

# Endogenous Persuasion with Rational Verification

Mike FELGENHAUER\*

July 16, 2017

## Abstract

This paper studies a situation in which a sender tries to persuade a receiver with evidence that is generated via public or private experimentation. The sender can design the properties of the experiments. The receiver chooses whether to verify the design. The paper finds that (i) communication breaks down under public experimentation if there is no restriction on the experiment's design, (ii) persuasion is possible under private experimentation and (iii) in any equilibrium with persuasion under private experimentation, the sender's argument contains cheap talk or hard evidence, each with a positive probability.

**Keywords:** Experimentation, persuasion, cheap talk, information acquisition.

**JEL classification:** D82, D83

---

\*Plymouth University, Cookworthy Building, Drake Circus, Plymouth, Devon, PL4 8AA, UK; email: mike.felgenhauer@plymouth.ac.uk.

# 1 Introduction

This paper investigates a situation where a sender (e.g., a prosecutor, a student or an interested party) tries to persuade a receiver (e.g., a judge, a teacher or a politician) to take an action (e.g., the conviction of a defendant, a better mark, a policy). For the receiver, the optimal action is in the sender’s favour in one state and it is against the sender in another state. There is a threshold belief above which the receiver will take the sender’s preferred action. Her prior belief is below the threshold.

Evidence (or an argument) used for persuasion has an inherent meaning and is not per se cheap talk. In this paper evidence consists of a recommendation and reasoning supporting the recommendation. As an illustration suppose a sender provides evidence (or argues) “This is the right decision, because ...”.<sup>1</sup> The recommendation “This is the right decision” is straightforward. The reasoning following the “because” on the other hand may be involved and require effort to understand. The receiver may not pay attention to the sender’s reasoning. It can be tedious for the receiver to grasp its meaning, as for example reading a paper or listening to a presentation. The receiver may just follow the recommendation and think about something more pleasant instead of trying to understand the reasoning.<sup>2</sup> The sender may try to exploit the receiver’s incentive to ignore the reasoning. Instead of searching for hard to find favorable high quality evidence, he may search for easy to find favorable superficial reasoning that does not have sufficient persuasive power to convince a receiver who pays attention.

The evidence used for persuasion in this paper stems from experimentation. An experimental outcome correctly predicts the decision-relevant state of the world with

---

<sup>1</sup>Straightforward recommendations followed by some sort of reasoning are common in practice. Recent examples are the persuasion attempts of the remain campaign (“Vote for remain, because ...”) and the leave campaign (“Vote for leave, because ...”) in the Brexit debate. Other examples are “I am innocent, because ...”, “TTIP is good, because ...” etc.

<sup>2</sup>Not paying attention matters in practice. A drastic example is jurors falling asleep during a trial. King (1996) finds that sixty nine percent of 562 judges who returned her survey reported cases (during the last three years) where jurors had fallen asleep. Another example is Wolfgang Schäuble, the German Minister of Finance, playing Sudoku in the Bundestag during a debate on financial aid worth billions of Euros for Greece (Darnstädt, 2012). In the context of the recent presidential election in the U.S. I also suspect that many voters did not pay attention to the winner’s arguments during the electoral campaign.

endogenous error probabilities. The sender designs the error probabilities of each experiment that he runs. The outcome of an experiment is interpreted as a recommendation for a decision. The experiment's design determines the quality of the information that the experiment generates. The design of the experiment, i.e., the probabilities with which it correctly predicts the states of the world, is viewed as the reasoning. The evidence of an experiment, hence, consists of the outcome of this experiment and the experiment's design. If the sender uses an experiment's outcome for persuasion in his message to the receiver, then the receiver directly observes the revealed outcome. The receiver then chooses whether to verify the message. If she verifies the message, then she learns the design of the experiment with which the presented outcome was generated.<sup>3</sup> Verifying the message is costly, whereas no verification is costless.

A message that contains an outcome from an uninformative experiment is interpreted as cheap talk. As in the cheap talk literature, such messages do not have an inherent meaning in the sense that they are not directly related to a state of the world. Deductions regarding the state of the world can only be made via equilibrium behavior. The relation that a message has to the state is, hence, not verifiable. However, it is verifiable that the nature of the message is cheap talk: A receiver (or a judge) studying the message in detail realizes the lack of an inherent meaning.

The paper distinguishes between public and private experimentation. Experimentation is public if, starting from a situation without information, the sender designs one single experiment and then discloses its outcome to the receiver. For example, a trial can be viewed as a public experiment where a prosecutor tries to persuade a judge to convict a defendant (Kamenica and Gentzkow, 2011). The prosecutor can, e.g., influence the error probabilities by choosing the structure of the examination of witnesses. In contrast, experimentation is private if the sender may secretly run multiple experiments and selectively reveal an outcome. The sender's decision whether to run a further experiment with a design of his choice is contingent on the experimentation

---

<sup>3</sup>E.g., by thoroughly reading a paper or by carefully listening to a presentation, the receiver can assess the quality of the presented argument.

history. For example, thought experiments that yield logical arguments are run sequentially in private and the results can be revealed selectively (Felgenhauer and Schulte, 2014). The sender can influence the experiments' quality by choosing the conceptual framework from which specific assumptions are drawn. The receiver under private experimentation does not directly observe the experimentation history even if she verifies the message, as the sender may privately have collected additional evidence. Verification of the sender's message under private experimentation can be viewed as learning the face value of the sender's recommendation.

The paper finds that persuasion is impossible under public experimentation if the design of the experiments is flexible.<sup>4</sup> This result is due to two commitment problems.<sup>5</sup> First, the receiver cannot commit to ignore information that she learns: if her belief is pushed above the threshold to choose the sender's preferred action, it is sequentially rational to take that action. Second, the sender cannot commit to designing an experiment which is sufficiently valuable for the receiver to justify the verification cost.<sup>6</sup> If the sender anticipates that the receiver verifies the message, then the sender's best experiment renders the receiver indifferent between both actions if she takes the sender's preferred action and it renders her certain that the state is bad when she chooses against the sender. Such an experiment maximizes the persuasion probability. However, this experiment has no value for the receiver: without it, she would choose against the sender; with it, either she is indifferent to take the sender's preferred action, or she chooses against the sender. Thus, the receiver does not have an incentive to pay the cost to learn the experiment's design in the first place. As the receiver cannot commit to punishing the sender for designing a low quality experiment, the sender cannot com-

---

<sup>4</sup>Kamenica and Gentzkow (2011) analyze persuasion under public experimentation with complete design flexibility, but without endogenous verification. They find that the design flexibility can give substantial power to the sender. The current paper shows that this point is not robust if endogenous verification matters. In this case the design flexibility is the cause of the communication breakdown.

<sup>5</sup>I thank an anonymous referee for suggesting this intuition.

<sup>6</sup>For persuasion to occur in equilibrium it is necessary that the receiver (sometimes) verifies the sender's message. Suppose the opposite. The sender's best response to no verification is to run an uninformative experiment that always yields a favorable outcome. By equilibrium reasoning the receiver anticipates this behavior and, therefore, a favorable outcome does not push her belief above the threshold.

mit to design an experiment that is sufficiently valuable for the receiver to justify the verification costs. Key for the communication breakdown under public experimentation is the flexibility of the experiment's design. If exogenous restrictions force the sender to choose among highly informative experiments, then persuasion becomes possible.<sup>7</sup>

Persuasion is possible if experimentation occurs in private given that experimentation and verification costs are sufficiently low. With private experimentation the design flexibility can be reduced endogenously. For example, if the sender chooses an unexpected (off-the-equilibrium path) design for persuasion, then the receiver who verifies the message may believe that the sender privately collected additional high quality evidence with adverse outcomes that are not revealed. Given these beliefs the receiver's sequentially rational action upon observing an unexpected design is against the sender. The sequentially rational punishment given such beliefs can force the sender to choose a high quality design that makes it worthwhile for the receiver to pay the verification costs. There is no pure strategy equilibrium with persuasion, but there can be mixed strategy equilibria. The paper derives the sender preferred equilibrium. This equilibrium has a simple structure. The sender first runs an informative experiment. If it's outcome is positive, then it is used for persuasion. Otherwise, the sender mixes between stopping experimentation unsuccessfully and running an uninformative experiment that always yields a positive outcome. The receiver mixes between verification of a positive outcome and no verification. In the former case she is only persuaded if the positive outcome stems from the informative experiment. In the latter case she is only persuaded by positive outcomes.

The different results under public and private experimentation are due to off-the-equilibrium path beliefs. If the receiver observes an off-the-equilibrium path design under private experimentation, then it is not obvious whether she has access to all the information that is available, as the sender may have run additional experiments

---

<sup>7</sup>In practice procedural rules in public trials that specify what arguments can be presented in court or what type of questions are admissible that a prosecutor may ask exogenously reduce the degrees of freedom regarding the design of public experiments. A continuous adjustment of the flexibility may also not be possible by the nature of the evidence generation process.

without revealing the outcomes. This allows off-the-equilibrium path beliefs that make a punishment for using a low quality design sequentially rational and can force the sender to choose a high quality design. Under public experimentation, in contrast, such off-the-equilibrium path beliefs are not possible. If the receiver pays the verification costs, then she has access to all the information that is available, which determines her beliefs and she cannot commit to choose against what she thinks is best. If the receiver verifies a favorable outcome, then the sender exploits the lack of credible punishment by choosing a design that just persuades her.

The paper provides a simple model that combines aspects of the literature on persuasion and cheap talk. It shows that in any equilibrium with persuasion under private experimentation the sender designs informative and uninformative experiments. The former yield hard decision-relevant evidence. The latter yield outcomes that do not have an inherent meaning. An outcome that does not have an inherent meaning can be viewed as cheap talk. In any equilibrium with persuasion the sender sends messages that contain cheap talk with a positive probability. Cheap talk is, therefore, not entirely crowded out by valuable hard evidence if the receiver may choose not to listen.

## 2 Literature

There is an extensive literature on persuasion (e.g., Jovanovic, 1982, Milgrom and Roberts, 1986, Glazer and Rubinstein, 2001, 2004, 2006). Many papers assume that the sender is exogenously endowed with private decision-relevant hard information. If arguments are viewed as decision-relevant evidence, as in the present paper, then it is natural that they have to be acquired in the first place (as, e.g., in Celik, 2003, Brocas and Carillo, 2007, and Henry, 2009). In this branch of the literature, persuasion via experimentation with a *flexible design* of the experiments has received a growing attention. As arguments play a role in many decision environments and the properties of arguments can be influenced, it is relevant to understand how the design flexibility affects decision making. The papers with these features that are most closely related

are Kamenica and Gentzkow (2011) and Felgenhauer and Loerke (2016).

Kamenica and Gentzkow (2011), in the following KG, study the best experimental design from a sender's perspective with which he can persuade a receiver under public experimentation. By being able to freely design the public experiment the sender has substantial influence in the decision making. The current paper on the other hand illustrates that the complete design flexibility does not necessarily imply power for the sender. It shows that the flexibility causes a communication breakdown in a parsimonious version of KG's model where in addition the receiver may rationally not pay attention to the sender's persuasion attempt. The complete design flexibility harms the sender under public experimentation with endogenous verification. Gentzkow and Kamenica (2014) show that KG is robust to a broad class of cost functions. The present paper finds that this result is not robust to endogenous verification, even if verification costs are arbitrarily small but positive.

Felgenhauer and Loerke (2016), in the following FL, analyze sequential private experimentation with a flexible design of the experiments.<sup>8</sup> Arguments are typically acquired in private and, hence, private experimentation may be viewed as more natural than public experimentation. The modelling approach here is similar to FL, but the current model considers endogenous verification. Their paper also derives the sender preferred equilibrium. In this equilibrium the sender runs only one experiment. Here, the sender starts with the same experiment as in FL and stops successfully if the outcome is positive. If the outcome of the first experiment is adverse on the other hand, then the sender mixes between running an uninformative experiment and stopping unsuccessfully.

Due to the receiver's endogenous verification, the equilibria with persuasion under private experimentation exhibit features of the persuasion and the cheap talk literature. In the cheap talk literature (e.g., Crawford and Sobel, 1982) messages do not have an inherent meaning, but obtain their meaning through equilibrium behavior. In any

---

<sup>8</sup>Felgenhauer and Schulte (2014) study sequential private experimentation with an exogenous precision of the experiments.

equilibrium with persuasion in the present paper, the sender sends messages containing outcomes from informative and uninformative experiments. The former can be viewed as hard evidence and the latter as cheap talk. Kartik (2009), Esó and Galambos (2013), Dzuida (2012) and Dewatripont and Tirole (2005) also combine features of both strands of the literature. In contrast to the present paper they do not consider endogenous evidence production.

Kartik (2009) studies how a sender tries to influence a receiver's decision by sending a message about his exogenously given type. Lying about the type is costly. Kartik (2009) shows that there are equilibria, in which "low" types separate and "high" types pool. In an extension the sender has the option to costlessly withhold information and the receiver observes whether information is withheld. This setting embeds both verifiable disclosure and cheap talk as limiting cases. Esó and Galambos (2013) extend Crawford and Sobel's (1982) cheap talk model such that the sender has the option to send hard evidence and soft signals about his exogenously given type and the receiver observes the message type. In equilibrium there are sender types sending soft signals and others sending hard evidence. These papers do not focus on whether the receiver pays attention to the sender's persuasion attempt.

Dzuida (2012) introduces an exogenous detection probability into a setting where a sender sends a cheap talk message to a receiver. If a lie is always detected, then all types are truthfully revealed, which mirrors full information revelation in persuasion games. If a lie is never detected, then the game is one with pure cheap talk and there is no information transmission. For each exogenous detection probability in between, there is some information revelation. In the present paper the argument that can be used can either be cheap talk or hard evidence (which is absent in Dzuida, 2012) and it is the sender's choice what type of evidence to send. Therefore, the present paper takes into account the different message spaces in the cheap talk and the persuasion literature.

Dewatripont and Tirole (2005) study communication as a moral hazard problem in a team with endogenous verification. The sender is endowed with evidence that may or

may not be strongly in his favor. The probability that the receiver correctly observes the sender’s evidence depends on the receiver’s costly message elaboration effort and the sender’s costly persuasion effort.<sup>9</sup> Persuasion is in principle possible. Endogenous verification is similar to the current model, but the present paper has a different focus. It studies whether persuasion is possible under private and public experimentation given that the design of the experiments is flexible.

### 3 Assumptions

#### 3.1 Preferences

There is a state of the world  $\omega \in \Omega$ , with  $\Omega = \{\omega_1, \omega_2\}$ . Ex ante both states are equally likely. The receiver chooses action  $a \in A$ , with  $A = \{a_1, a_2\}$ , and obtains gross utility

	$\omega = \omega_1$	$\omega = \omega_2$
$a = a_1$	1	$1 - p_d$
$a = a_2$	$p_d$	1

with  $p_d \in (\frac{1}{2}, 1)$ . The receiver’s optimal action is  $a = a_1$  if and only if her posterior that the state is  $\omega = \omega_1$  is weakly above her “threshold of doubt”  $p_d$ . As ex ante both states are equally likely, she chooses  $a = a_2$  if she does not have access to additional information.

The sender prefers action  $a = a_1$  in each state. His gross utility is 1 if  $a = a_1$  and it is 0 if  $a = a_2$ .

Experimentation and verification costs have to be subtracted from the respective gross utilities.

#### 3.2 Experimentation

The sender may run experiments that generate signals about  $\omega$ . An experiment  $\tau$  yields outcome  $\sigma_\tau \in \{s_1, s_2\}$ . Outcome  $s_1$  is called a “positive outcome” and  $s_2$  an

---

<sup>9</sup>Similarly, in Perez-Richet and Prady (2012), who study complexity in persuasion, the receiver’s learning costs are influenced by the sender’s choice of the complexity level.

“adverse outcome”. The design of an experiment  $\tau$  is  $\pi_\tau = (\pi_\tau(s_1 | \omega_1), \pi_\tau(s_2 | \omega_2))$ , with  $\pi_\tau(s_j | \omega_j) = \text{prob}\{\sigma_\tau = s_j | \omega = \omega_j\}$ ,  $\pi_\tau(s_j | \omega_j) \in [0, 1]$ ,  $j \in \{1, 2\}$ . Let  $\pi_\tau(s_1 | \omega_1) \geq 1 - \pi_\tau(s_2 | \omega_2)$ , i.e., a positive outcome  $\sigma_\tau = s_1$  occurs with a higher probability if  $\omega = \omega_1$  than if  $\omega = \omega_2$ . Denote by  $(\sigma_\tau, \pi_\tau)$  the “evidence” of experiment  $\tau$ . The evidence of some experiment  $\tau$ , thus, contains its outcome and its design. Denote by  $h = \{(\sigma_\tau, \pi_\tau)\}_{\tau=1, \dots, t}$  the experimentation history after the first  $t$  experiments.

Unless stated otherwise, experimentation costs are as follows. Running an uninformative experiment  $\tau$  with design  $\pi_\tau = (1, 0)$ ,  $\pi_\tau = (0, 1)$  or  $\pi_\tau = (\frac{1}{2}, \frac{1}{2})$  costs  $c_U \in (0, c_I)$ , with  $c_I \in (0, \frac{1}{2})$ . Running an informative experiment  $\tau$  costs  $c_I$ .

Under public experimentation the sender runs a single experiment and then stops after either outcome as in KG.<sup>10</sup>

The experimentation technology under private experimentation is as in FL. The sender can run as many experiments as desired. He chooses the design of each experiment that he runs and this choice is history dependent. The decision to continue experimenting or to stop running experiments is also history dependent.

### 3.3 Messages and verification

Denote the sender’s message by  $\hat{m} = \sigma_\tau$ , where  $\sigma_\tau$  is the outcome of some experiment  $\tau$ .<sup>11</sup> A message  $\hat{m} = \sigma_\tau$  is “feasible” if  $\sigma_\tau$  is the outcome of an experiment  $\tau$  that the sender ran. The sender may only send a feasible message.

The receiver observes message  $\hat{m} = \sigma_\tau$ . She may verify the message. If the receiver does not verify  $\hat{m}$ , then she exclusively observes  $\hat{m} = \sigma_\tau$ . Verifying  $\hat{m}$  corresponds to learning the design  $\pi_\tau$  of experiment  $\tau$  that yielded the presented outcome  $\sigma_\tau$ . If she verifies  $\hat{m} = \sigma_\tau$ , then she observes the evidence  $(\sigma_\tau, \pi_\tau)$ .<sup>12</sup> The interpretation is that

---

<sup>10</sup>The analysis can be extended to sequential public experimentation. By an analogous argument it can be shown that persuasion is impossible in this extension. The proof is available upon request.

<sup>11</sup>Under private experimentation the sender may run more than one experiment. As in FL, the model could, thus, allow him to reveal more than one outcome. In order to keep the notation simple the focus here is on messages that contain one outcome. This simplification does not affect the results. The proof is available upon request.

<sup>12</sup>Note that the receiver does not learn about other experiments that the sender may have run. If the message is interpreted as a presented argument, like a paper containing a theoretical model, then

by learning the design  $\pi_\tau$  of an experiment  $\tau$  that yielded an outcome  $\sigma_\tau$  she directly learns the meaning of this outcome. If she does not verify the message, then she can only deduce (a possibly less accurate) meaning from equilibrium behavior. Verifying message  $\hat{m}$  costs  $c_v > 0$ .

### 3.4 Timing and strategies

The sender moves first. Under public experimentation a pure strategy for the sender specifies the design of the experiment that he runs. The sender's message is mechanical in the sense that it contains the outcome of this experiment. Under private experimentation the sender's strategy specifies his behavior at each experimentation history  $h$  that he may observe. At each such history the sender may either continue experimenting with an experiment with a history dependent design, or he may stop experimenting and send a feasible message.

The receiver moves after observing the sender's message. At each such message  $\hat{m} = \sigma_\tau$  the receiver may either not verify  $\hat{m}$  and choose an action  $a(\sigma_\tau)$ , or she may verify  $\hat{m}$  and choose an action  $a(\sigma_\tau, \pi_\tau)$ . For each message that the receiver may observe her pure strategy, thus, specifies (i) whether to verify the message, (ii) an action for each message that she does not verify and (iii) an action for each underlying evidence of each message that she verifies.

### 3.5 Equilibrium concept

The equilibrium concept is weak perfect Bayesian equilibrium. In equilibrium the players' strategies are sequentially rational. On the equilibrium path beliefs are Bayesian. If the receiver observes some off-the-equilibrium path evidence  $(\sigma_\tau, \pi_\tau)$  (respectively, message  $\hat{m} = \sigma_\tau$ ), then it is assumed that she only considers experimentation histories  $h$  as possible with  $(\sigma_\tau, \pi_\tau) \in h$ . Off-the-equilibrium path beliefs have to be such that

---

by paying attention, e.g., by carefully reading the paper, the receiver learns the details of the model that is presented. However, by studying a particular model, she naturally does not learn anything about other thought experiments that the sender may have conducted and that he did not present.

the sender cannot signal what he does not know.<sup>13</sup>

An equilibrium in which the receiver takes the sender's preferred action  $a = a_1$  with a positive probability on the equilibrium path is called an *equilibrium with persuasion*.

## 4 Public experimentation

In the following it is shown that there is no equilibrium with persuasion under public experimentation if the sender does not face restrictions regarding the design of his experiments. Key for the communication breakdown is that the receiver, once she verifies the message, observes all the evidence  $(\sigma_\tau, \pi_\tau)$  that is available and, therefore, regardless of whether this evidence is on or off-the-equilibrium path, this receiver's Bayesian update that the state is  $\omega = \omega_1$  is  $prob\{\omega = \omega_1 \mid (\sigma_\tau = s_1, \pi_\tau)\} = \frac{\pi_\tau(s_1|\omega_1)}{\pi_\tau(s_1|\omega_1) + (1 - \pi_\tau(s_2|\omega_2))}$  after observing a positive outcome and  $prob\{\omega = \omega_1 \mid (\sigma_\tau = s_2, \pi_\tau)\} = \frac{1 - \pi_\tau(s_1|\omega_1)}{1 - \pi_\tau(s_1|\omega_1) + \pi_\tau(s_2|\omega_2)}$  after observing an adverse outcome.

In any potential equilibrium with persuasion there are two commitment problems. First, the receiver cannot commit to choose against the sender if the observed evidence is such that her threshold of doubt is just passed. Second, anticipating the receiver's commitment problem allows the sender to design any informative experiment with a relatively low predictive power. The sender cannot commit to run a high quality experiment that rarely yields a positive outcome. The low predictive power is not sufficiently valuable for the receiver to justify the verification costs. Hence, there is no equilibrium with persuasion in which the receiver verifies messages. In the appendix it is shown that there is also no equilibrium with persuasion without verification.

**Proposition 1** *There is no pure and no mixed-strategy equilibrium with persuasion under public experimentation.*

---

<sup>13</sup>Consider, e.g., public experimentation and suppose the receiver verifies the message. If the sender runs an off-the-equilibrium path experiment, then this assumption implies that this receiver cannot have arbitrary beliefs regarding  $\omega$ , as the sender does not know more about  $\omega$ . Therefore, this receiver's off-the-equilibrium path beliefs regarding  $\omega$  have to be Bayesian based exclusively on the experimentation history.

The commitment problems are now analyzed in turn in a potential equilibrium with persuasion in which the receiver verifies some messages with a positive probability on the equilibrium path. It is shown that the receiver does not have an incentive to verify messages, yielding a contradiction.

Consider the receiver's commitment problem. If she verifies  $\hat{m} = s_1$ , then a comparison of the payoffs from each action  $a$  shows that it is sequentially rational to choose  $a = a_1$  if and only if

$$prob\{\omega = \omega_1 \mid (\sigma_\tau = s_1, \pi_\tau)\} = \frac{\pi_\tau(s_1 \mid \omega_1)}{\pi_\tau(s_1 \mid \omega_1) + (1 - \pi_\tau(s_2 \mid \omega_2))} \geq p_d.$$

Hence, a receiver who observes evidence  $(\sigma_\tau = s_1, \pi_\tau)$  cannot commit to choose against the sender as soon as her threshold of doubt  $p_d$  is passed. If the receiver verifies  $\hat{m} = s_2$ , then it is optimal to choose  $a = a_2$  for each design that the corresponding experiment may have.<sup>14</sup> As verification is costly, it follows that it is sequentially rational for the receiver to choose  $a = a_2$  without verification upon observing  $\hat{m} = s_2$ . The receiver may potentially be persuaded in equilibrium upon observing  $\hat{m} = s_1$  (with and/or without verification), since a positive outcome  $s_1$  is more likely if  $\omega = \omega_1$  than if  $\omega = \omega_2$ .

Next, consider the sender's commitment problem. Anticipating the receiver's sequentially rational behavior, the sender in equilibrium designs any informative experiment  $\tau$  that he runs with a positive probability such that the probability of a positive outcome  $\frac{1}{2}\pi_\tau(s_1 \mid \omega_1) + \frac{1}{2}(1 - \pi_\tau(s_2 \mid \omega_2))$  is maximized subject to  $prob\{\omega = \omega_1 \mid (\sigma_\tau = s_1, \pi_\tau)\} \geq p_d$ . The solution to this maximization problem is  $\pi_\tau(s_1 \mid \omega_1) = 1$  and  $\pi_\tau(s_2 \mid \omega_2) = \frac{2p_d - 1}{p_d}$ .<sup>15</sup> This experimental design implies  $prob\{\omega = \omega_1 \mid (\sigma_\tau = s_1, \pi_\tau)\} = p_d$ . That is, the predictive power of a positive outcome is such that the threshold of doubt

---

<sup>14</sup>Due to the assumption  $\pi_\tau(s_1 \mid \omega_1) \geq 1 - \pi_\tau(s_2 \mid \omega_2)$ , which implies that a positive outcome is more likely in state  $\omega = \omega_1$  than in state  $\omega = \omega_2$ , we have  $prob\{\omega = \omega_1 \mid (\sigma_\tau = s_2, \pi_\tau)\} = \frac{1 - \pi_\tau(s_1 \mid \omega_1)}{1 - \pi_\tau(s_1 \mid \omega_1) + \pi_\tau(s_2 \mid \omega_2)} \leq \frac{1 - \pi_\tau(s_1 \mid \omega_1)}{1 - \pi_\tau(s_1 \mid \omega_1) + 1 - \pi_\tau(s_1 \mid \omega_1)} = \frac{1}{2} < p_d$ .

<sup>15</sup>This part of the analysis is analogous to KG. At the optimum we have  $prob\{\omega = \omega_1 \mid (\sigma_\tau = s_1, \pi_\tau)\} = \frac{\pi_\tau(s_1 \mid \omega_1)}{\pi_\tau(s_1 \mid \omega_1) + (1 - \pi_\tau(s_2 \mid \omega_2))} = p_d$ . Otherwise, the sender could increase the probability of a positive outcome by increasing  $\pi_\tau(s_1 \mid \omega_1)$  or decreasing  $\pi_\tau(s_2 \mid \omega_2)$ , without violating the constraint.

is just passed such that the receiver who verifies a positive outcome from this experiment is just persuaded to choose  $a = a_1$ . Sequential rationality requires that any other experiment  $\tau'$  that the sender runs with a positive probability is uninformative, with  $prob\{\omega = \omega_1 \mid (\sigma_{\tau'} = s_1, \pi_{\tau'})\} = \frac{1}{2}$ . The sender cannot commit to run higher quality experiments.

In order to complete the argument that there is no equilibrium with persuasion as described above, it is now shown that the receiver's gross payoff in the hypothetical equilibrium is the same regardless of whether she verifies a message with a positive outcome. As verification is costly, the receiver who verifies a message has a profitable deviation not to verify  $\hat{m} = s_1$ .

Figure 1 shows the gross payoff of the receiver as a function of the posterior  $\mu_1$  that the state is  $\omega = \omega_1$ . For all  $\mu_1 \in [0, p_d)$  it is optimal for the receiver to choose  $a = a_2$ . Her gross payoff in this region is  $\mu_1 p_d + (1 - \mu_1)$ , which is linearly decreasing in  $\mu_1$ . For all  $\mu_1 \in [p_d, 1]$  it is optimal for the receiver to choose  $a = a_1$ . Her gross payoff in this region is  $\mu_1 + (1 - \mu_1)(1 - p_d)$ , which is linearly increasing in  $\mu_1$ . We have  $\mu_1 p_d + (1 - \mu_1) = \mu_1 + (1 - \mu_1)(1 - p_d)$  at  $\mu_1 = p_d$ .

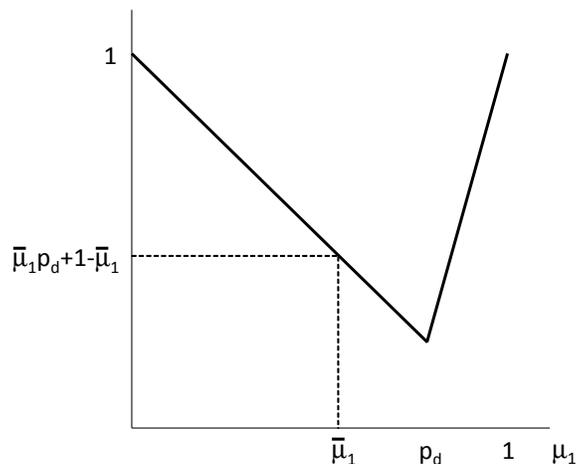


Figure 1: Receiver's gross value function.

Consider a receiver who observes a message  $\hat{m} = s_1$ , but does not verify it. A receiver who does not verify the message does not observe the experiment's design. From her perspective a positive outcome may stem from an informative experiment with design

$(1, \frac{2p_d-1}{p_d})$  or from an uninformative experiment. According to the sender's commitment problem the informative experiment  $\tau$  implies  $prob\{\omega = \omega_1 \mid (\sigma_\tau = s_1, \pi_\tau)\} = p_d$  and the uninformative experiment  $\tau'$  implies  $prob\{\omega = \omega_1 \mid (\sigma_{\tau'} = s_1, \pi_{\tau'})\} = \frac{1}{2}$ . Denote her posterior belief by  $prob\{\omega = \omega_1 \mid \hat{m} = s_1\} \equiv \bar{\mu}_1$ . Bayesian plausibility requires  $\bar{\mu}_1 \in [\frac{1}{2}, p_d]$  if  $\hat{m} = s_1$ . Her payoff is equal to the gross value function evaluated at  $\bar{\mu}_1$ , i.e., it is  $\bar{\mu}_1 p_d + (1 - \bar{\mu}_1)$ , as illustrated in Figure 1.

Consider next the receiver who verifies a message containing a positive outcome  $\hat{m} = s_1$ . Bayesian plausibility requires that the expected posterior for the receiver who verifies  $\hat{m} = s_1$  is equal to  $\bar{\mu}_1$  before she verifies it. In an equilibrium with persuasion her Bayesian updates that the state is  $\omega = \omega_1$  are  $p_d$  if the experimental design is informative and  $\frac{1}{2}$  if the experiment is uninformative, with  $\frac{1}{2} < p_d$ . As all combinations of the posteriors and the corresponding gross payoffs are on the downward sloping (linear) part of the gross value function this implies that the expected gross utility of the receiver who verifies the message  $\hat{m} = s_1$  is equal to the gross value function evaluated at belief  $\bar{\mu}_1$ , i.e., it is equal to  $\bar{\mu}_1 p_d + (1 - \bar{\mu}_1)$ .

The receiver's expected gross utility from verifying a message containing a positive outcome is, thus, equal to her payoff if she does not verify it.<sup>16</sup> Since paying attention is costly, it is better for the receiver not to verify a message with a positive outcome. Hence, there is no equilibrium with persuasion in which the receiver sometimes verifies  $\hat{m} = s_1$  on the equilibrium path.

The flexibility that the sender has regarding the design of the experiment's error probabilities excludes equilibria with persuasion in which the sender and the receiver mix. The flexibility allows the sender to make the receiver who verifies the message

---

<sup>16</sup>Key for this point is that a positive outcome of the informative experiment leads to a posterior equal to  $p_d$ . As an illustration suppose hypothetically that a positive outcome of the informative experiment instead yields a posterior  $\mu'_1$  above  $p_d$ . In this case the following could occur. We could still have  $\bar{\mu}_1 < p_d$ , implying that the receiver who does not verify chooses  $a = a_2$  and her payoff is equal to the gross value function (on the downward sloping part) evaluated at  $\bar{\mu}_1$ . The receiver who verifies the message, however, obtains a higher ex ante gross payoff. Consider a line between the gross value function (on the downward sloping part) evaluated at  $\frac{1}{2}$  and the gross value function (on the upward sloping part) evaluated at  $\mu'_1 (> p_d)$ . The former corresponds to the gross payoff from choosing  $a = a_2$  if the experiment is uninformative and the latter to the gross payoff from choosing  $a = a_1$  if it is informative. The ex ante gross payoff of the receiver who verifies the message is on this line evaluated at  $\bar{\mu}_1$ , which clearly is strictly above the payoff of the receiver who does not verify.

indifferent to choose in his favor after a favorable recommendation, which in turn depresses the ex ante value of the message too much to justify the costs of verifying it. The following example shows that cases can be constructed where an equilibrium with persuasion exists if the degrees of freedom regarding the design of the experiments are reduced.

**Example 1** Suppose the sender may exclusively run an experiment  $\tau$  with design  $\pi_\tau = (1, 0)$  or an experiment  $\tau$  with design  $\pi_\tau = (1, 1)$ . An experiment with  $\pi_\tau = (1, 0)$  yields  $\sigma_\tau = s_1$  regardless of the state and an experiment with  $\pi_\tau = (1, 1)$  perfectly predicts the state. Let  $c_v \in (0, 3p_d - 2p_d^2 - 1]$ . There is an equilibrium with persuasion in mixed strategies with the following properties: The sender mixes between running an experiment with design  $\pi_\tau = (1, 0)$  and an experiment with  $\pi_\tau = (1, 1)$ . The probability that the sender runs the informative experiment decreases in the verification costs  $c_v$  and it approaches one if  $c_v$  goes to zero. The receiver mixes regarding verification of a message containing a single positive outcome and chooses  $a = a_1$  if she does not verify the evidence. The verification probability of a message in the sender's favor is greater than  $1/2$ . It increases in the costs of running the informative experiment and it decreases in the costs of conducting the uninformative experiment.  $\square$

In Example 1 the sender has less flexibility. If he runs an informative experiment, then this experiment mirrors the state of the world. The receiver who observes the outcome of such an experiment and chooses the corresponding optimal action obtains a high gross payoff. Before verifying a message in favor of the sender, the receiver has a benefit from verifying it that has the potential to justify the verification costs. The equilibrium probability that the message is informative has to be such that the receiver is indifferent to verify it. The sender on the other hand faces a trade-off between obtaining a positive outcome from the uninformative experiment with certainty that only persuades the receiver who does not verify the message and obtaining a positive

outcome with probability  $1/2$  that persuades the receiver regardless of whether she verifies the message. The receiver’s probability to verify the message has to make the sender indifferent between running either experiment.

If the design of the sender’s experiment is flexible, then the lack of commitment power on the receiver’s part is important for the communication breakdown. As an illustration suppose that the experiment’s design is endogenous again, but that the receiver can costlessly commit to a rule: accept after verification if and only if the posterior is above a threshold  $p$ . The other assumptions are maintained. It can be shown that the receiver’s optimal threshold posterior is  $p = 1$ . As in Example 1, the sender mixes between running an uninformative experiment with  $\pi_\tau = (1, 0)$  and an experiment with  $\pi_\tau = (1, 1)$  that perfectly predicts the state, and persuasion occurs in equilibrium.<sup>17</sup>

Gentzkow and Kamenica (2014) study the robustness of KG’s concavification approach for the derivation of the sender’s optimal design of the experiment to diverse (experimentation) cost functions.<sup>18</sup> They describe a family of cost functions that is compatible with KG’s concavification approach. For these cost functions the receiver benefits from communication. In the following it is shown that this result is not robust to endogenous verification, even if the cost of verification is arbitrarily small but positive. Suppose that an increase of the experiment’s informativeness according to the Blackwell criterion implies higher experimentation costs. Uninformative experiments are assumed to be equally expensive and cheapest.<sup>19</sup>

Consider a hypothetical equilibrium with persuasion. The sender’s optimal informative experiment still makes the receiver indifferent between  $a_1$  and  $a_2$  if she knows that  $\hat{m} = s_1$  stems from this experiment.<sup>20</sup> Suppose the opposite. The sender could deviate

---

<sup>17</sup>As in Example 1 verification costs have to be sufficiently low for such an equilibrium to exist. The proof is available upon request.

<sup>18</sup>KG assume costless experimentation, but their results hold if there are (sufficiently low) constant experimentation costs for informative experiments. Constant experimentation costs as above can be viewed as being normalized to zero in KG. It also does not matter if uninformative experiments are cheaper than informative experiments: Persuasion has a value for the sender. Persuasion is impossible in KG with an uninformative experiment, however, it is possible with some informative experiments.

<sup>19</sup>It is easy to show that cost functions in Gentzkow and Kamenica (2014) have these properties.

<sup>20</sup>The design of the optimal informative experiment depends on the curvature of the cost function.

by marginally lowering  $\pi_\tau(s_2 | \omega_2)$ . The receiver's threshold of doubt upon the observation of  $\hat{m} = s_1$  is then still passed, if she knows that  $\hat{m} = s_1$  stems from this alternative experiment. Lowering  $\pi_\tau(s_2 | \omega_2)$  reduces the experiment's informativeness according to the Blackwell criterion and, hence, its costs. The sender in addition benefits, as a reduction of  $\pi_\tau(s_2 | \omega_2)$  increases the probability of a positive outcome. The receiver in the hypothetical equilibrium, thus, has to be indifferent between  $a = a_1$  and  $a = a_2$  if she verifies  $\hat{m} = s_1$  and observes that it stems from the informative experiment. In the hypothetical equilibrium the optimal uninformative experiment again has precision  $(1, 0)$ , as it is cheapest and yields a positive outcome with certainty.

These are all the ingredients that are required for the communication breakdown. Consider a receiver who verifies  $\hat{m} = s_1$ . If experiment  $\tau$  is informative, then the indifference condition implies  $\text{prob}\{\omega = \omega_1 | (\sigma_\tau = s_1, \pi_\tau)\} = p_d$ , i.e., that the threshold of doubt is just passed. If experiment  $\tau'$  is uninformative, then we have  $\text{prob}\{\omega = \omega_1 | (\sigma_{\tau'} = s_1, \pi_{\tau'})\} = \frac{1}{2}$ . There is no equilibrium with persuasion by the same logic as above.

## 5 Private experimentation

In contrast to public experimentation the receiver under private experimentation does not directly observe the outcomes of the experiments that the sender ran. Private experimentation allows the sender to disclose a subset of the acquired outcomes. The receiver in general also cannot deduce the number of experiments that the sender ran upon observing the message, as the sender's decision to continue experimenting is history dependent. This renders unraveling à la Milgrom and Roberts (1986) in general impossible. Due to the power of off-the-equilibrium path beliefs there can be many equilibria. In the following the properties of equilibria with persuasion are investigated and the sender preferred equilibrium is derived. In this section it is again assumed that running an informative experiment costs  $c_I$  and running an uninformative experiment costs  $c_U$ .

Let  $\nu(\widehat{m})$  be the equilibrium probability with which message  $\widehat{m}$  is verified and let  $\rho(\widehat{m})$  be the equilibrium probability with which the evidence underlying message  $\widehat{m}$  is such that the receiver who verifies  $\widehat{m}$  chooses  $a = a_1$  with a positive probability in equilibrium.

**Lemma 1** *There is no pure strategy equilibrium with persuasion.*

As an illustration consider a hypothetical pure strategy equilibrium in which the sender runs exclusively one experiment  $\tau$  on the equilibrium path. Suppose  $\text{prob}\{\omega = \omega_1 \mid (\sigma_\tau = s_1, \pi_\tau)\} > p_d$  and suppose the receiver verifies  $\widehat{m} = s_1$  in equilibrium, i.e.,  $\nu(\widehat{m} = s_1) = 1$ . Upon observing evidence  $(\sigma_\tau = s_1, \pi_\tau)$  the receiver in equilibrium deduces that the sender ran only experiment  $\tau$  with a design  $\pi_\tau$  implying  $\text{prob}\{\omega = \omega_1 \mid (\sigma_\tau = s_1, \pi_\tau)\} > p_d$ . In this contingency it is sequentially rational to choose  $a = a_1$ , implying  $\rho(\widehat{m} = s_1) = 1$ . There is no such equilibrium, however. A profitable deviation for the receiver who verifies message  $\widehat{m} = s_1$  is to choose  $a = a_1$  without costly verification of this message. The remaining cases are excluded in the appendix.

Even though there is no pure strategy equilibrium with persuasion, the following example shows that persuasion can occur in a mixed strategy equilibrium.

**Example 2** Let  $c_v \in (0, p_d(1 - p_d))$ . There is an equilibrium with persuasion with the following properties: The sender starts experimenting with an experiment with design  $\pi_\tau = (1, 1)$ . If the outcome of this experiment is positive, then the sender stops experimenting and sends a message with this outcome. If the outcome of the first experiment is adverse, then there is a strictly positive probability smaller than one that the sender runs an experiment with design  $\pi_\tau = (1, 0)$ . If he runs this experiment, then the sender sends a message containing the outcome of this experiment. The receiver mixes regarding verification and may only verify messages with a positive outcome. The receiver who does not verify such a message chooses  $a = a_1$ . The receiver who verifies the message on the equilibrium path only chooses  $a = a_1$  if the outcome stems from an experiment with design

$\pi_\tau = (1, 1)$ . The probability that the sender runs the uninformative experiment increases in the verification costs  $c_v$  and it approaches zero if  $c_v$  goes to zero. The verification probability of an on the equilibrium path message in the sender's favor is greater than  $1/2$ . It decreases in the costs of conducting the uninformative experiment and it does not depend on the costs an running informative experiment.  $\square$

There are some differences between Example 1 and 2. In Example 1 the sender runs only one experiment, whereas in Example 2 the sender may run one or two experiments. Under private experimentation there is, in general, no equilibrium where the sender runs only one experiment on the equilibrium path and mixes between designs  $\pi_\tau = (1, 1)$  and  $\pi_\tau = (1, 0)$  for this experiment. If there were such an equilibrium, then the sender has to be indifferent between running either of these experiments, he has to have an incentive to start experimenting and he must not continue experimenting after observing an adverse outcome of the first experiment. If he has a strict incentive to start with an experiment with  $\pi_\tau = (1, 1)$ , then this implies via the indifference condition and the fact that an experiment with  $\pi_\tau = (1, 0)$  yields a positive outcome regardless of the state that there is a strict incentive to run the experiment with  $\pi_\tau = (1, 0)$  if the outcome of the first experiment is adverse.<sup>21</sup> There is, thus, a profitable deviation to continue experimenting if the outcome of the first experiment is adverse. In the equilibrium in Example 2 the verification probability has to be such that the sender is indifferent between stopping experimentation unsuccessfully and running the uninformative experiment whose outcome can be used to persuade the receiver who does not verify the message.

There can be other equilibria with persuasion, e.g., equilibria where the sender does not run an experiment with  $\pi_\tau = (1, 1)$  on the equilibrium path. In order to show that outcomes from uninformative experiments are sent in any equilibrium with persuasion, it is first established that not all messages that are verified are based on evidence that

---

<sup>21</sup>The sender would then hide the adverse outcome of the first experiment and reveal the positive outcome of the second experiment.

persuades the receiver who verifies them. That is, we cannot have an equilibrium with persuasion with  $\rho(\hat{m}) = 1$  for all messages  $\hat{m}$  that the receiver verifies on the equilibrium path. Suppose the opposite is true, i.e.,  $\rho(\hat{m}) = 1$  for all messages  $\hat{m}$  that the receiver verifies on the equilibrium path. The receiver's best response implies  $v(\hat{m}) = 0$  for all such  $\hat{m}$ , i.e., not to verify  $\hat{m}$ , and choosing  $a = a_1$  upon observing each such message. The sender's best response to the receiver's behavior is to run exclusively uninformative experiments, which contradicts  $\rho(\hat{m}) = 1$  for all messages  $\hat{m}$  that the receiver verifies on the equilibrium path. Lemma 2 follows.

**Lemma 2** *There is no equilibrium with persuasion with  $\rho(\hat{m}) = 1$  for all messages  $\hat{m}$  that the receiver verifies on the equilibrium path.*

Lemma 2 implies that there is a positive probability that the evidence underlying some message  $\hat{m}$  that the receiver verifies is such that this receiver is not persuaded upon observing it. The properties of such evidence, however, are not obvious. Experimentation may serve several purposes. There may be benefits from learning about the state or benefits from reaching posteriors with more promising continuation payoffs. Learning and changing the posterior requires informative experiments. The outcomes of such informative experiments could also be used for persuading the receiver who does not verify the message, even if they are not sufficiently informative to persuade the receiver who verifies it. An equilibrium with persuasion clearly requires that the sender runs informative experiments. Yet, the following proposition finds that there is no equilibrium with persuasion, in which the sender exclusively sends messages in his favor that are based on informative experiments.

**Proposition 2** *Equilibria with persuasion exist if experimentation and verification costs are sufficiently low. In any equilibrium with persuasion, the sender sends messages with a positive probability on the equilibrium path that contain an outcome from an uninformative experiment.*

An outcome of an uninformative experiment does not relate to the state of the world and in this sense it does not have an inherent meaning. It can be interpreted

as cheap talk. From this perspective Proposition 2 asserts that any message in the sender’s favor contains some cheap talk element with a positive probability. This cheap talk element is not costless, as it results from a costly uninformative experiment. The costs of running this uninformative experiment in this model can also be viewed as information transmission costs or costs of making an uninformative statement appear informative. As the messages in equilibrium contain informative and uninformative components, the model combines aspects of the literature on persuasion and the cheap talk literature.<sup>22</sup>

In the following it is argued that a certain equilibrium with persuasion exists. Then it is established that this equilibrium is sender preferred. “Sender preferred” means that there is no other equilibrium with persuasion in which the sender’s payoff is higher.

Consider an equilibrium in the spirit of Example 2, where the sender (i) starts experimenting with an informative experiment with precision  $(1, 1 - c_I)$ , (ii) stops experimenting successfully after observing a positive outcome from this experiment and (iii) runs an uninformative experiment with  $\pi_\tau = (1, 0)$  with a positive probability smaller than one after observing an adverse outcome of the first experiment. Off-the-equilibrium path the sender’s behavior is sequentially rational. In equilibrium the receiver mixes between verifying a message  $\hat{m} = s_1$  and not verifying it. If she does not verify this message, then she chooses  $a = a_1$ . If she verifies it, then she only chooses  $a = a_1$  if the outcome stems from an informative experiment with precision  $(1, 1 - c_I)$  or  $(p, 1)$ , with  $p \in (0, 1]$ .<sup>23</sup> In all other cases the receiver chooses  $a = a_2$ . Her belief upon observing an off-the-equilibrium path precision that is not some precision  $(p, 1)$  is such that it is optimal for her to choose  $a = a_2$ .<sup>24</sup>

In order to confirm that this is an equilibrium consider first the sender. An equi-

---

<sup>22</sup>In fact, the model can be easily reformulated such that the sender can only run informative experiments and in addition send positive cheap talk outcomes that do not stem from experimentation. The strategic effects are the same.

<sup>23</sup>If the precision is  $(p, 1)$ , then a positive outcome can only realize in state  $\omega = \omega_1$ . I.e., upon observing a positive outcome from such an experiment the receiver must believe that  $\omega = \omega_1$ , in which case choosing  $a = a_1$  is optimal.

<sup>24</sup>E.g., she may believe that the sender privately ran an additional experiment which perfectly predicts the state of the world, i.e., with precision  $(1, 1)$ , that yielded an adverse outcome that is not revealed.

librium condition is that the sender does not run a further experiment with precision  $(1, 1 - c_I)$  after observing an adverse outcome of the first experiment and then stops unsuccessfully. After observing an adverse outcome of the first experiment the sender knows that the state of the world is  $\omega = \omega_2$ . Running a further experiment with precision  $(1, 1 - c_I)$  yields continuation payoff  $(1 - (1 - c_I)) - c_I$ , as he persuades the receiver with a positive outcome regardless of whether she verifies this outcome. Therefore, a necessary condition for equilibrium existence is  $(1 - (1 - c_I)) - c_I \leq 0$ , which is satisfied with equality. That is, the precision of the informative experiment is just high enough to deter him from running a further such experiment upon observing an adverse outcome from the first experiment. The other equilibrium conditions for the sender are satisfied analogous to the proof for the existence of the equilibrium in Example 2.

Consider next the receiver. Suppose she observes a message containing a positive outcome. The probability with which the sender runs the uninformative experiment after observing an adverse outcome of the first experiment determines the receiver's value of message  $\hat{m} = s_1$ . In equilibrium the receiver has to be indifferent between choosing  $a = a_1$  without verification after observing  $\hat{m} = s_1$  and verifying with consequent sequentially rational behavior. Let  $\xi$  be the probability that the sender runs the uninformative experiment after observing an adverse outcome of the first experiment. Her utility from choosing  $a = a_1$  without verification upon observing  $\hat{m} = s_1$  is  $\frac{1}{1+(1-(1-c_I))+(1-c_I)\xi} + (1 - \frac{1}{1+(1-(1-c_I))+(1-c_I)\xi})(1 - p_d)$ .<sup>25</sup> Her utility from verifying and consequent sequentially rational behavior is  $\frac{1}{1+(1-(1-c_I))+(1-c_I)\xi} + (1 - \frac{1}{1+(1-(1-c_I))+(1-c_I)\xi})((1 - (1 - c_I))(1 - p_d) + (1 - c_I)) - c_v$ .<sup>26</sup> In equilibrium  $\xi^*$  is such

<sup>25</sup>If  $\omega = \omega_1$ , her utility from choosing  $a = a_1$  is 1. If  $\omega = \omega_2$ , her payoff from choosing  $a = a_1$  is  $1 - p_d$ . We have  $\text{prob}\{\omega = \omega_1 \mid \hat{m} = s_1\} = \frac{1}{1+(1-(1-c_I))+(1-c_I)\xi}$ , given that the receiver does not verify  $\hat{m} = s_1$ .

<sup>26</sup>We have  $\text{prob}\{\omega = \omega_1 \mid \hat{m} = s_1\} = \frac{1}{1+(1-(1-c_I))+(1-c_I)\xi}$  before the receiver verifies  $\hat{m} = s_1$ . If the receiver observes  $\hat{m} = s_1$  and verifies it, then she observes whether it was generated with an experiment with  $\pi_\tau = (1, 1 - c_I)$  or an experiment with  $\pi_\tau = (1, 0)$ . In state  $\omega = \omega_1$  the informative experiment always yields a positive outcome, the correct decision  $a = a_1$  is made with certainty and the payoff is 1. Consider state  $\omega = \omega_2$ . In this case the receiver makes the wrong decision  $a = a_1$  yielding payoff  $1 - p_d$  if the informative experiment generated a positive outcome  $s_1$ , which occurs with probability  $(1 - (1 - c_I))$ . If the informative experiment generates the adverse outcome  $s_2$ , which occurs with probability  $1 - c_I$ , then the receiver makes the correct decision  $a = a_2$ . The reason is that in this case she observes that the positive outcome in the message stems from the uninformative experiment.

that these two utilities are equal, which corresponds to

$$\xi^* = \frac{c_v - p_d c_I + c_v c_I + p_d c_I^2}{p_d - p_d c_I - c_v - p_d c_I + c_v c_I + p_d c_I^2}.$$

We have  $\xi^* \in (0, 1)$  if  $c_v \in (0, \frac{1}{2} p_d (1 - c_I))$ .<sup>27</sup>

Finally, consider the receiver's behavior upon observing either message. Upon observing message  $\widehat{m} = s_2$  the receiver at the optimum chooses  $a = a_2$  without verification, as it implies  $\omega = \omega_2$  with certainty. Consider next message  $\widehat{m} = s_1$ . The receiver who does not verify the message only chooses  $a = a_1$  if her threshold of doubt is passed. This is the case if  $\text{prob}\{\omega = \omega_1 \mid \widehat{m} = s_1\} \geq p_d$ , which is equivalent to  $1 - \frac{c_v}{p_d(1-c_I)} \geq p_d$  and which holds for all  $c_v \in (0, p_d(1-p_d)(1-c_I))$ .<sup>28</sup> The receiver who verifies a message containing a positive outcome from the on the equilibrium path experiment then clearly has a strict incentive to choose  $a = a_1$ . Off-the-equilibrium path behavior of the receiver who verifies  $\widehat{m} = s_1$  is sequentially rational as described above.

In the following it is shown that the equilibrium, in which  $\pi_\tau(s_2 \mid \omega_2) = 1 - c_I$ , is sender preferred. For this purpose the following two lemmas are useful.

**Lemma 3** *In any equilibrium with persuasion the sender stops experimenting at any history  $h$  where the last experiment in  $h$  is the uninformative experiment with precision  $\pi_t = (1, 0)$  and outcome  $\sigma_t = s_1$ .*

The intuition for Lemma 3 is that if the sender instead ran a further informative experiment after running the uninformative experiment, then he could profitably deviate by modifying his strategy such that it replicates the persuasion probability at lower expected experimentation costs. The modification is such that he “reverses the order” of these experiments in the sense that he first runs the informative experiment and he runs the uninformative experiment only in case the outcome of the informative

---

Via equilibrium behavior she deduces that the first experiment must have yielded an adverse outcome, implying  $\omega = \omega_2$  with certainty. In this case  $a = a_2$  is optimal.

<sup>27</sup>  $\frac{d\xi^*}{dc_v} > 0$  and  $\lim_{c_v \rightarrow 0} \xi^* = \frac{c_I}{1-c_I} \in (0, 1)$  (since  $c_I < \frac{1}{2}$  by assumption) and  $\xi^* = 1$  if  $c_v = \frac{1}{2} p_d (1 - c_I)$ .

<sup>28</sup> We have  $\text{prob}\{\omega = \omega_1 \mid \widehat{m} = s_1\} = \frac{1}{1+(1-(1-c_I))+(1-c_I)\xi^*}$ , which simplifies to  $\text{prob}\{\omega = \omega_1 \mid \widehat{m} = s_1\} = 1 - \frac{c_v}{p_d(1-c_I)}$  after plugging in  $\xi^*$ .

experiment is adverse. If the outcome of the informative experiment is positive instead, then he can persuade the receiver regardless of whether she verifies  $\hat{m} = s_1$  and he does not run the costly uninformative experiment in this case.

The next lemma compares the sender's payoff in any equilibrium with persuasion from running the uninformative experiment and stopping unsuccessfully at any history.

**Lemma 4** *In any equilibrium with persuasion the sender is indifferent between running the uninformative experiment with precision  $\pi_\tau = (1, 0)$  and stopping experimentation unsuccessfully at any history.*

Intuitively, if the sender instead strictly prefers running the uninformative experiment over stopping unsuccessfully, then he would find a positive outcome regardless of the state of the world. Such an outcome then does not have the power to persuade the receiver who does not verify a message containing a positive outcome. It also does not have the power to persuade a receiver who verifies the message. But then running the costly uninformative experiment is pointless for the sender and he cannot strictly prefer running it over stopping unsuccessfully. If the sender instead strictly prefers stopping unsuccessfully over running the uninformative experiment, then he would never run the latter, which contradicts Proposition 2.

FL derive the sender preferred equilibrium in a model where the receiver exogenously always observes the precision of the revealed evidence. They show that in this equilibrium the sender starts experimenting with an experiment of precision  $(1, 1 - c_I)$  and stops experimenting after either outcome. If the outcome is positive, then it is revealed and the receiver is persuaded. If the outcome is adverse, then the sender stops experimenting unsuccessfully and the receiver is not persuaded. Using Lemmas 3 and 4, it can now be shown that for each equilibrium with persuasion in the current model there is a corresponding equilibrium in FL that is payoff equivalent for the sender. It follows that if an equilibrium in the current model yields the same payoff for the sender as the sender preferred equilibrium in FL, then it is also sender preferred here. This is the case for the above equilibrium where he starts with an experiment with precision

$(1, 1 - c_I)$  and where he may only run the uninformative experiment after an adverse outcome of the first experiment.

According to Lemma 3 the sender always stops experimenting after running the uninformative experiment. According to Lemma 4 the sender is just indifferent between running the uninformative experiment and stopping experimentation unsuccessfully in any equilibrium with persuasion. Consider any equilibrium with persuasion in the current model. Modify the sender's equilibrium strategy as follows such that he does not run the uninformative experiment at any history. At any history where the sender (i) mixes exclusively between running the uninformative experiment and stopping unsuccessfully he now stops experimenting unsuccessfully, (ii) runs the uninformative experiment with certainty he now stops experimenting unsuccessfully and (iii) mixes between running the uninformative experiment and an informative experiment he now runs the informative experiment. If the receiver's strategy is not changed, then the modified sender's strategy is sequentially rational: At any history where he now stops experimenting unsuccessfully instead of running the uninformative experiment (cases (i) and (ii)), the continuation payoff is the same with the equilibrium strategy and the modified strategy due to the indifference condition of Lemma 4. At any history where the sender now runs an informative experiment instead of running the uninformative experiment (case (iii)) he receives the same continuation utility, as otherwise mixing with the original strategy at such a history is not sequentially rational. Hence, at any other on the equilibrium path history the same continuation decision is also optimal. The modified strategy is payoff equivalent to the equilibrium strategy for the sender if the receiver's strategy is not changed. If the receiver now exogenously observes the precision of the experiment of a revealed positive outcome, then it is optimal for her to behave like the receiver who verifies this outcome. Since the receiver who verifies the positive outcome is persuaded to choose  $a = a_1$  if it stems from the appropriate informative experiment, a receiver who exogenously observes the outcome also is persuaded to choose  $a = a_1$ . It follows that for each equilibrium with persuasion in the current model, there is a payoff equivalent equilibrium with persuasion in FL for the sender where the sender plays the

modified strategy and the receiver behaves like the receiver who verifies the message. In particular, for the equilibrium here, where the sender starts with an experiment with precision  $(1, 1 - c_I)$  and where he may only run the uninformative experiment after an adverse outcome of the first experiment, the payoff equivalent equilibrium for the sender in FL is where he exclusively runs the first experiment with precision  $(1, 1 - c_I)$  and then stops after either outcome. The reason is that here the sender is indifferent between running the uninformative experiment and stopping unsuccessfully and the informative experiment has the same design. Proposition 3 directly follows.

**Proposition 3** *If experimentation and verification costs are sufficiently small, then we have the following in the sender preferred equilibrium. The sender starts experimenting with an experiment with precision  $(1, 1 - c_I)$ . If this experiment yields a positive outcome, then it is revealed to the receiver and the receiver chooses  $a = a_1$  regardless of whether she verifies the message. If the outcome is adverse, then the sender runs an uninformative experiment with precision  $(1, 0)$  with a strictly positive probability smaller than one and stops experimenting unsuccessfully otherwise. In the former case the receiver who does not verify the message chooses  $a = a_1$  and the receiver who verifies the message chooses  $a = a_2$ .*

The sender's behavior in the sender preferred equilibrium here is similar to the sender's behavior in the corresponding equilibrium in FL. He starts with the same informative experiment. Instead of stopping after either outcome, though, he here may run an uninformative experiment in case the first outcome is adverse. Both equilibria are payoff equivalent for the sender because here it is an equilibrium condition that the sender is indifferent between stopping unsuccessfully and running the uninformative experiment. If this were not the case then he would either always present a positive outcome regardless of the state of the world or he would never run an uninformative experiment. Both cases are inconsistent with an equilibrium with persuasion in this game. Due to the indifference condition the sender does not benefit from running the uninformative experiment. He is as well off as in the case where the receiver always

verifies the sender's message.

If the receiver exogenously always observes the experiment design of the revealed outcome, then the sender in the sender preferred equilibrium only runs an informative experiment. With endogenous verification on the other hand the sender's message is diluted by cheap talk: There is a chance that a positive outcome stems from the uninformative experiment. The receiver who does not verify the message cannot distinguish whether the outcome stems from an informative or uninformative experiment. The receiver who verifies the message on the other hand can deduce that the state is  $\omega = \omega_2$  if he observes a positive outcome from an uninformative experiment: The uninformative experiment is only run if the outcome from the first informative experiment is adverse and the design of the informative experiment is such that an adverse outcome perfectly predicts  $\omega = \omega_2$ . Overall the receiver who verifies the message has the same gross utility (i.e., neglecting verification costs) as in the case where she exogenously observes the revealed outcome's design. If the receiver does not verify then her utility if she observes a positive outcome is equal to this gross utility minus verification costs due to the receiver's indifference condition (she only mixes between verification and no verification if she is indifferent). Saving costs by not verifying is just offset by making more mistakes due to not observing the cheap talk component of the message.

## 6 Discussion

Communication is important for achieving economic efficiency. For example, in the mechanism design literature it is often assumed that agents can exclusively send cheap talk messages, which severely limits economic efficiency (e.g., Myerson and Satterthwaite, 1983). In practice communication often contains hard components, like logical arguments or regressions on a public database, and cheap talk. Yet, the persuasion and the cheap talk literature have evolved rather separately. This paper and a few other recent models encompass features of both strands of the literature.

The focus here is on a receiver who may not pay attention to a sender's attempt

to persuade her. If the collection of evidence that can be used for persuasion occurs in private, as with thought experiments or private regressions on a public database, then communication is possible if running experiments and paying attention is not too costly. In any equilibrium with persuasion the sender's message contains cheap talk with a positive probability. The cheap talk element dilutes the value of the message that the sender sends.

The paper highlights the role of the flexibility of the design of the experiments that the sender may run. Full flexibility leads to a communication breakdown under public experimentation. The paper finds that communication and persuasion is possible under public experimentation if the flexibility is reduced. Under private experimentation an exogenous reduction of the sender's flexibility is not necessary. It can be obtained endogenously and, therefore, the flexibility per se is not an obstacle for persuasion.

## APPENDIX

**Proof of Proposition 1:** Consider a potential equilibrium with persuasion in which the receiver never verifies any message. Without loss of generality suppose that the receiver chooses  $a = a_1$  after observing  $\hat{m} = s_1$ . It is the (unique) best response for the sender to run an experiment  $\tau$  with  $\pi_\tau = (1, 0)$ : The sender then obtains a positive outcome regardless of  $\omega$ , i.e., it maximizes the persuasion probability, and expected experimentation costs are at the minimum, given that the sender starts experimenting. As the message containing a positive outcome is uninformative regarding  $\omega$ , it is not optimal for the receiver to choose  $a = a_1$  upon observing such a message, which contradicts that she is persuaded. Q.E.D.

**Verification of Example 1:** The receiver observes the outcome of the experiment that the sender ran. Let  $\rho$  be the probability with which the sender runs the experiment with design  $\pi_\tau = (1, 1)$ . Let  $\nu$  be the probability with which the receiver verifies the message containing outcome  $s_1$ .

Step (1) analyzes the receiver's behavior. Step (2) analyzes the sender's behavior.

(1) Receiver behavior:

If the receiver observes a message with outcome  $s_1$  and verifies it, then she observes whether the sender generated it with an experiment with  $\pi_\tau = (1, 1)$  or an experiment with  $\pi_\tau = (1, 0)$ . In the former case her optimal choice is  $a = a_1$  yielding utility 1 and in the latter case her posterior equals her prior belief and the optimal choice is  $a = a_2$  yielding expected utility  $\frac{1}{2}p_d + \frac{1}{2}$ . Given that the receiver observes  $s_1$ , the probability that the underlying experiment is informative is  $\frac{\rho}{\rho + 2(1-\rho)} = \frac{\rho}{2-\rho}$ . Her expected utility from verifying the message is, thus,  $\frac{\rho}{2-\rho} + (1 - \frac{\rho}{2-\rho})(\frac{1}{2}p_d + \frac{1}{2}) - c_v$ .

If the receiver observes a message with outcome  $s_1$  and does not verify it, then she chooses  $a = a_1$ . The probability that the underlying experiment is informative is  $\frac{\rho}{2-\rho}$ . In this case she obtains utility 1 from  $a = a_1$ . If the experiment is uninformative, then she obtains utility  $(\frac{1}{2} + \frac{1}{2}(1 - p_d))$  from  $a = a_1$ . The receiver's expected utility from not verifying the message and choosing  $a = a_1$  is, therefore,  $\frac{\rho}{2-\rho} + (1 - \frac{\rho}{2-\rho})(\frac{1}{2} + \frac{1}{2}(1 - p_d))$ .

The receiver is indifferent between verifying and not verifying upon observing a message with outcome  $s_1$  if

$$\frac{\rho}{2-\rho} + \left(1 - \frac{\rho}{2-\rho}\right)\left(\frac{1}{2} + \frac{1}{2}(1-p_d)\right) = \frac{\rho}{2-\rho} + \left(1 - \frac{\rho}{2-\rho}\right)\left(\frac{1}{2}p_d + \frac{1}{2}\right) - c_v,$$

which is equivalent to  $\rho = \frac{2c_v - 2p_d + 1}{c_v - 2p_d + 1}$ . We have  $\rho \in (0, 1)$  if  $c_v \in (0, p_d - \frac{1}{2})$ .

It remains to show that  $a = a_1$  is indeed optimal for the receiver if she does not verify the message with outcome  $s_1$ . This choice is optimal if  $\text{prob}\{\omega = \omega_1 \mid \hat{m} = s_1\} = \frac{1}{2-\rho} \geq p_d$ . As the equilibrium  $\rho$  decreases in  $c_v$ , with  $\lim_{c_v \rightarrow 0} \rho = 1$ , there are  $c_v > 0$  sufficiently low such that  $\text{prob}\{\omega = \omega_1 \mid \hat{m} = s_1\} = \frac{1}{2-\rho} \geq p_d$  is satisfied. The condition is satisfied if  $c_v \in (0, 3p_d - 2p_d^2 - 1]$ .

As  $3p_d - 2p_d^2 - 1 < p_d - \frac{1}{2}$  for all  $p_d \in (\frac{1}{2}, 1)$ , equilibrium existence requires  $c_v \in (0, 3p_d - 2p_d^2 - 1]$ .

(2) Sender behavior:

It has to be shown that (i) the sender is indifferent between running either experiment and that (ii) the sender's benefit from running an experiment with either  $\pi_\tau = (1, 1)$  or  $\pi_\tau = (1, 0)$  is at least 0.

(i) If the sender runs the experiment with  $\pi_\tau = (1, 1)$ , then  $s_1$  realizes with probability  $\frac{1}{2}$  and  $s_1$  induces  $a = a_1$  regardless of whether the receiver verifies the message, yielding expected utility  $\frac{1}{2} - c_I$ .

If the sender runs the experiment with  $\pi_\tau = (1, 0)$ , then outcome  $s_1$  realizes with probability 1, but it only induces  $a = a_1$  if the receiver does not verify the message, i.e., with probability  $1 - \nu$ , yielding expected utility  $1 - \nu - c_U$ .

The sender is indifferent between running either an experiment with  $\pi_\tau = (1, 0)$  or an experiment with  $\pi_\tau = (1, 1)$  if

$$1 - \nu - c_U = \frac{1}{2} - c_I.$$

Solving this equation for  $\nu$  yields  $\nu = \frac{1}{2} - c_U + c_I$ , which is in the interval  $(\frac{1}{2}, 1)$  since  $c_U \in (0, c_I)$  and  $c_I \in (0, \frac{1}{2})$ .

(ii) The sender's expected utility if he runs the experiment with  $\pi_\tau = (1, 1)$  is  $\frac{1}{2} - c_I$ , as derived in (i). He is better off to run this experiment than not to run any experiment if

$$\frac{1}{2} - c_I \geq 0,$$

which is satisfied, as  $c_I \in (0, \frac{1}{2})$  by assumption.

Due to the indifference condition in (i) the sender is also better off to run an experiment with  $\pi_\tau = (1, 0)$  than not to run any experiment. Q.E.D.

**Proof of Lemma 1:** Denote by  $\widehat{m}^*$  a message that is sent with a positive probability on the equilibrium path and that induces the receiver who does not verify it to choose  $a = a_1$  with a positive probability. Message  $\widehat{m}^*$  can be interpreted as a recommendation to choose the sender's preferred action. Let  $\widehat{M}^*$  be the set of all messages  $\widehat{m}^*$ .

Consider a potential pure strategy equilibrium with persuasion where  $\widehat{M}^*$  is not empty. First, it is shown that there is no such equilibrium with  $v(\widehat{m}^*) = 0$  for all  $\widehat{m}^* \in \widehat{M}^*$ . It follows that in any such potential pure strategy equilibrium we have  $v(\widehat{m}^*) = 1$  and  $\rho(\widehat{m}^*) = 0$  for some  $\widehat{m}^* \in \widehat{M}^*$  and / or we have  $v(\widehat{m}^*) = 1$  and  $\rho(\widehat{m}^*) = 1$  for some  $\widehat{m}^* \in \widehat{M}^*$ . These cases are excluded next.

There is no equilibrium with  $v(\widehat{m}^*) = 0$  for all  $\widehat{m}^* \in \widehat{M}^*$ : Suppose the opposite. The sender's best response is to run an experiment  $\tau$  with  $\pi_\tau = (1, 0)$  if  $\widehat{m}^* = s_1$  (respectively, an experiment  $\tau$  with  $\pi_\tau = (0, 1)$  if  $\widehat{m}^* = s_2$ ), since  $c_U < c_I$  and  $\text{prob}\{\sigma_\tau = s_1 \mid \pi_\tau = (1, 0)\}$  (respectively,  $\text{prob}\{\sigma_\tau = s_2 \mid \pi_\tau = (0, 1)\}$ ) is greater than if outcome  $s_1$  (respectively,  $s_2$ ) was generated with an experiment that is informative regarding  $\omega$ .<sup>29</sup> The receiver deduces that the message is uninformative and it is optimal for her to choose  $a = a_2$ , i.e., she is not persuaded.

There is no equilibrium with  $v(\widehat{m}^*) = 1$  and  $\rho(\widehat{m}^*) = 0$  for some  $\widehat{m}^* \in \widehat{M}^*$ : The receiver's best response to  $\rho(\widehat{m}^*) = 0$ , i.e., the evidence underlying  $\widehat{m}^*$  being always uninformative, is to choose  $a = a_2$  upon observing  $\widehat{m}^*$  without verifying it.

There is no equilibrium with  $v(\widehat{m}^*) = 1$  and  $\rho(\widehat{m}^*) = 1$  for some  $\widehat{m}^* \in \widehat{M}^*$ : The

---

<sup>29</sup>Running an uninformative experiment with  $\pi_\tau = (\frac{1}{2}, \frac{1}{2})$  is weakly worse for the sender than to run an uninformative experiment with  $\pi_\tau = (1, 0)$  (respectively,  $\pi_\tau = (0, 1)$ ).

receiver's best response to  $\rho(\widehat{m}^*) = 1$ , i.e., the evidence underlying  $\widehat{m}^*$  being always persuasive for a receiver who verifies the message, is to choose  $a = a_1$  upon observing  $\widehat{m}^*$  without verifying it.

Finally, consider a potential pure strategy equilibrium with persuasion where  $\widehat{M}^*$  is empty, i.e., where the receiver who does not verify messages chooses  $a = a_2$ . In such an equilibrium only the receiver who verifies the message may choose  $a = a_1$ . Consider a message that this receiver verifies. We cannot have that the underlying evidence of this message is always persuasive, as in this case the receiver who verifies the message has a profitable deviation not to verify it and to choose  $a = a_1$  saving the verification costs. Analogously, we cannot have that the underlying evidence of such a message is always not persuasive. Hence, there is no such pure strategy equilibrium. We also do not have a pure strategy equilibrium with persuasion where the receiver never verifies any message and always chooses  $a = a_2$ . Q.E.D.

**Verification of Example 2:** Step (1) introduces beliefs that support the equilibrium behavior. Step (2) analyzes the receiver's on the equilibrium path behavior. Step (3) analyzes the sender's on the equilibrium path behavior. Denote by  $\xi$  the probability that the sender runs an experiment with design  $\pi_\tau = (1, 0)$  if he observes an adverse outcome  $s_2$  of the first experiment (which has design  $\pi_\tau = (1, 1)$ ). Let  $\nu$  be the probability with which the receiver verifies a message that contains outcome  $s_1$ .

(1) On the equilibrium path beliefs are formed in accordance with Bayes' law. Off-the-equilibrium path beliefs can be specified that support the equilibrium behavior. An off-the-equilibrium path event is where the receiver who verifies a message observes an off-the-equilibrium path design. Consider a message containing a positive outcome.

If the design of the corresponding experiment is not  $(p, 1)$ , with  $p \in (0, 1)$ , then the receiver who verifies the message forms a probability assessment over experimentation histories such that the probability that  $\omega = \omega_1$  conditional on this assessment is below the threshold of doubt. E.g., she may believe that the sender privately ran a single additional experiment with design  $(1, 1)$  that yielded an adverse outcome. Given these beliefs the receiver's best response is to choose  $a = a_2$ . The sender, thus, does not

benefit from sending a message containing an outcome stemming from an experiment with such an off-the-equilibrium path design.

If the off-the-equilibrium path design is  $(p, 1)$ , with  $p \in (0, 1)$ , then beliefs upon observing a positive outcome are such that the receiver who verifies the message thinks that  $\omega = \omega_1$ , which induces action  $a = a_1$ .<sup>30</sup> The probability that at least one positive outcome realizes from running any number of such experiments is lower than from running a single experiment with design  $(1, 1)$ . The sender also cannot learn more about the state from running any sequence of experiments compared to a single experiment with design  $(1, 1)$ . Therefore, the sender does not benefit from running such off-the-equilibrium path experiments. Due to these off-the-equilibrium path beliefs, only experiments with design  $(1, 1)$  and  $(1, 0)$  may be run on the equilibrium path.

(2) If the receiver observes  $\hat{m} = s_1$  and verifies it, then she observes whether it was generated with an experiment with  $\pi_\tau = (1, 1)$  or an experiment with  $\pi_\tau = (1, 0)$ . In the former case she knows that  $\omega = \omega_1$  and her optimal choice is  $a = a_1$  yielding utility 1. In the latter case she deduces that the first experiment with design  $(1, 1)$  yielded an adverse outcome  $s_2$ , implying  $\omega = \omega_2$ . Her optimal choice in this case is  $a = a_2$  yielding expected utility 1. Conditional on observing  $\hat{m} = s_1$ , the probability that this message was generated with an experiment with design  $\pi_\tau = (1, 1)$  is  $\frac{1}{1+\xi}$ . Her expected utility from verifying the message is, thus,  $\frac{1}{1+\xi} + (1 - \frac{1}{1+\xi}) - c_v = 1 - c_v$ .

If the receiver observes  $\hat{m} = s_1$  and does not verify it, then she chooses  $a = a_1$  according to the postulated equilibrium. If  $\omega = \omega_1$ , her utility from choosing  $a = a_1$  is 1. If  $\omega = \omega_2$ , her payoff from choosing  $a = a_1$  is  $1 - p_d$ . We have  $\text{prob}\{\omega = \omega_1 \mid \hat{m} = s_1\} = \frac{1}{1+\xi}$ , given that the receiver does not verify  $\hat{m} = s_1$ . The receiver's expected utility from not verifying  $\hat{m} = s_1$  and choosing  $a = a_1$  is, therefore,  $\frac{1}{1+\xi} + (1 - \frac{1}{1+\xi})(1 - p_d)$ .

The receiver is indifferent between verifying and not verifying message  $\hat{m} = s_1$  if

$$\frac{1}{1+\xi} + (1 - \frac{1}{1+\xi})(1 - p_d) = 1 - c_v,$$

---

<sup>30</sup>If an experiment with design  $(p, 1)$  yields a positive outcome, then it implies  $\omega = \omega_1$ .

which is equivalent to  $\xi = \frac{c_v}{p_d - c_v}$ . We have  $\xi \in (0, 1)$  if  $c_v \in (0, \frac{1}{2}p_d)$ .

It remains to show that  $a = a_1$  is indeed optimal for the receiver if she does not verify message  $\widehat{m} = s_1$ . This choice is optimal if  $\text{prob}\{\omega = \omega_1 \mid \widehat{m} = s_1\} = \frac{1}{1+\xi} \geq p_d$ . As the equilibrium  $\xi$  increases in  $c_v$ , with  $\lim_{c_v \rightarrow 0} \xi = 0$ , there are  $c_v > 0$  sufficiently low such that  $\text{prob}\{\omega = \omega_1 \mid \widehat{m} = s_1\} = \frac{1}{1+\xi} \geq p_d$  is satisfied. The inequality holds if  $c_v \leq p_d - p_d^2$ .

As  $p_d - p_d^2 < \frac{1}{2}p_d$  for all  $p_d \in (\frac{1}{2}, 1)$ , equilibrium existence requires  $c_v \in (0, p_d(1 - p_d))$ .

(3) Consider the sender's message. If the sender stops experimenting and the stock of collected experimental outcomes contains an outcome that induces  $a_1$  if the appropriate message is sent, then he sends such a message as  $a_1$  is the preferred action. If the sender stops experimenting unsuccessfully, then he induces action  $a_2$  with any feasible message. Therefore, sending any feasible message is optimal after unsuccessful experimentation.

The sender in equilibrium may exclusively run experiments with  $\pi_\tau = (1, 1)$  and  $\pi_\tau = (1, 0)$ . On the equilibrium path the sender's first experiment has design  $\pi_\tau = (1, 1)$ . If this experiment yields outcome  $s_1$ , then he stops experimenting and reveals  $\widehat{m} = s_1$  containing the outcome of the first experiment. If the first experiment yields outcome  $s_2$ , then he runs a second experiment with design  $\pi_\tau = (1, 0)$  with a probability  $\xi \in (0, 1)$ . Then he stops experimenting. It has to be shown that (i) the sender starts experimenting with design  $\pi_\tau = (1, 1)$  and (ii) the sender after observing outcome  $s_2$  of an experiment with design  $\pi_\tau = (1, 1)$  is indifferent to run an experiment with design  $\pi_\tau = (1, 0)$  (versus no further experiment) and does not want to run an experiment with  $\pi_\tau = (1, 1)$ .

(i) Anticipating his and the receiver's following behavior, the sender's gross payoff from running the first experiment with  $\pi_\tau = (1, 1)$  is  $\frac{1}{2} + \frac{1}{2}\xi(1 - \nu)$ . Experimenting according to the equilibrium is better than not experimenting if

$$\frac{1}{2} + \frac{1}{2}\xi(1 - \nu) - c_I - \frac{1}{2}\xi c_u \geq 0. \quad (1)$$

The best deviation, given that the sender starts experimenting, is to run the first

experiment with  $\pi_\tau = (1, 0)$ , which yields utility  $(1 - \nu) - c_U$ .<sup>31</sup> This deviation is not profitable, if

$$\frac{1}{2} + \frac{1}{2}\xi(1 - \nu) - c_I - \frac{1}{2}\xi c_U \geq 1 - \nu - c_U. \quad (2)$$

(ii) Outcome  $s_2$  of the first experiment implies  $\omega = \omega_2$ . After observing outcome  $s_2$  and knowing that  $\omega = \omega_2$ , the sender does not run a further costly experiment with  $\pi_\tau = (1, 1)$  as the probability to obtain an outcome  $s_1$  from such an experiment is zero. After observing outcome  $s_2$  of the first experiment, the sender is indifferent between not running a further experiment and to run an experiment with  $\pi_\tau = (1, 0)$  if

$$(1 - \nu) - c_U = 0$$

which is equivalent to

$$\nu = 1 - c_U.$$

Inequalities (1) and (2) are both equivalent to  $\frac{1}{2} - c_I \geq 0$  if  $\nu = 1 - c_U$ . For  $c_I \in (0, \frac{1}{2})$  the inequality is satisfied. Q.E.D.

**Proof of Proposition 2:** Existence follows from the discussion in the text.

Consider a potential equilibrium with persuasion in which the sender on the equilibrium path exclusively runs informative experiments. Denote some evidence that induces the receiver who verifies the message to choose  $a = a_2$  with certainty in equilibrium by  $m_2$ . Due to Lemma 2 in any equilibrium with persuasion the sender sends some message  $\hat{m}$  that the receiver verifies on the equilibrium path, where  $\hat{m}$  is based on some evidence  $m_2$  with a positive probability. Consider a hypothetical on the equilibrium path experimentation history  $h_t$  where the last experimental outcome completes some  $m_2$ . The sender may stop experimenting at  $h_t$  and send a message with some underlying evidence  $m_2$  or the sender may continue experimenting at  $h_t$ .<sup>32</sup> Consider the last

<sup>31</sup>With such an experiment he can only persuade the receiver who does not verify the message. This occurs with probability  $(1 - \nu)$ . The costs of running the experiment are  $c_U$ , regardless of whether the receiver verifies the evidence.

<sup>32</sup>E.g., he may try to find some underlying evidence inducing  $a = a_1$  with some probability. Revealing  $m_2$  could then be the outside option that potentially persuades the receiver who does not verify the

experiment  $t$  in  $h_t$ . This experiment has design  $\pi_t = (p, q)$ . Suppose that the sender only lacks one outcome  $s_1$  in order to complete  $m_2$  before running experiment  $t$ . The argument if he lacks one outcome  $s_2$  is analogous. Let  $\mu_1$  be the posterior belief that the state is  $\omega = \omega_1$  at history  $h_{t-1}$ . Denote by  $u_1$  the sender's continuation payoff if experiment  $t$  yields outcome  $s_1$ . Similarly, denote by  $u_2$  the continuation payoff if this experiment yields outcome  $s_2$ .

In the following a deviation is constructed, where the sender mixes between two experiments at  $t$ , that makes the sender better off. With this deviation, the (expected) continuation payoffs after either outcome of experiment  $t$  are the same as without the deviation. The probability that either outcome realizes is also the same with the deviation as without. However, expected experimentation costs at  $t$  are lower with mixing than without the deviation, rendering the deviation profitable.

Suppose the sender mixes at  $t$ . With probability  $\alpha > 0$  he runs an experiment with  $\pi_t = (1, 0)$  and with probability  $(1 - \alpha)$  he runs an experiment with  $\pi_t = (p - \varepsilon_1, q + \varepsilon_2)$ , with  $\varepsilon_j > 0$ ,  $j = 1, 2$ . An experiment with  $\pi_t = (1, 0)$  always yields outcome  $s_1$  and is uninformative regarding  $\omega$ . Parameters  $\varepsilon_1$  and  $\varepsilon_2$  are now chosen such that (i) the posterior belief  $prob\{\omega = \omega_1 \mid \mu_1, (\sigma_t = s_2, \pi_t)\}$  if the experiment with design  $\pi_t = (p - \varepsilon_1, q + \varepsilon_2)$  is run is the same as in case the experiment with design  $\pi_t = (p, q)$  is conducted and (ii) outcome  $s_1$  is less likely with the experiment with design  $\pi_t = (p - \varepsilon_1, q + \varepsilon_2)$  than with the experiment with design  $\pi_t = (p, q)$ . By definition we have  $prob\{\omega = \omega_1 \mid \mu_1, (\sigma_t = s_2, \pi_t)\} = \frac{\mu_1(1-p)}{\mu_1(1-p) + (1-\mu_1)q}$  if experiment  $t$  has design  $\pi_t = (p, q)$ . Consider a small decrease of  $p$ , i.e.,  $\varepsilon_1$  is small and positive. The posterior  $prob\{\omega = \omega_1 \mid \mu_1, (\sigma_t = s_2, \pi_t)\}$  with this modification can only stay constant if  $q$  increases sufficiently, i.e.,  $\varepsilon_2$  is positive. For a sufficiently small and positive  $\varepsilon_1$ , there is, thus, a  $\varepsilon_2(\varepsilon_1) > 0$  such that  $prob\{\omega = \omega_1 \mid \mu_1, (\sigma_t = s_2, \pi_t)\}$  is constant. Parameters  $\varepsilon_j > 0$ ,  $j = 1, 2$ , ensures that (ii) is satisfied. As the posterior at outcome  $s_2$  of experiment  $t$  is the same with the deviation and without, the continuation payoffs in both cases are equal to  $u_2$ . Next set  $\alpha = \alpha(\varepsilon_1)$  such that the probability that  $s_1$  realizes 

---

message in case the search for evidence inducing  $a = a_1$  is not successful.

with the deviation is the same as with the hypothetical equilibrium. With this  $\alpha$ ,  $\varepsilon_1$  and  $\varepsilon_2$ , the probability that  $s_j$  realizes is the same with and without the deviation for each  $j = 1, 2$ . This implies that the expected posterior (from  $h_{t-1}$ 's perspective) if  $s_1$  realizes is the same with and without the deviation.<sup>33</sup> Now suppose that the sender in the deviation continues after outcome  $s_1$  (regardless of the design of experiment  $t$ ) as if this outcome were generated by an experiment with design  $(p, q)$ .<sup>34</sup> The expected continuation payoff if the outcome of experiment  $t$  in the deviation is  $s_1$  is, thus,  $u_1$  as without the deviation.

Each outcome  $s_j$ ,  $j = 1, 2$ , by construction realizes with the same probability with and without the deviation. The expected continuation payoffs after each outcome  $s_j$  are the same with and without the deviation. The deviation is profitable, as an experiment with design  $\pi_t = (1, 0)$  costs  $c_U$  which by assumption is lower than running an informative experiment at costs  $c_I$ . Q.E.D.

**Proof of Lemma 3:** At  $h$  the sender does not run exclusively further costly uninformative experiments with precision  $(1, 0)$ , since he already has a positive outcome from such an experiment and he can only send a message containing one outcome.

Suppose that the sender in a hypothetical equilibrium runs a further informative experiment  $\tau$  with certainty after  $h$ , that he persuades the receiver with a message containing a positive outcome from this informative experiment (regardless of whether she verifies the outcome) and that the sender stops experimenting if the outcome of this informative experiment is adverse in which case he sends a message containing the outcome from the uninformative experiment. Furthermore, suppose that the uninfor-

---

<sup>33</sup>Note that with the deviation, the actual posterior after outcome  $s_1$  depends on whether experiment  $t$  has design  $(p - \varepsilon_1, q + \varepsilon_2)$  or design  $(1, 0)$ . By construction the expected posterior, however, is as without the deviation.

<sup>34</sup>To illustrate consider the following three histories that have the same subhistory  $h_{t-1} = \{(\sigma_\tau, \pi_\tau)\}_{\tau=1, \dots, t-1}$ : (1) History  $h_t = \{(\sigma_\tau, \pi_\tau)\}_{\tau=1, \dots, t-1, t}$ , where the last outcome in  $h_t$  is  $s_1$  stemming from an experiment with design  $(p, q)$ . (2) History  $h'_t = \{(\sigma_\tau, \pi_\tau)\}_{\tau=1, \dots, t-1, t}$ , where the last outcome in  $h_t$  is  $s_1$  stemming from an experiment with design  $(1, 0)$ . (3) History  $h''_t = \{(\sigma_\tau, \pi_\tau)\}_{\tau=1, \dots, t-1, t}$ , where the last outcome in  $h_t$  is  $s_1$  stemming from an experiment with design  $(p - \varepsilon_1, q + \varepsilon_2)$ .

The statement means that the continuation decision at any history  $h_z = \{(\sigma_\tau, \pi_\tau)\}_{\tau=1, \dots, t-1, t, \dots, z}$  that has subhistory  $h'_t$  or  $h''_t$  is set equal to the continuation decision at a history that equals  $h_z$  with the exception that experiment  $t$  has design  $(p, q)$ .

mative experiment  $t$  in  $h$  is run with certainty after the sender observes the outcome of experiment  $t - 1$  in  $h$ . Modify the sender's strategy as follows. After observing the outcome of experiment  $t - 1$  in  $h$  the sender runs the informative experiment  $\tau$  with certainty. If the outcome of this experiment is positive, then he stops experimenting and sends a message containing this outcome. If the outcome of this experiment is adverse, then he runs the uninformative experiment with precision  $(1, 0)$ , after which he stops experimenting and where sends a message containing the positive outcome from the uninformative experiment. Consider the sender's expected payoff after observing the outcome from experiment  $t - 1$  in  $h$  from the hypothetical equilibrium strategy and the modified strategy. The persuasion probability is the same. However, the expected experimentation costs are lower with the modified strategy, as the informative experiment is run with the same probability, but the uninformative experiment is run with a lower probability with the modified strategy. Hence, there is no such equilibrium.

In other potential equilibria where the sender does not stop after running the uninformative experiment it is possible to replicate the persuasion probability at lower expected costs at the history before the uninformative experiment is run by constructing an analogous modified strategy. Q.E.D.

**Proof of Lemma 4:** According to Lemma 3 the sender anticipates that it is sequentially rational to stop experimenting after having run the uninformative experiment. He can only potentially persuade the receiver who does not verify a message with a positive outcome from the uninformative experiment with  $\pi_\tau = (1, 0)$ . The sender's benefit from running the uninformative experiment with  $\pi_\tau = (1, 0)$  is independent of the state of the world, as it yields a positive outcome regardless of the state, and, hence, his benefit is independent of the experimentation history.

Suppose there is a history where the sender strictly prefers running the uninformative experiment with  $\pi_\tau = (1, 0)$  (anticipating stopping afterwards) over stopping unsuccessfully. In this case the sender strictly prefers running the uninformative experiment (anticipating stopping afterwards) over stopping unsuccessfully at any experimentation history. Hence, the sender never stops experimenting without finding a

positive outcome. In such an equilibrium the receiver who does not verify the message knows that she observes a positive outcome regardless of the state of the world. Hence, the positive outcome is uninformative and it is a best response to choose  $a = a_2$  upon observing it. In such an equilibrium the sender cannot persuade the receiver with an outcome from the uninformative experiment regardless of whether the receiver verifies the message or not. As the uninformative experiment is costly, the sender prefers not running it over running it, yielding a contradiction.

Suppose there is a history where the sender strictly prefers stopping experimenting unsuccessfully over running the uninformative experiment with  $\pi_\tau = (1, 0)$  (anticipating stopping afterwards). In this case the sender strictly prefers stopping experimentation over running the uninformative experiment at any experimentation history. Consequently, he never runs the uninformative experiment, contradicting Proposition 2.

As the sender runs the uninformative experiment with  $\pi_\tau = (1, 0)$  with a positive probability according to Proposition 2, it follows that he must be indifferent between running this experiment and stopping experimentation unsuccessfully. Q.E.D.

## REFERENCES

- BROCAS, I. and CARILLO, J. D. (2007). “Influence through ignorance”, *RAND Journal of Economics*, Vol. 38, No. 4, pp. 931-947.
- CELIK, G. (2003). “Interested experts: Do they know more?”, *working paper*.
- CRAWFORD, V. P. and SOBEL, J. (1982). “Strategic information transmission”, *Econometrica*, Vol. 50, pp. 1431-1451.
- DARNSTÄDT, T. (2012). “Verfassung: Das Sudoku-Problem”, *Spiegel Online*, <http://www.spiegel.de/politik/deutschland/verfassung-das-sudoku-problem-a-818382.html>, accessed on 08/07/2016.
- DEWATRIPONT, M. and TIROLE, J. (2005). “Modes of communication”, *Journal of Political Economy*, Vol. 113, No. 6, pp. 1217-1238.
- DZUIDA, W. (2012). “Communication with detectable deceit”, *working paper*.
- ESŐ, P. and GALAMBOS, Á. (2013). “Disagreement and evidence production in strategic information transmission”, *International Journal of Game Theory*, Vol. 42, No. 1, pp. 263-282.
- FELGENHAUER, M. and LOERKE, P. (2016). “Bayesian persuasion with private experimentation”, *International Economic Review*, forthcoming.
- FELGENHAUER, M. and SCHULTE, E. (2014). “Strategic private experimentation”, *American Economic Journal: Microeconomics*, Vol. 6, No. 4, pp. 74-105.
- GENTZKOW, M. and KAMENICA, E. (2014). “Costly persuasion”, *American Economic Review Papers and Proceedings*, Vol. 104, No. 5, pp. 457-462.
- GLAZER, J. and RUBINSTEIN, A. (2001). “Debates and decisions: On a rationale of argumentation rules”, *Games and Economic Behavior*, Vol. 36, pp. 158-173.
- GLAZER, J. and RUBINSTEIN, A. (2004). “On optimal rules of persuasion”, *Econometrica*, Vol. 72, No. 6, pp. 1715-1736.
- GLAZER, J. and RUBINSTEIN, A. (2006). “A study in the pragmatics of persuasion: A game theoretical approach”, *Theoretical Economics*, Vol. 1, pp. 395-410.
- HENRY, E. (2009). “Strategic disclosure of research results: the cost of proving your honesty”, *The Economic Journal*, Vol. 119, pp. 1036-1064.

JOVANOVIĆ, B. (1982). “Truthful disclosure of information”, *Bell Journal of Economics*, Vol. 13, No. 1, pp. 36-44.

KAMENICA, E. and GENTZKOW, M. (2011). “Bayesian persuasion”, *American Economic Review*, Vol. 101, No. 6, pp. 2590-2615.

KARTIK, N. (2009). “Strategic communication with lying costs”, *The Review of Economic Studies*, Vol. 76, No. 4, pp. 1359-1395.

KING, N. J. (1996). “Juror delinquency in criminal trials in America, 1796-1996”, *Michigan Law Review*, Vol. 94, No. 8, pp. 2673-2751.

MILGROM, P. and ROBERTS, J. (1986). “Relying on the information of interested parties”, *The RAND Journal of Economics*, Vol. 17, No. 1, pp. 18-32.

MYERSON, R. B. and SATTERTHWAITE, M. A. (1983). “Efficient mechanisms for bilateral trading”, *Journal of Economic Theory*, Vol. 29, No. 2, pp. 265-281.

PEREZ-RICHET, E. and PRADY, D. (2012). “Complicating to persuade?”, working paper.