2: $\widehat{\theta}^l_{N'}=\bar{\theta}$ and $\widehat{\theta}^l_{B'}<\bar{\theta}$ so that $F\in[\bar{\theta}\frac{P-c}{4-\kappa},0.5\bar{\theta}(P-c))$ (by Lemma 13). Then, by (151)

$$U^{r,l}=\kappa(0.5P+0.5c\frac{\widehat{\theta}^l_{B'}}{\bar{\theta}})+(1-\kappa)(0.5P+0.5c). \tag{191}$$

We have

$$\frac{dU_L^r}{d\kappa}=\frac{\partial U^{r,l}}{\partial\kappa}+\frac{\partial U^{r,l}}{\partial\widehat{\theta}^l_{B'}}\left(\widehat{\theta}^l_{B'}\right)', \tag{192}$$

where,

$$\frac{\partial U^{r,l}}{\partial\kappa}=\frac{-c(\bar{\theta}-\widehat{\theta}^l_{B'})}{2\bar{\theta}}>0, \tag{193}$$

$$\frac{\partial U^{r,l}}{\partial\widehat{\theta}^l_{B'}}=\frac{c\kappa}{2\bar{\theta}}<0. \tag{194}$$

At the same time,

$$\left(\widehat{\theta}^l_{B'}\right)'=-\frac{\partial\varpi(\widehat{\theta}^l_{B'})/\partial\kappa}{\partial\varpi(\widehat{\theta}^l_{B'})/\partial\widehat{\theta}^l_{B'}}=-\frac{(\bar{\theta}-\widehat{\theta}^l_{B'})\widehat{\theta}^l_{B'}}{(2-\kappa)\bar{\theta}}<0, \tag{195}$$

where the first equality is by (179), and the second equality is obtained by substituting for $\eta(m_{G'}|\widehat{\theta}^l_{B'})$ in $\varpi(\widehat{\theta}^l_{B'})$ (see (167)) and simplifying. Finally, (192)-(195) imply that whenever $F\in[\bar{\theta}\frac{P-c}{4-\kappa},0.5\bar{\theta}(P-c))$

$$\frac{dU_L^r}{d\kappa}>0. \tag{196}$$

Case 3: $\widehat{\theta}^l_{N'}=\bar{\theta}$ and $\widehat{\theta}^l_{B'}=\bar{\theta}$ so that $F\geq 0.5\bar{\theta}(P-c)$ (by Lemma 13). Then, by (151)

$$\begin{aligned}U^{r,l}&=\kappa(0.5P+0.5c)+(1-\kappa)(0.5P+0.5c)\\&=0.5P+0.5c.\end{aligned} \tag{197}$$

Hence, in this case

$$\frac{dU_L^r}{d\kappa}=0. \tag{198}$$

The results of Cases 1-3 together imply Claim i) of the proposition.

ii) *The evasion equilibrium.*
Case 1: $\widehat{\theta}^e_{N'} = 2\widehat{\theta}^e_{B'}$ so that $F < \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$ (by Lemma 16). Then, by (151)

$$U^{r,e} = \kappa(0.5P + 0.5c\frac{\widehat{\theta}^e_{B'}}{\bar{\theta}}) + (1-\kappa)\frac{2\widehat{\theta}^e_{B'}}{\bar{\theta}}(0.5P + 0.5c). \tag{199}$$

We have

$$\frac{dU^r_E}{d\kappa} = \frac{\partial U^{r,e}}{\partial \kappa} + \frac{\partial U^{r,e}}{\partial \widehat{\theta}^e_{B'}}\left(\widehat{\theta}^e_{B'}\right)'. \tag{200}$$

At the same time,

$$\frac{\partial U^{r,e}}{\partial \kappa} = \frac{P(\bar{\theta} - 2\widehat{\theta}^e_{B'}) - c\widehat{\theta}^e_{B'}}{2\bar{\theta}} > 0, \tag{201}$$

$$\frac{\partial U^{r,e}}{\partial \widehat{\theta}^e_{B'}} = \frac{c(2-\kappa) + 2P(1-\kappa)}{2\bar{\theta}} > 0, \tag{202}$$

with the latter inequality by $\kappa \leq \frac{2(P+c)}{2P+c}$ (a necessary condition for the evasion equilibrium by Proposition 2). Finally, by the implicit function theorem and the fact that $\phi(\widehat{\theta}^e_{B'}) = 0$ by Lemma 14,

$$\left(\widehat{\theta}^e_{B'}\right)' = -\frac{\partial\phi(\widehat{\theta}^e_{B'})/\partial\kappa}{\partial\phi(\widehat{\theta}^e_{B'})/\partial\widehat{\theta}^e_{B'}}. \tag{203}$$

Substituting for $\eta(m_{N'}|\widehat{\theta}^e_{B'})$ in $\phi(\widehat{\theta}^e_{B'})$ (see (107)) and simplifying we get

$$-\frac{\partial\phi(\widehat{\theta}^e_{B'})/\partial\kappa}{\partial\phi(\widehat{\theta}^e_{B'})/\partial\widehat{\theta}^e_{B'}} = \frac{\widehat{\theta}^e_{B'}}{4(1-\kappa)^2 + \kappa(1-\kappa)} > 0. \tag{204}$$

(200)-(204) imply that whenever $F < \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$.

$$\frac{dU^r_E}{d\kappa} > 0. \tag{205}$$

Case 2: $\widehat{\theta}^e_{N'} = \bar{\theta}$ and $\widehat{\theta}^e_{B'} < \bar{\theta}$ so that $F \in [\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}, \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa})$ (by Lemma 16). Then,

$$U^{r,e} = \kappa(0.5P + 0.5c\frac{\widehat{\theta}^e_{B'}}{\bar{\theta}}) + (1-\kappa)(0.5P + 0.5c). \tag{206}$$

We have

$$\frac{dU^r_E}{d\kappa} = \frac{\partial U^{r,e}}{\partial \kappa} + \frac{\partial U^{r,e}}{\partial \widehat{\theta}^e_{B'}}\left(\widehat{\theta}^e_{B'}\right)'. \tag{207}$$

At the same time,

$$\frac{\partial U^{r,e}}{\partial \kappa} = \frac{-c(\bar{\theta} - \widehat{\theta}^{e}_{B'})}{2\bar{\theta}}, \tag{208}$$

$$\frac{\partial U^{r,e}}{\partial \widehat{\theta}^{e}_{B'}} = \frac{c\kappa}{2\bar{\theta}}. \tag{209}$$

As in the previous case, by the implicit function theorem

$$\left(\widehat{\theta}^{e}_{B'}\right)' = -\frac{\partial \phi(\widehat{\theta}^{e}_{B'})/\partial \kappa}{\partial \phi(\widehat{\theta}^{e}_{B'})/\partial \widehat{\theta}^{e}_{B'}} = \frac{\left(\widehat{\theta}^{e}_{B'}\right)^2}{2(1-\kappa)^2\bar{\theta}}, \tag{210}$$

where the last equality is obtained substituting (93) for $\eta(m_{N'}|\widehat{\theta}^{e}_{B'})$ in the expression for $\phi(\widehat{\theta}^{e}_{B'})$. Substituting (208), (209) and (210) into (207) we get

$$\frac{dU^r_E}{d\kappa} = \frac{-c(\bar{\theta} - \widehat{\theta}^{e}_{B'})}{2\bar{\theta}} + \frac{c\kappa}{2\bar{\theta}}\frac{\left(\widehat{\theta}^{e}_{B'}\right)^2}{2(1-\kappa)^2\bar{\theta}}. \tag{211}$$

Let us find $\widehat{\theta}^{e}_{B'}$. Solving the indifference condition

$$\phi(\widehat{\theta}^{e}_{B'}) = F - \widehat{\theta}^{e}_{B'}\eta(m_{N'}|\widehat{\theta}^{e}_{B'})(P-c) = F - \widehat{\theta}^{e}_{B'}\frac{\bar{\theta}(1-\kappa)}{\kappa\widehat{\theta}^{e}_{B'} + 2\bar{\theta}(1-\kappa)}(P-c) = 0 \tag{212}$$

yields

$$\widehat{\theta}^{e}_{B'} = \frac{2F(1-\kappa)\bar{\theta}}{(P-c)(1-\kappa)\bar{\theta} - F\kappa}. \tag{213}$$

Substituting this into (211) and simplifying we obtain

$$\frac{dU^r_E}{d\kappa} = \frac{-c}{2(F\kappa - (P-c)(1-\kappa)\bar{\theta})^2}(a_1\bar{\theta}^2 + a_2\bar{\theta} + a_3), \tag{214}$$

where

$$\begin{aligned} a_1 &= ((P-c)(1-\kappa))^2, \\ a_2 &= -2(P-c)F(1-\kappa), \\ a_3 &= -F^2\kappa^2. \end{aligned}$$

Since the fraction in (214) is strictly positive, $\frac{dU^r_E}{d\kappa} < 0$ whenever $a_1\bar{\theta}^2 + a_2\bar{\theta} + a_3 < 0$. The only positive real root of the corresponding quadratic equation is

$$\bar{\theta} = \frac{F(1+\sqrt{1+\kappa^2})}{(P-c)(1-\kappa)}. \tag{215}$$

Consequently, $\frac{dU_E^r}{d\kappa} < 0$ if

$$\bar{\theta} < \frac{F(1+\sqrt{1+\kappa^2})}{(P-c)(1-\kappa)} \Leftrightarrow F > \frac{(P-c)(1-\kappa)\bar{\theta}}{1+\sqrt{1+\kappa^2}}. \tag{216}$$

One can show that the RHS of 216 is smaller than the upper bound of F in the considered case:

$$\frac{(P-c)(1-\kappa)\bar{\theta}}{1+\sqrt{1+\kappa^2}} < \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}. \tag{217}$$

At the same time, it can be both larger and smaller than the lower bound $\frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}$ depending on the parameters. Consequently, $\frac{dU_E^r}{d\kappa} < 0$ whenever

$$F \in \left(\max[\frac{(P-c)(1-\kappa)\bar{\theta}}{1+\sqrt{1+\kappa^2}}, \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}], \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}\right).$$

Case 3: $\widehat{\theta}^l_{N'} = \bar{\theta}$ and $\widehat{\theta}^l_{B'} = \bar{\theta}$ so that $F \geq \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}$ (by Lemma 16). As in the case of the lying equilibrium, it is then straightforward to show that

$$\frac{dU_E^r}{d\kappa} = 0. \tag{218}$$

In total, Cases 1-3 imply that in the evasion equilibrium $\frac{dU_E^r}{d\kappa} \leq 0$ whenever $F \geq \max[\frac{(P-c)(1-\kappa)\bar{\theta}}{1+\sqrt{1+\kappa^2}}, \frac{\bar{\theta}(1-\kappa)(P-c)}{4-3\kappa}]$ (with a strict inequality if $F < \frac{\bar{\theta}(1-\kappa)(P-c)}{2-\kappa}$) and $\frac{dU_E^r}{d\kappa} > 0$ otherwise. ■

Proof of Proposition 12. Under the lying prohibition no type in states B' and N' can send $m_{G'}$ (i.e., pool with types in state G'), so that the lying equilibrium is eliminated. At the same time, it is easy to verify that both the evasion and the evasive babbling equilibria become possible (under certain parameter conditions), while no other equilibria exist (by analogous arguments as in the proof of Proposition 7).

If the lying prohibition induces a switch from the lying to the evasion equilibrium, then the equilibrium cutoffs do not change, since they are determined by (14) in both equilibria. While the ex ante utility of the receiver is uniquely determined by the cutoffs (see (151)), the equality of the cutoffs in both equilibria implies the equality of the ex ante utility. Hence, the lying prohibition policy does not have any welfare effects in this case.

Consider the case when the lying prohibition induces a switch to the evasive babbling equilibrium (when the evasion equilibrium does not exist). Since the evasion equilibrium does not exist, it holds (given the cutoffs making the sender not willing to deviate in the evasion equilibrium)

$$\eta^e(m_{N'}) < \underline{\eta}, \tag{219}$$

where the upper index denotes the equilibrium type.[25] At the same time, since under

[25] This is the only possible case when the evasion equilibrium does not exist while the (pre-policy)

outcome-based preferences $\widehat{\theta}^e_{i^s} = \widehat{\theta}^l_{i^s} = \widehat{\theta}^0_{i^s}$, it is easy to verify that

$$\eta^e(m_{N'}) = \eta^l(m_{G'}|B' \cup N'). \tag{220}$$

Applying the law of total expectation we obtain

$$\begin{aligned} U^{r,l} - U^{r,eb} &= E[U^{r,l}|B' \cup N'] \Pr[B' \cup N'] + E[U^{r,l}|G'] \Pr[G'] \\ &\quad - E[U^{r,eb}|B' \cup N'] \Pr[B' \cup N'] - E[U^{r,eb}|G'] \Pr[G'] \\ &= E[U^{r,l}|B' \cup N'] \Pr[B' \cup N'] \\ &= E[U^{r,l}(m_{G'})|B' \cup N'] \Pr[m_{G'}|B' \cup N'] \Pr[B' \cup N']. \end{aligned} \tag{221}$$

where the second equality follows from $E[U^{r,l}|G'] = E[U^{r,eb}|G'] = P$ and $E[U^{r,eb}|B' \cup N'] = 0$ by construction of equilibria, and the last equality follows from $U^{r,l}(m_{B'}) = 0$ and the law of total expectation. At the same time,

$$E[U^{r,l}(m_{G'})|B' \cup N'] = \eta^l(m_{G'}|B' \cup N')B + (1 - \eta^l(m_{G'}|B' \cup N')c < 0, \tag{222}$$

where the inequality follows from (219) and (220). This together with (221) implies

$$U^{r,l} - U^{r,eb} < 0. \tag{223}$$

Thus, the lying prohibition policy always increases welfare in this case. ■

Proof of Proposition 13. According to (151),

$$E[U^{r,0}] = \kappa(0.5P + 0.5c\frac{\widehat{\theta}^0_{B'}}{\bar{\theta}}) + (1 - \kappa)\frac{\widehat{\theta}^0_{N'}}{\bar{\theta}}(0.5P + 0.5c). \tag{224}$$

Since $\widehat{\theta}^0_{B'}$ and $\widehat{\theta}^0_{N'}$ do not depend on κ by (14), we have

$$\frac{\partial E[U^{r,0}]}{\partial \kappa} = \frac{1}{2\bar{\theta}}(P(\bar{\theta} - \widehat{\theta}^0_{N'}) - c(\widehat{\theta}^0_{N'} - \widehat{\theta}^0_{B'})) \geq 0, \tag{225}$$

where the inequality follows from $\widehat{\theta}^0_{N'} \geq \widehat{\theta}^0_{B'}$ and $c < 0$. ■

References

Abeler, J., D. Nosenzo, and C. Raymond (2016): "Preferences for Truth-Telling – A Meta Study," CESifo Working Paper Series 6087, CESifo Group Munich.

Austen-Smith, D. (1994): "Strategic Transmission of Costly Information," Econometrica, 62, 955–963.

lying equilibrium does. The other constraint $\eta(m_{B'}) < \underline{\eta}$ is always satisfied in the evasion equilibrium whenever it is satisfied in the lying equilibrium (due to the equality of the cutoffs).

Battigalli, P., G. Charness, and M. Dufwenberg (2013): “Deception: The Role of Guilt,” Journal of Economic Behavior & Organization, 93, 227–232.

Battigalli, P., and M. Dufwenberg (2007): “Guilt in Games,” American Economic Review, 97, 170–176.

Battigalli, P., and M. Dufwenberg (2009): “Dynamic Psychological Games,” Journal of Economic Theory, 141, 1–35.

Blume, A., O. J. Board, and K. Kawamura (2007): “Noisy Talk,” Theoretical Economics, 2, 395–440.

Bolton, P., X. Freixas, and J. Shapiro (2007): “Conflicts of Interest, Information Provision, and Competition in Banking,” Journal of Financial Economics, 85, 297–330.

Bolton, G. E., and A. Ockenfels (2000): “ERC: A Theory of Equity, Reciprocity, and Competition,” American Economic Review, 90, 166–193.

Charness, G., and M. Dufwenberg (2006): “Promises and Partnership,” Econometrica, 74, 1579–1601.

Chater, N., R. Inderst, and S. Huck (2010): “Consumer Decision-Making in Retail Investment Services,” Report to the European Commission.

Collinson, P. (2012): “FSA Ban on Commission-Based Selling Sparks "Death of Salesman" Fears,” The Guardian, December 30.

Crawford, V. P., and J. Sobel (1982): “Strategic Information Transmission,” Econometrica, 50, 1431–1451.

Dufwenberg, M., Jr. and M. Dufwenberg, Sr. (2016): “Lies in Disguise – A Theoretical Analysis of Cheating,” CESifo Working Paper Series 6208, CESifo Group Munich.

Dye, R. A. (1985): “Disclosure of Nonproprietary Information,” Journal of Accounting Research, 23, 123–145.

Dziuda, W. (2011): “Strategic Argumentation,” Journal of Economic Theory, 146, 1362–1397.

Ellingsen, T., M. Johannesson, S. Tjøtta, and G. Torsvik (2010): “Testing Guilt Aversion,” Games and Economic Behavior, 68, 95–107.

Fehr, E., and K. M. Schmidt (1999): “A Theory of Fairness, Competition and Cooperation,” Quarterly Journal of Economics, 114, 817–868.

Geanakoplos, J., D. Pearce, and E. Stacchetti (1989): “Psychological Games and Sequential Rationality,” Games and Economic Behavior, 1, 60–79.

Gneezy, U. (2005): “Deception: The Role of Consequences,” American Economic Review, 95, 384–394.

Gneezy, U., A. Kajackaite, and J. Sobel (2018): “Lying Aversion and the Size of the Lie,” American Economic Review, 108(2), 419–453.

Gruber, J., J. Kim, and D. Mayzlin (1999): “Physician Fees and Procedure Intensity: The Case of Cesarean Delivery,” Journal of Health Economics, 18, 473–490.

Guerra, G., and D. J. Zizzo (2004): “Trust Responsiveness and Beliefs,” Journal of Economic Behavior and Organization, 55, 25–30.

Inderst, R., and M. Ottaviani (2012): “How (not) to Pay for Advice: A Framework for Consumer Financial Protection,” Journal of Financial Economics, 105, 393–411.

Inderst, R., and M. Ottaviani (2013): “Sales Talk, Cancellation Terms, and the Role of Consumer Protection,” Review of Economic Studies, forthcoming.

Kartik, N. (2009): “Strategic Communication with Lying Costs,” Review of Economic Studies, 76, 1359–1395.

Kartik, N., M. Ottaviani, and F. Squintani (2007): “Credulity, Lies, and Costly Talk,” Journal of Economic Theory, 134, 93–116.

Khalmetski, K. (2016): “Testing Guilt Aversion with an Exogenous Shift in Beliefs,” Games and Economic Behavior, 97, 110–119.

Khalmetski, K., A. Ockenfels, and P. Werner (2015): “Surprising Gifts: Theory and Laboratory Evidence,” Journal of Economic Theory, 159, 163–208.

Khalmetski, K., B. Rockenbach, and P. Werner (2017): “Evasive Lying in Strategic Communication,” Journal of Public Economics, 156, 59–72.

Khalmetski, K., and D. Sliwka (2017): "Disguising Lies - Image Concerns and Partial Lying in Cheating Games,” CESifo Working Paper Series 6347, CESifo Group Munich.

Loginova, U. (2012): “Four Essays on Strategic Communication,” Doctoral dissertation, Columbia University.

McGuire, T. G. (2000): “Physician Agency,” In Handbook of Health Economics, Volume 1A, edited by Anthony J. Culyer and Joseph P. Newhouse, 461–536. Amsterdam; New York and Oxford: Elsevier Science, North-Holland.

Reuben, E., P. Sapienza, and L. Zingales (2009): “Is Mistrust Self-Fulfilling?” Economics Letters, 104(2), 89–91.

Vanberg, C. (2008): “Why do People Keep their Promises? An Experimental Test of Two Explanations,” Econometrica, 76, 1467–1480.