Signaling Expertise*

(preliminary and incomplete; please do not circulate)

Alexandros Rigos[†] Maria Kozlovskaya[‡] Matteo Foschi[§]

February 15, 2019

Abstract

This paper models strategic information transmission between buyers in need of a service and an expert seller who can provide it. Buyers are heterogeneously informed about their needs and so the seller can try to offer them unnecessary services. We assume better-informed buyer types can differentiate themselves from worse-informed types by providing verifiable evidence of their expertise. We show that, in equilibrium, there can be strong incentives to hide one's expertise. By selectively hiding their knowledge, partially-informed buyers can even completely protect themselves from seller fraud. Additionally, more-informed buyers can protect less-informed buyers from fraud by concealing their knowledge.

Keywords: Signaling; Hard Evidence; Credence Goods; Persuasion; Communication; Information Disclosure; Experts

JEL classification: C72, D82, D83, L15

*We thank Martin Kaae Jensen and Willemien Kets for the useful discussions and suggestions. We also thank audiences at the 2017 Lancaster Game Theory Conference, the 2018 International Conference on Game Theory at Stony Brook, Aston, ETH Zürich, Lund, and Manchester for the useful comments. All errors are our own. The usual disclaimer applies.

†Department of Economics, Lund University School of Economics and Management, Box 7082, 220 07, Lund, Sweden; alexandros.rigos@nek.lu.se.

‡Aston University, Birmingham, B4 7ET, UK; m.kozlovskaya@aston.ac.uk.

§Charles River Associates, 8 Finsbury Circus, London, EC2M 7EA, UK; mfoschi@crai.com.

1 Introduction

In markets with asymmetric information, better-informed parties may attempt to take advantage of less-informed ones. Typical examples come from the medical care, car repair and IT industries (Dulleck and Kerschbamer 2006). Moreover, larger information gaps between the parties can lead to more severe opportunistic behavior and resulting inefficiencies. If the size of the gap is unknown, can a less informed party affect the outcome by signaling their expertise to the better informed one? Would they like to appear as informed as possible? If signals are verifiable, will expertise be fully revealed in equilibrium?

We address these questions by modeling a heterogeneously informed population of buyers who may produce and selectively disclose verifiable evidence while participating in an transaction with fully informed sellers.

In our model, a Customer is in need of an Expert's services. There are two distinct kinds of services that Customer may require—service 1 and service 2—but he may need both. The Customer has one of four types: the clueless type receives no information about his needs; partially informed type 1 can observe whether or not he needs service 1, but receives no information about service 2; partially informed type 2 gets informed about service 2 but not service 1; the knowledgeable type is perfectly informed. When approached by the Customer, the Expert gets fully informed about the Customer's needs and makes a take-it-or-leave-it offer to provide either or both services. Therefore, the Expert can act opportunistically and attempt to provide unnecessary services. We allow Customers to selectively disclose verifiable *evidence about their information* before receiving the Expert's offer.

Our main results, in brief, are the following. Firstly, we find that heterogeneity of informedness can create situations where Customer with some (but not all) information are strictly better-off withholding some of their limited evidence from the fully informed Expert, thus preferring to appear less informed than they are. This is goes against intuition and received knowledge of information unraveling (e.g., Milgrom 1981). Secondly, we show that populations of Customers that hold enough shared knowledge can become fully protected from fraud by using pooling strategies (essentially withholding their level of informedness from sellers). Importantly, full protection does not require the presence—not even the existence—of fully-informed Customers.

The paper is related to the rich, but mostly distinct literatures on communication games, and credence goods models. Persuasion with hard evidence has been extensively studied in game-theoretic models of communication (Glazer and Rubinstein 2004; Hagenbach and Koessler 2017; Hart, Kremer, and Perry 2017; Rubinstein and Glazer 2006). Our approach is complementary to existing literature: compared to the classical sender-receiver model, we consider a more specific state space but a more general communication structure, where both parties can send messages to each other. The main insights of our approach is that a less-informed party in a two-way communication scenario may want to withhold some of their limited evidence from a better-informed party.

In a more applied credence goods context, the exploitation of poorly-informed buyers by better-informed sellers has been considered in the industrial organisation literature (Dulleck and Kerschbamer 2006; Emons 1997; Hyndman and Ozerturk 2011; Schneider, Meub, and Bizer 2016; Wolinsky 1993). We contribute to this strand of research by introducing heterogeneity in buyers' information, and allowing the buyers to communicate their knowledge to the seller. Our results broaden the spectrum of policy instruments aimed at improving consumer welfare on credence goods markets. Existing research has suggested various seller-targeted remedies to buyer exploitation, including warranties, industry standards, and separation of diagnosis and treatment (Dulleck and Kerschbamer 2006). Our analysis implies that global buyer protection can be achieved by targeted educational interventions, by improving the knowledge of a small group of buyers. Importantly for an institutional designer, we show that fraud can be prevented even in single-seller economies with one-shot transactions. This is the main applied contribution of our paper to existing literature, which established competition and reputation as the two main channels in disciplining sellers (Dulleck and Kerschbamer 2006; Wolinsky 1993).

The rest of the paper is organised as follows. The model of signaling expertise is introduced in Section 2, signaling incentives are derived in Section 3 and the model is solved for equilibrium in Section 4. We briefly discuss our findings and point at possible paths for future research in Section 5. The Appendix contains the derivation of best responses on which the paper's propositions are based.

2 The Model

We consider a situation where a (potentially) less informed agent needs the help of an expert agent to treat a problem he is faced with. To illustrate the idea, think of the driver of a broken-down car and a mechanic. There are two parts that may need replacement in the car: the engine and the transmission. On the one hand, the mechanic knows which parts need replacement. On the other hand, unbeknown to the mechanic, the driver may be able to observe the state of both parts, a single part, or neither part. The timing of the interaction is as follows (see also Figure 1). The driver sends a message to the mechanic, selectively disclosing pieces of information that he may have (see Section 2.7). Upon receiving the driver's message, the mechanic offers to fix one or both parts of the car. Finally, the driver either accepts or rejects the mechanic's offer, thus ending the interaction. The model is laid out formally in the remainder of this section. Table 1 is provided as a quick reference to the notation used. Throughout the paper, notation in Greek lowercase letters is used to denote probability distributions.

Table 1: Exogenous variables and their meaning

	Concept	Notation	Instance	Interpretation
Model elements	Problem	w	w_{10}	Issue 1 needs fixing; Issue 2 does not need fixing
	Customer type	t	t_{10}	Can observe issue 1; Cannot observe issue 2
			t_{*0}	Set of types who cannot observe issue 2 ($\{t_{00}, t_{10}\}$)
	Customer info set	I	I_{10}	Issue 1 needs fixing; Issue 2 does not need fixing
			I_{*1}	Issue 2 needs fixing
		$I(w,t)$		Info set of type t when the problem is w
	Treatment	o	o_{10}	An offer to fix issue 1 and to not fix issue 2
	Message	m	m_{10}	Disclose issue 1's state; Do not disclose issue 2's state
Parameters	Problem prior	π	$\pi_{10}, \pi(w_{10})$	(Customer's) Prior probability of problem w_{10}
	Type prior	ϱ	$\varrho_{10}, \varrho(t_{10})$	(Expert's) Prior probability of type t_{10}
	Treatment price	p	p_{10}	The price of fixing issue 1 and not fixing issue 2
	Price of honesty	h	h_{10}	Premium of having w_{10} fixed by the honest expert

2.1 Players and problems

The players in our model are a Customer (he) and an Expert (she). The Customer suffers from a problem and requires the Expert's help in order to have the problem treated. A problem is defined by the state of each of two issues—issue 1 and issue 2. Each issue's state can be either 0 or 1. If the state of an issue is 0, this indicates that the Customer does not need to have this issue treated whereas if the state of the issue is 1, the issue needs to be treated. We assume that the Customer knows that he needs the Expert's services, so at least one of the two issues should require treatment. Therefore, a problem (or state of the world) is an element $w \in \Omega \equiv \{0,1\}^2 \setminus (0,0)$. For notational clarity and brevity reasons, we use w_{10}, w_{01}, and w_{11} to denote the relevant problems (instead of $(1,0)$, $(0,1)$, and $(1,1)$, respectively). The way problem w_{10} should be interpreted is that the Customer suffers from issue 1 and does not suffer from issue 2, whereas in state w_{11} the Customer suffers from both issue 1 and issue 2. The Customer has a prior belief $\pi \in \Delta(\Omega)$ over the problem space.[1] So, $\pi_{10} \equiv \pi(w_{10})$ is the probability with which problem w_{10} occurs. The belief π is assumed to be common knowledge and has full support.

[1]Throughout the paper we use $\Delta(X)$ to denote the set of probability distributions over space X.

2.2 Customer types

The Customer can be one of four types based on his ability to identify the problem at hand. A type is an element $t \in T \equiv \{0,1\}^2$ and is characterized by the set of issues whose state he can observe. So, type t_{11} is the knowledgeable/expert type and is always able to identify the exact problem he suffers from. Types t_{10} and t_{01} are partially informed types who can only observe the state of one of the two issues (issue 1 and 2, respectively). Finally, type t_{00} is a clueless type and does not get to observe any of the issues. The Expert does not know which type she is facing and has a prior belief $\varrho \in \Delta(T)$ over the type space. We write $\varrho_{10} \equiv \varrho(t_{10})$ to denote the ex-ante probability with which the Expert believes that she is facing a t_{10}-type Customer. Types and problems are independently distributed, so that all types face the different problems with the same probabilities. The prior ϱ is assumed to be common knowledge and correct, i.e., it reflects the probability that a person in the population is of a particular type.

2.3 Information sets

According to the above typology, given his type and the state of the world, the Customer can find himself in one of several information sets. Each information set contains all information that the Customer has when he goes to the Expert. We denote information sets by I and appropriate subscripts, using the asterisk ($*$) to denote missing information. For example, I_{*1} is the information set where the Customer cannot observe the state of issue 1 but can observe the state of issue 2 being 1. Put in words, the Customer at this information set knows that issue 2 needs fixing but is unsure about whether issue 1 also needs fixing or not). There are two states consistent with this information set: w_{11} and w_{01}. Clearly, the only type that can be at this information set is t_{01}. This is more general: *there is only one type that can be at any given information set.* A list of the types and the information sets in which they can find themselves is given in Table 2. We also use the notation $I(w,t)$ to denote the information set in which type t finds himself when the problem is w. The set of information sets is denoted by $\mathscr{I}$. Observe that since it is common knowledge that at least one issue needs treatment (i.e. only states w_{01}, w_{10}, and w_{11} are possible), the Customer at information set I_{0*} knows that the state *has to be* w_{01}—even though he cannot directly observe this. Similarly, at information set I_{*0} the Customer knows the state to be w_{10}.

2.4 Treatments and prices

The Expert will offer to treat the Customer's problem with one of three treatments. Similar to a problem, a treatment is an element $o \in \{0,1\}^2 \setminus (0,0)$ with o_{10} having the intepretation

Table 2: Customer types, possible information sets for each type, and messages available to each type

Type $[t]$	Information sets	Available messages $[M(t)]$
t_{00}	I_{**}	m_{00}
t_{10}	I_{0*}, I_{1*}	m_{10}, m_{00}
t_{01}	I_{*0}, I_{*1}	m_{01}, m_{00}
t_{11}	I_{01}, I_{10}, I_{11}	$m_{11}, m_{01}, m_{10}, m_{00}$

that the Expert offers to treat issue 1 and to not treat issue 2. Receiving a treatment is costly. The price of treatment o_i is denoted by $p(o_i)$ and abbreviated as p_i. Without loss of generality, treatment o_{01} is (weakly) more expensive than treatment o_{10}. Moreover, treating both issues is more expensive than treating a single issue i.e. $p_{11} > p_{01} \geq p_{10} > 0$.

Importantly, the Expert is only allowed to offer treatments that can actually fix the Customer's problem (but potentially more). So, if the problem is w_{10} (w_{01}), the Expert can offer either o_{10} (o_{01}) or o_{11} whereas if the problem is w_{11}, then the only treatment that the Expert can offer is o_{11}. This assumption is referred to as *liability* in the credence goods literature (introduced by Pitchik and Schotter 1987; see Dulleck and Kerschbamer 2006, for more literature using this assumption). It captures the idea that the Expert is legally liable if she provides inadequate treatment to the Customer. Another way to think about this in the mechanic example is that the driver can realize that, if undertreated, the car does not function the way it should, in which case he does not have to pay for its treatment. This assumption allows us to rule out the case of undertreatment. We further assume that the treatment is *verifiable*, i.e. that the Customer can ensure that he is actually receiving the treatment for which he pays (see Dulleck and Kerschbamer 2006). With this assumption we rule out the possibility for overcharging, i.e., the possibility of the Expert charging for treatments that she hasn't carried out.

2.5 Honest expert

As an outside option, if the Customer rejects the Expert's offer, he can get the problem fixed by an honest expert, who is unmodelled. It is assumed that the honest expert always fixes the actual problem from which the Customer suffers, but at a price higher than the one that the Expert charges. The premium that the honest expert charges to fix problem w_i is denoted by $h(w_i)$ and abbreviated as $h_i > 0$. We will often refer to h as the price of honesty. In our working example, the honest expert represents a car dealership: they will always fix the exact problem that the car has but the treatment is more expensive.

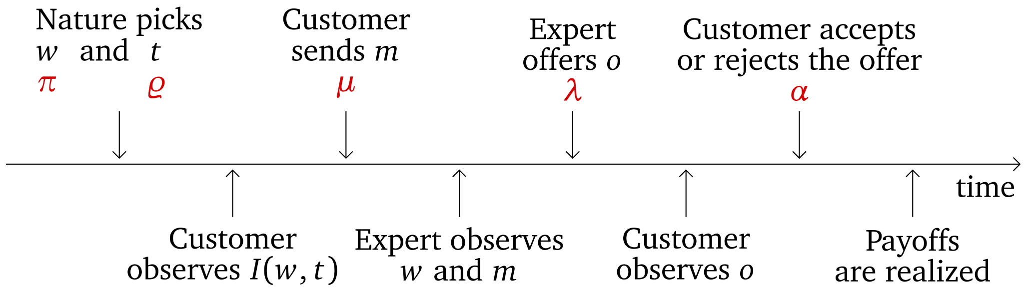


Figure 1: Timing of events and decisions. Variables in red Greek characters indicate the distributions that the respective variables follow.

Assumption 1. *In order to rule out trivial situations in which all Customer types accept any offer that the Expert makes, we require that the honesty premium be not too large. In particular, we assume that*

$$A_i \equiv p_{11} - (p_i + h_i) > 0 \qquad \textit{for } i \in \{10, 01\}. \tag{1}$$

The above condition implies that if the Customer knows that his problem is w_{10} (by being either at information set I_{10} or I_{*0}) and the Expert offers treatment o_{11}, the Customer will certainly reject the Expert's offer (see 2.6 and A below); similarly for w_{01}.

2.6 Payoffs

The assumptions of Sections 2.4 and 2.5 guarantee that the Customer's problem is always going to be treated sufficiently, either by the Expert or by the outside-option honest expert.

If the state is w_i and the Expert's offer is o_j, accepting the offer costs the Customer p_j whereas rejecting the offer bears a cost of $p_i + h_i$ (as this is how much the honest expert charges). On the flipside, if the Customer accepts the offer, the Expert receives a revenue of p_j whereas if the Customer rejects the offer, the Expert receives a revenue of 0.

The Customer's goal is to get the problem fixed at the lowest possible cost. In particular, the Customer is trying to *minimize expected payment*. As for the Expert, we assume that she is trying to *maximize her expected revenue.*[2]

2.7 Messages/Language

The Customer sends a message to the Expert before she decides which treatment to offer (see Figure 1). The messages that the Customer can send are type-dependent. We consider

[2]There is an underlying assumption that the Expert's profit (and, therefore, cost) is linear in the price charged. More complicated cost structures would only convolute the analysis without providing additional intuition.

a language structure of *hard evidence*: the Customer can selectively disclose information that he has about one or both issues.

Messages are denoted by m and appropriate subscripts, indicating which pieces of information the Customer is disclosing. When, for example, the Customer sends message m_{10}, he discloses that he knows the state of issue 1 and does not disclose the state of issue 2. Naturally, in order to disclose the state of an issue, the Customer needs to know it.[3] So, the aforementioned message can only be sent by types that can observe issue 1, i.e., types t_{10} and t_{11}. An exhaustive list of the messages available to each type is given in Table 2. The message space is denoted by M while the set of messages available to type t is denoted by $M(t)$. The set of types that can send message m is denoted by $T(m)$.

Definition 1. *A tuple (p, h, π, ϱ) will be called an expertise signaling game.*

Based on the above, in Sections 2.8 and 2.9, we describe what agents in our model take into account when they make decisions, the format of the strategies they follow, and the way they form and update beliefs.

2.8 Strategies

Customer: Signaling The Customer's signaling strategy prescribes with what probability he should send each message, conditional on his information set. It is, therefore, a mapping $\mu : \mathscr{I} \to \Delta(M)$ under the restriction that the type t sending the message (as identified by the information set, see Section 2.3) assigns positive probability only to messages that he can send (i.e., in $M(t)$, see Table 2 for details). We will write $\mu(m|I)$ to denote the probability with which the Customer sends message m at information set I.

Expert: Cheating The Expert's strategy prescribes with what probability she should offer each treatment, conditional on the problem she observes and on the message she received. Given the assumptions of Section 2.4, if the state is w_{11}, then the only treatment she can offer is o_{11}. Moreover, if the state is w_{10} or w_{01}, then she can only offer to fix the actual problem that the Customer faces (o_{10} or o_{01}, respectively) or to fix both issues (o_{11}). Therefore, the Expert's strategy can be summarized by $\lambda : \Omega \times M \to [0, 1]$. The interpretation of $\lambda(w_i, m_j)$ is the probability with which the Expert offers treatment o_{11} when the Customer's problem is w_i and he sent message m_j. Trivially, it has to be that $\lambda(w_{11}, m) = 1$ for any $m \in M$. When $i \in \{10, 01\}$, $\lambda(w_i, m)$ is interpreted as the Expert's cheating probability when the problem is w_i and the message she received is m.

[3] In this sense, the Customer provides *hard proof* that he knows the state of the issue. This could be, for example, through demonstrating that a particular part of the car is functioning as it should or not.

Customer: Accepting/Rejecting The Customer's strategy should prescribe with what probability to accept each particular treatment offered by the Expert, conditional on the Customer's information set and the message he sent. Given the assumptions of Section 2.4, if the Expert offers treatment o_i with $i \in \{10, 01\}$, then the problem has to be w_i and o_i is the cheapest treatment that can fix the problem. So, in an optimal strategy, the Customer should always accept offers o_{10} and o_{01}. The only question that remains for the Customer to answer is whether to accept an offer of o_{11} by the Expert or not. We, therefore, summarize the Customer's strategy by $\alpha : \mathscr{I} \times M \to [0, 1]$. The interpretation of $\alpha(I, m)$ is the probability with which the Customer accepts an offer of o_{11} by the Expert when his information set is $I \in \mathscr{I}$ and he has sent message m. It is implied that the Customer always accepts offers of o_{10} and o_{01}.

2.9 Beliefs

Upon observing his information set I, the Customer updates his belief about the problem. This happens by excluding any problems that are inconsistent with I and adjusting using Bayes's rule. Explicitly, let Ω_{ij} denote the set of problems that are consistent with information set I_{ij}, i.e., $\Omega_{ij} = \{w_{kl} \in \Omega : i \in \{k, *\} \text{ and } j \in \{l, *\}\}$. Then the Customer's belief is updated to be

$$\pi'_{kl}(I_{ij}) \equiv \Pr(w = w_{kl}|I_{ij}) = \begin{cases} \pi(w_{kl})\left(\sum_{w \in \Omega_{ij}} \pi(w)\right)^{-1} & \text{if } w_{kl} \in \Omega_{ij} \\ 0 & \text{otherwise} \end{cases}. \tag{2}$$

Similarly, upon receiving the Customer's message, say m, the Expert updates her belief about the Customer's type. The Customer has to be of a type that can send the message, i.e., $t \in T(m)$. Moreover, the beliefs are updated using Bayes's rule. This leads to the Expert's updated beliefs being

$$\varrho'(t|w, m) = \varrho(t)\mu(m|I(w, t))\left(\sum_{t' \in T(m)} \varrho(t')\mu\big(m|I(w, t')\big)\right)^{-1}. \tag{3}$$

Finally, upon receiving the Expert's offer o, the Customer updates his beliefs once more. If he is offered treatment o_{10} or o_{01}, then he knows that the state is certainly w_{10} or w_{01}, respectively, as a result of the assumptions of Section 2.4. After being offered treatment o_{11} by the Expert, however, the Customer updates his belief using Bayes's rule:

$$\pi''_{kl}(I, m, o_{11}) \equiv \Pr(w = w_{kl}|I, m, o_{11}) = \pi'_{kl}(I)\lambda(w_{kl}, m)\left(\sum_{w' \in \Omega} \Pr(w = w'|I)\lambda(w, m)\right)^{-1}. \tag{4}$$

The incentives presented in the next Section and the equilibria of Section 4 are derived through application of the best responses that are found in the Appendix.

3 Signaling incentives

In this section we examine the signaling incentives of the different Customer types. The solution concept we use throughout is perfect Bayesian equilbirium. We refine beliefs off the equilibrium path using the intuitive criterion (Cho and Kreps 1987), when needed. The way these incentives interact with one another to yield equilibrium outcomes is studied in Section 4. In the interest of expositional clarity, we focus on expertise signaling games under generic parameter configurations and on equilibria in which each Customer type follows a *pure signaling strategy*. No additional intuitions are gained through studying mixed signaling strategies.

Notice that the signal structure of our model allows more sophisticated Customer types to separate themselves from less sophisticated ones. Moreover, it also allows them to condition their decision to separate or not on the information that they have. However, less sophisticated types cannot actively make a decision to separate themselves from other types. Naturally, one might wonder whether this is important since signaling one's expertise should always be beneficial. We show that there are cases in which hiding one's expertise is beneficial. In fact, these cases are not extreme.

3.1 Knowledgeable Customers (t_{11})

The knowledgeable Customer (type t_{11}) can ensure the best payoff for himself by always sending message m_{11}, i.e., by disclosing his type. In this case, he achieves the smallest possible payment of p_i when the state is w_i. This happens because the type t_{11} Customer optimally rejects treatment o_{11} if the problem is w_{10} or w_{01} and optimally accepts treatment o_{11} when the problem is w_{11}. That is,

$$\alpha(I_{10},m)=0 \qquad \alpha(I_{01},m)=0 \qquad \alpha(I_{11},m)=1 \qquad \text{any } m\in M.$$

So, knowing that (since message m_{11} uniquely identifies type t_{11} and type t_{11} rejects any other offer), the Expert can receive a positive payment only if she offers $o_i = w_i$. So, in the Expert's best response to m_{11} we have that

$$\lambda(w_{10},m_{11})=\lambda(w_{01},m_{11})=0.$$

Therefore, a t_{11}-type Customer would choose to send some message m' other than m_{11} only if he could guarantee that he gets the same payoff.

If $w=w_{11}$, the message sent by the Customer does not make any difference for the Expert's decision as she will always offer o_{11}, independently of the message. If, $w\in\{w_{10},w_{01}\}$, though, for the type t_{11} Customer to send message $m'\neq m_{11}$, it has to be that (see equation (10))

$$p_i+\lambda(w_i,m')h_i=p_i \quad\Rightarrow\quad \lambda(w_i,m')=0.$$

3.2 Partially informed Customers (t_{10} and t_{01})

In a similar fashion, when at information set I_{0*} (I_{*0}), a Customer of type t_{10} (t_{01}) can guarantee himself the best outcome by sending m_{10} (m_{01}). The only types that can send the message m_{10} (m_{01})—types t_{10} (t_{01}) and t_{11}—have a posterior π' that assigns all probability to problem w_{01} (w_{10}) and, hence, reject a treatment of o_{11}. Therefore, in her best response, the Expert offers treatment o_{01} (o_{10}) after receiving message m_{10} (m_{01}) when the problem is w_{01} (o_{10}), i.e.,

$$\lambda(w_{10}, m_{01}) = \lambda(w_{01}, m_{10}) = 0.$$

Consequently, when at information set I_{0*} (I_{*0}), the Customer of type t_{10} (t_{01}) will hide the information he has (i.e., send message m_{00}) only if $\lambda(w_{10}, m_{11}) = 0$.

On the contrary, when at information set I_{1*} (I_{*1}), a t_{10} (t_{01}) type Customer can have incentives to hide his information. In particular, we provide the following proposition.

Proposition 1. *In the absence of knowledgeable types ($\varrho(t_{11}) = 0$), in any equilibrium, a Customer at information set I_{1*} or I_{*1} has a (possibly weak) incentive to withhold his information.*

Proof. All proofs are available upon request. □

The intuition behind the above result is the following. In a state of the world where only one of the two issues need fixing, the offer to fix both is fraudulent. There are two Customer types which could fall for the fraudulent offer: the clueless type, and one of the partially informed types (in particular, the one who can only observe the issue which actually needs fixing). The clueless type assigns a prior probability to all problems, whereas the partially informed type can rule out one of the problems (where only the issue he does not observe needs fixing) and hence updates his prior on the state where both issues need fixing upwards. It follows that, before receiving the Expert's offer, the clueless Customer assigns a strictly lower probability to both issues needing a fix than a partially informed Customer. Hence, the Expert is less likely to lie to the clueless type, because he is less willing to accept a fraudulent offer. As a result, it benefits the partially informed type to pool with the clueless type to decrease the probability of being defrauded.

The result of Proposition 1 is stark: in a world without knowledgeable/expert Customers (which is most likely the case in the real world), partially informed Customers *never* have a strong incentive to signal themselves out when they are uncertain about the state. In fact, in Section 4.3.2 we show that witholding information in this manner can be sufficient to completely shield the partially informed type from fraud.

4 Equilibrium

In this section we look into equilibrium signaling strategies under different parameter combinations. We begin by giving some intuitions of outcomes in homogenous Customer populations, in which types can be perfectly identified by the Expert (Section 4.1). We proceed by identifying conditions under which separating equilibria exist (Section 4.2). Finally, consituting our major results, we characterize equilibria in which certain Customer types pool together, leading to a decrase in fraud by the Expert (Section 4.3). We provide Figure 2 as a visual depiction of the signaling strategies and Expert cheating behavior in our different equilibrium classes.

4.1 Homogenous populations

The analysis of 3.1 shows that in any equilibrium, the knowledgeable Customer type is never cheated.

Populations that consist of a single, imperfectly informed type $t \neq t_{11}$ will always be cheated on with some positive probability in equilibrium. In particular, a partially informed Customer, when identified by the Expert, will never be cheated when the issue he can observe does not need fixing but will be cheated with positive probability when the issue he can observe is the only one that needs fixing.

In equilibrium, a clueless Customer—when identified by the Expert—will be cheated with positive probability when the problem is w_{10}. Depending on parameter conditions, he may or may not be cheated with positive probability when the problem is w_{01}. In particular, he is going to be cheated with positive probability under problem w_{01} if and only if $\pi_{11}h_{11} > \pi_{10}A_{10}$. In this case, the Expert always cheats under problem w_{10}.

4.2 Full separation

We now turn to heterogeneous Consumer populations in which all types are present (ϱ has full support). A feature of our signaling structure used is that more informed types are always able to separate themselves from less informed ones, whereas the converse is not true. Therefore, if revealing that one is informed is the "best" thing one can do, we would expect that the only equilibria of an expertise signaling game would be separating, i.e., equilibria in which each type sends a message that no other type sends.[4] The next proposition provides a necessary and sufficient condition for the existence of such equilibria.

[4]The signaling strategy in a separating equilibrium is given by: $\mu(m_i|I(w,t_j)) = 1$ if $i = j$, $\mu(m_i|I(w,t_j)) = 0$ if $i \neq j$ $(i, j \in \{01, 10, 11\}, w \in \Omega)$.

Proposition 2. *The expertise signaling game $G = (p, h, \pi, \varrho)$ has a separating equilibrium iff*

$$\pi_{11}h_{11} \geq \pi_{10}A_{10} + \pi_{01}A_{01}. \tag{5}$$

So, separating equilibria exist only if w_{11} occurs too often or h_{11} is too high. In particular, condition (5) says that the ex-ante expected benefit from accepting treatment o_{11} when it is truthful (and, therefore, avoiding paying the extra h_{11}) outweighs the expected loss of accepting a fraudulent offer of o_{11}. This makes clueless Customers always willing to accept o_{11} and drives away the strong incentive to hide information that type t_{01} has under other parameter combinations (see Proposition 1). The signaling strategy followed by Consumers in a separating equilibrium is graphically depicted in Figure 2a.

In a separating equilibrium the Expert offers treatment o_{11} unless she knows that the type she is facing is aware that the problem is not w_{11}, while she offers to fix the actual problem w otherwise. The Customer rejects an offer iff he knows the offer to be fraudulent.

4.3 Pooling

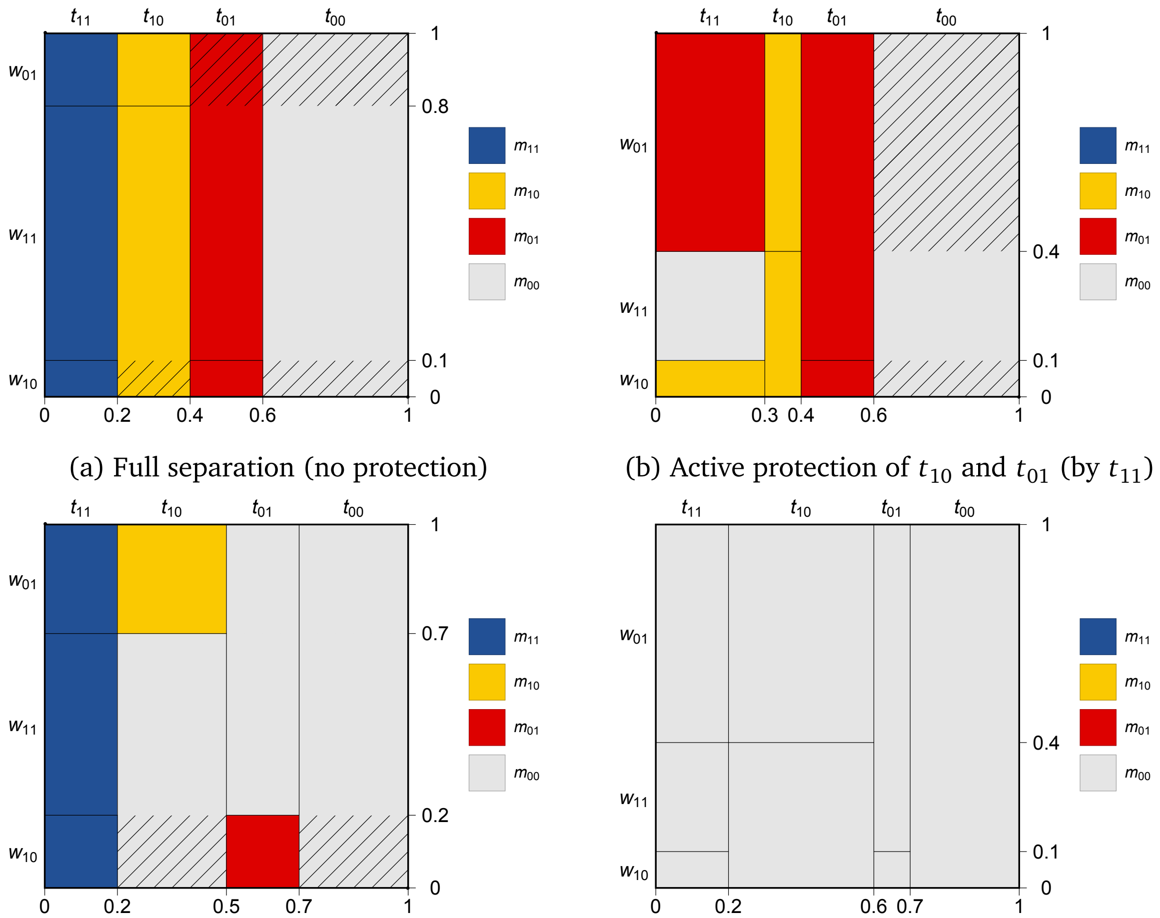
When the Customer population is heterogeneous, there is scope for (fully or partially) pooling equilibria to appear. Importantly, contrasting separating equilibria, there are equilibria in which when the problem is $w \in \{w_{10}, w_{01}\}$ there is at least one non-fully-informed type t who is never cheated—even though t would be cheated with positive probability under problem w if the other types were absent (see 4.1). In these equilibria, each such type t is able to pool with some other type t' who rejects fraudulent o_{11} offers with high enough probability, thus making the Expert to never cheat. This pooling is achieved by the two types sending the same message m at their respective information sets ($I(w, t)$ and $I(w, t')$). We will say that type t is protected by type t' under problem w (or that t' protects t under problem w).

There are two reasons why type t' may reject a fraudulent offer of o_{11} with high enough probability: (a) because he can correctly identify that one of the two issues does not require fixing (e.g. $t' \in \{t_{10}, t_{11}\}$ and $w = w_{01}$) or (b) because he believes that offer o_{11} is very likely to be fraudulent. We explore these different kinds of protection in what follows.

4.3.1 Active protection

In the first case, the "protector" type t' is more informed than the protected type, as he is fully informed about what the problem is.[5] Protection happens by t' hiding some information that he has, essentially not disclosing the issue that he knows does not need fixing.

[5]This can happen directly, in case of $t' = t_{11}$, or indirectly in case of $t' = t_{10}$ under problem w_{01} or $t' = t_{01}$ under problem w_{10} (see 2.3 and 3.2).



(a) Full separation (no protection)

(b) Active protection of t_{10} and t_{01} (by t_{11})

(c) Passive protection of t_{01} (by t_{00})

(d) Full protection

Figure 2: Examples of (pure) signaling strategies in equilibria. Each rectangular region is associated with a Customer information set with its size reflecting the probability with which it occurs according to the prior distributions of problems and Customer types (π and ϱ, respectively). The color of each region indicates which message the Customer sends at that information set. Shaded rectangles indicate that the Expert fraudulently offers o_{11} with positive probability. In all diagrams the following values are assumed: $p_{11} = 1$, $p_{10} = 0.5$, $p_{01} = 0.8$, $h_{11} = h_{10} = h_{01} = 0.1$.

In doing so, he allows the less-informed type t to pool with him. In this sense, this type of protection is *active*, as it is a result of choice (in contrast to passive protection of 4.3.2 below).[6]

To be more explicit, active protection of partially informed t_{10} (t_{01}) Customers by knowledgeable t_{11} Customers happens by the two types pooling on message m_{10} (m_{01}) under

[6] Notice that since types t_{10} and t_{11} can guarantee themselves to be offered the "fair" treatment when the problem is w_{01}, they will choose to not disclose issue 1 (by sending m_{00} or—in case of t_{11}—m_{01}) only if they are guaranteed to get offered o_{01} under that message (see 3.1 and 3.2). Similarly for types t_{01} and t_{11} when the problem is t_{10}.

problem w_{10} (w_{01}). Referring to our mechanic example, when a knowledgeable driver observes that his car's engine needs fixing but the transmission is working fine, he only demonstrates that the engine needs fixing, without mentioning the transmission. Similarly, when an engine-observing driver observes that his car's engine needs fixing, he demonstrates that. Consequently, upon receiving the message/demonstration that the engine needs fixing, the mechanic believes that the driver she is facing is very likely to be knowledgeable and would thus reject a fraudulent offer. Considering that, she makes the fair offer to only fix the engine. An example of the signaling strategy in an equilibrium where knowledgeable Customers actively protect both partially-informed-type Customers is depicted in Figure 2b.

Active protection can also extend to sets of (protector or protected) types. In particular, protection of t_{10} (t_{01}) and t_{00} types by t_{01} (t_{10}) and/or t_{11} under problem w_{10} (w_{01}) happens by the types pooling on message m_{00}, i.e., by not disclosing any information (see Figure 2d).

Let t_{1*} (t_{*1}) denote the set of types who can observe issue 1 (2). Similarly, let t_{0*} (t_{*0}) denote the set of types who cannot observe issue 1 (2). Let us also abuse notation and use $\varrho(t_{1*})$ to denote the probability $\Pr(t \in t_{1*})$ according to the prior ϱ. Similarly for $\varrho(t_{*1})$ and so on. Then necessary and sufficient conditions for active protection to occur are given in the following proposition.

Proposition 3. *A pooling equilibrium where less-informed Customers (t_{0x}) are protected from fraud under problem w_{01} by more-informed Customers (t_{1x}) exists iff the relative share of more-informed Customers exceeds the Expert's relative benefit from lying:*

$$\frac{\varrho(t_{1x})}{\varrho(t_{1x})+\varrho(t_{0x})} > \frac{p_{11}-p_{10}}{p_{11}}, \tag{6}$$

where $x = *$ if both types in t_{0*} are protected and $x = 1$ if only the partially informed Customers t_{01} are protected. Similarly for protection of t_{x0} by t_{x1} under problem w_{10}.

Notice that as soon as clueless Customers t_{00} are protected under problem w_{10} (w_{01}), partially informed t_{10} (t_{01}) Customers will opt to also send m_{00} when at information set I_{1*} (I_{*1}); unless they are protected by t_{11} directly through message m_{10} (m_{01}).[7] So, protection through m_{00} has to be able to "cover" both t_{00} and t_{10} (t_{01}) Customers.

Readily from Proposition 3 one obtains the following result.

Corollary 1. *An expertise signaling game has an equilibrium in which the Expert never cheats iff*

$$\varrho(t_{11})+\varrho(t_{10}) > \frac{p_{11}-p_{10}}{p_{11}} \quad \textit{and} \quad \varrho(t_{11})+\varrho(t_{01}) > \frac{p_{11}-p_{01}}{p_{11}}.$$

[7] In this case t_{11} protects t_{10} (t_{01}) by sending message m_{10} (m_{01}), whereas type t_{01} (t_{10}) protects t_{00} by sending message m_{00}. The parametric requirements for such an equilibrium are stricter than the requirements for an equilibrium where types t_{11} and t_{01} (t_{10}) protect t_{10} (t_{01}) and t_{00} by sending message m_{00} and so the relevant version of condition (6) also holds.

We will refer to such an equilibrium as a full protection equilibrium as all types are protected from fraud. An example of a signaling strategy in a full protecion equilibrium is one where all types send message m_{00} in all states of the world (see Figure 2d).[8]

It is intuitive and straightforward that one's own payoff should get better as one gets more expertise. Proposition 3 shows that more expertise in some part of the population can also help others achieve better outcomes. In this sense, increasing expertise in the Customer population can have positive externalities on less-expert Customers. Importantly, if Customers collectively have enough expertise, they can ensure full protection from fraud, even in the absence of knowledgeable t_{11} types.

4.3.2 Passive protection

Due to the incentive analysed in Proposition 1, unless type t_{01} is actively protected under problem w_{01} by type t_{11} (by t_{11} sending m_{01}), he has at least a weak incentive to send message m_{00} when at information set I_{*1}. Importantly, this incentive can be strong if t_{00} is likely enough to reject an offer of o_{11}—even to the extent that t_{01} can get completely protected from fraud via pooling with t_{00}. The next proposition provides sufficient conditions for this to happen.

Proposition 4. *Let $G = (p, h, \pi, \varrho)$ be an expertise signaling game and let π and ϱ have full support. Let also $\frac{\varrho(t_{01})}{\varrho(t_{00})+\varrho(t_{01})} < \frac{p_{01}}{p_{11}}$ and $\pi_{11}h_{11} < \pi_{10}A_{10}$. The game G has an equilibrium in which t_{01} is protected by t_{00} under w_{01} iff one of the following conditions holds.*

A) $\frac{\varrho(t_{10})}{\varrho(t_{00})+\varrho(t_{10})} > \frac{p_{10}}{p_{11}}$.

B) $\frac{\varrho(t_{10})}{\varrho(t_{00})+\varrho(t_{10})} < \frac{p_{10}}{p_{11}}$ *and* $p_{10} < p_{01}\frac{\varrho(t_{00})+\varrho(t_{01})}{\varrho(t_{00})+\varrho(t_{10})} + p_{11}\frac{\varrho(t_{10})-\varrho(t_{01})}{\varrho(t_{00})+\varrho(t_{10})}$.

Symmetrically (substituting index $_{01}$ for $_{10}$ and vice versa) for protection of t_{10} by t_{00}.

In the equilibria of Proposition 4, both partially informed types (t_{10} and t_{01}) send m_{00} when uncertain about the state of the world (information sets I_{1*} and I_{*1}, respectively). While—when uncertain about his problem—one of the two types (t_{10}) is indifferent between disclosing his information and not, the other type (t_{01}) *strictly prefers to not disclose his information,* as the Expert never cheats under w_{01} after receiving message m_{00}.

The presence of the clueless types t_{00} is crucial in this result: their inability to distinguish between problems w_{10} and w_{01} is precisely the reason why the Expert is reluctant to cheat under w_{01} (too much cheating will be rejected; therefore the expert stops cheating in the

[8] There can be full protection equilibria where the signaling strategy differs from sending m_{00} at all info sets. For example, the message used by the Customer at information set I_{11} is payoff-irrelevant as the Expert will offer treatment o_{11} independently of what message she receives. Nevertheless, the conditions of Corollary 1 should be satisfied in any such equilibrium.

least profitable scenario). This reaction of the Expert, consequently, gives a strong incentive to type t_{01} to not disclose his information, thus pooling with t_{00} and receiving t_{00}'s passive protection. An example of the signaling strategy in an equilibrium with passive protection is depicted in Figure 2c.

Notice that the conditions of Proposition 4 imply that type t_{01} can be passively protected even in the absence of t_{10} types. A possible symmetric case where t_{10} is passively protected in the absence of t_{01} *does not* occur. This has to do with the asymmetry of prices ($p_{10} < p_{01}$) and the way this asymmetry affects type t_{00}'s indifference conditions (see also Section 4.1).

Finally, there also exist protection equilibria in which partially informed Customers resolve their indifference between disclosing and withholding information in favor of disclosing. We deem these equilibria fragile as even the slightest cost of information disclosure would suffice for them to break down. On the contrary, the passive protection equilibria of Proposition 4 are robust to the introduction of (small) disclosure costs as the types involved disclose no information to begin with. A further discussion of protection robustness follows.

4.3.3 Protection is robust

At the end of Section 4.3.2, we briefly touched upon which equilibria we consider plausible. While the discussion there related to passive protection, here discuss the robustness of active protection equilibria.

On the one hand, the result of Proposition 3 shows that active protection can occur *for any prior over problems* π. On the other hand, other types of equilibria can occur under the same parametric conditions. For example, by taking a look at Proposition 2, the necessary and sufficient condition for full separation is essentially a condition on π, a parameter orthogonal to ϱ. Moreover, as noted before, any type t' who is actively protecting some other type t is indifferent between sending the message that allows t to be protected and actually disclosing the issue that he knows is working fine: he is guaranteed to get the fair offer under both messages. Therefore, it is clear that under any parametric conditions that allow for active protection, there are also other equilibria where active protection is absent.

It is, thus, important to ask which of these equilibria are "more likely" to occur—especially since the aforementioned indifference of protectors between protecting and not protecting may make active protection equilibria appear "fragile." On the contrary, if information disclosure is costly, i.e., if sending some message $m \neq m_{00}$ comes at a cost $c > 0$, Customers would prefer to not disclose information rather than to do so.[9][10] In particular, if

[9]Hagenbach and Koessler (2017) break ties between disclosing and not disclosing in favor of not disclosing.

[10]We would expect this to be the case in our mechanic example where in order to demonstrate that some part of the car is functioning or not, the driver would need to open the car's hood and show, or even dismantle,

the conditions for full protection (see Corollary 1) are satisfied and the cost c is high enough, then the only equilibrium that survives is the full protection one. This is summarized in the following Observation.

Observation 1. *Let $G = (p, h, \pi, \varrho)$ be an expertise signaling game that satisfies the assumptions of Corollary 1. If signaling costs are high enough, then G has a unique equilibrium in which the Expert never cheats.*

In a similar fashion, if revealing more is more costly than revealing less (i.e., if it costs more to send m_{11} than to send m_{01} or m_{10}) protection under messages m_{01} and m_{10} can be shown to be robust. As a general result, separating equilibria ("reveal what you know") are less plausible compared to protection equilibria. Moreover, at least some partial pooling should be expected, especially in the presence of (even minimal) disclosure costs.

Proposition 5. *Let $G = (p, h, \pi, \varrho)$ be an expertise signaling game and let π and ϱ have full support. If G has a separating equilibrium, then G also has a pooling equilibrium where more informed Customers t_{01} (t_{10}) withhold information and pool with less informed Customers t_{00} under problem w_{01} (w_{10}).*

Proposition 5 shows that under any parametric condition where a separating equilibrium exists, pooling equilibria are also existent. Therefore, in the presence of costs, fully separating equilibria would be unlikely to be observed in the real world.

5 Discussion and Concluding Remark

This paper considered verifiable pre-play communication in a market with heterogeneously informed buyers and perfectly informed sellers. Better informed buyer types could distinguish themselves from worse informed buyer types by sending verifiable messages (but not vice versa). We show that partially informed types have an incentive to withhold information. In pooling equilibria (where some types do not reveal their expertise), protection from seller's opportunistic behaviour is achieved. This result is in contrast with some of the existing literature on persuasion, which highlights the incentives of senders to "masquerade" as higher types whenever there is an acyclical relationship between types (Hagenbach, Koessler, and Perez-Richet 2014; Milgrom 1981).

Agenda for future research Extending the model to an n-issue case is an obvious direction for future work. Fully solving the larger model is combinatorially-heavy as both the

the relevant part.

type- and state-spaces explode. Our intuition suggests that the main results of the paper should still hold in the more general case, yet new insights are likely to be gained.

Observation 1 and the analysis surrounding it shows that robustness of protection equilibria is not a concerning issue. The question is not so much on whether protection can be sustained but more about how to get to a protection equilibrium once a different equilibrium (with positive cheating) has been established. This requires a modeling of some dynamics and strategy-updating.

Moreover, can human decision-makers reach (active or passive) protection equilibria? Which methods (if any) can facilitate them to achieve full protection? These questions can be answered by experiments conducted in the lab or in the field (the latter being considerably more challenging from a design perspective than the former).

Finally, from a social planner's point of view, is it more beneficial to train partially informed customers to become experts or uninformed ones to become partially informed? The answer can vary from industry to industry and depend on whether training costs are super- or sub-additive. Answers to such questions can help inform policy on the provision of training programmes to fight fraud depending on particular market conditions.

Appendix

A Best responses

We solve the model "backwards," starting from the Customer's decision of whether to accept or reject an offer.

Accept/Reject Since $h_{10}, h_{01} > 0$, and given the assumptions of Section 2.4, the Customer should always accept an offer of o_{10} or o_{01}. The decision that needs more thorough calculation is on whether to accept an offer of o_{11}. Suppressing the history (I, m, o_{11}), let the Customer's belief be denoted by π''. The Customer should accept to receive treatment o_{11} only if his expected payment from rejecting o_{11} is at least as high as p_{11}, i.e., only if

$$p_{11} \leq \pi''_{10}(p_{10}+h_{10}) + \pi''_{01}(p_{01}+h_{01}) + \pi''_{11}(p_{11}+h_{11})$$

or, equivalently, only if (see Assumption 1)

$$\pi''_{11}h_{11} \geq \pi''_{10}A_{10} + \pi''_{01}A_{01}. \tag{7}$$

The Expert's strategy enters implicit in the above equation as it determines π''.

Cheat or not We proceed with identifying the optimal treatment that the Expert should offer. If the problem is w_{11}, then the Expert can only offer o_{11}. If the problem is w_i, $i \in \{10, 01\}$, the Expert has to choose whether to offer o_{11} (cheat) or offer to fix the actual problem o_i. As she is trying to maximize the payment she will receive, the Expert should offer o_{11} only if her expected payment from offering o_{11} is at least as large as p_i, i.e., only if

$$\sum_{t \in T} \varrho'(t|w_i, m)\alpha(I(w_i, t), m)p_{11} \geq p_i. \tag{8}$$

The Customer will always accept an offer of o_i.

Signaling Finally, we consider the Customer's decision on which pieces information to demonstrate that he has, if any. Let $P(I, m)$ denote the expected payment of the Customer of type t who finds himself at information set $I = I(w, t)$ and sends message m. The formula for P is given by

$$\begin{aligned} P(I,m) = {} & \pi'_{11}(I)\left(\alpha(I,m)p_{11} + (1-\alpha(I,m))(p_{11}+h_{11})\right) \\ & + \pi'_{10}(I)\left(\lambda(w_{10},m)(\alpha(I,m)p_{11} + (1-\alpha(I,m))(p_{10}+h_{10})) + (1-\lambda(w_{10},m))p_{10}\right) \\ & + \pi'_{01}(I)\left(\lambda(w_{01},m)(\alpha(I,m)p_{11} + (1-\alpha(I,m))(p_{01}+h_{01})) + (1-\lambda(w_{01},m))p_{01}\right). \end{aligned}$$

Or, more succintly, by

$$P(I,m) = \pi'_{11}(I)\left[(p_{11} + (1-\alpha(I,m))h_{11}\right] + \sum_{i \in \{10,01\}} \pi'_i(I)\left[p_i + \lambda(w_i,m)(h_i + \alpha(I,m)A_i\right]. \tag{9}$$

The Customer should send message $m \in M(t)$ only if

$$P(I,m) \leq P(I,m') \qquad \text{all } m' \in M(t). \tag{10}$$

References

Cho, I.-K. and D. M. Kreps (1987). “Signaling games and stable equilibria”. *The Quarterly Journal of Economics* 102 (2), pp. 179–221.

Dulleck, U. and R. Kerschbamer (2006). “On doctors, mechanics, and computer specialists: The economics of credence goods”. *Journal of Economic Literature* 44 (1), pp. 5–42.

Emons, W. (1997). “Credence goods and fraudulent experts”. *The RAND Journal of Economics*, pp. 107–119.

Glazer, J. and A. Rubinstein (2004). “On optimal rules of persuasion”. *Econometrica* 72 (6), pp. 1715–1736.

Hagenbach, J. and F. Koessler (2017). “Simple versus rich language in disclosure games”. *Review of Economic Design* 21 (3), pp. 163–175.

Hagenbach, J., F. Koessler, and E. Perez-Richet (2014). “Certifiable Pre-Play Communication: Full Disclosure”. *Econometrica* 82 (3), pp. 1093–1131.

Hart, S., I. Kremer, and M. Perry (2017). “Evidence games: truth and commitment”. *The American Economic Review* 107 (3), pp. 690–713.

Hyndman, K. and S. Ozerturk (2011). “Consumer information in a market for expert services”. *Journal of Economic Behavior & Organization* 80 (3), pp. 628–640.

Milgrom, P. R. (1981). “Good news and bad news: Representation theorems and applications”. *The Bell Journal of Economics*, pp. 380–391.

Pitchik, C. and A. Schotter (1987). “Honesty in a Model of Strategic Information Transmission”. *The American Economic Review* 77 (5), pp. 1032–1036.

Rubinstein, A. and J. Glazer (2006). “A study in the pragmatics of persuasion: a game theoretical approach”. *Theoretical Economics* 1 (4), pp. 395–410.

Schneider, T., L. Meub, and K. Bizer (2016). *Consumer information in a market for expert services: Experimental evidence*. Tech. rep. University of Goettingen, Department of Economics.

Wolinsky, A. (1993). “Competition in a market for informed experts’ services”. *The RAND Journal of Economics*, pp. 380–398.