Job Market Paper

Alexandros Rigos

Expertise Disclosure in Markets for Credence Goods*

Alexandros Rigos† Maria Kozlovskaya‡ Matteo Foschi§

November 19, 2019

(For the latest version please click here)

Abstract

In markets for credence goods buyers rely on expert sellers for diagnosis and treatment. This gives sellers incentives to provide buyers with unnecessary services. We theoretically study strategic information transmission in this environment, allowing differentially informed buyers to verifiably communicate their expertise to the sellers. We show that, in equilibrium, it is frequently not optimal for buyers to disclose their level of expertise, and that this—when sellers are unable to distinguish "feigned ignorance" from a genuine lack of expertise—may completely eliminate seller fraud in pooling equilibria. Our results can be used to identify what pieces of information should be provided in public information campaigns that aim to reduce consumer exploitation. They also suggest that such campaigns can be effective even if some buyers do not fully understand their content (because they affect sellers' beliefs).

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

JEL classification: C72, D82, D83, L15

---

*We thank Chris Wallace and Martin Kaae Jensen for the detailed comments on a previous draft, Heinrich Nax, Petra Thiemann, Pol Campos-Mercade, and Willemien Kets for useful discussions and suggestions, and audiences at the 2017 Lancaster Game Theory Conference, the 2018 Stony Brook Game Theory Conference, the 2019 CEPET Workshop, SING15, SAET 2019, ESEM 2019, and seminar participants at Aston, ETH Zürich, Lund, Manchester, and Cardiff 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

Credence goods markets are characterized by extreme information asymmetry as buyers typically lack knowledge about their own needs and rely on expert sellers for diagnosis and treatment. This information asymmetry gives sellers an incentive to abuse their superior knowledge and exploit consumers by providing them with unnecessary services, i.e., by overtreating them. These markets are ubiquitous and include many markets for services, such as medical care, car repair, and financial services (Dulleck and Kerschbamer 2006; Dulleck, Kerschbamer, and Sutter 2011).[1]

In credence goods markets, buyers are often heterogeneous in their expertise. For example, certain consumers may have more (or different) knowledge about their health, car's condition, or financial needs than others. In an attempt to avoid being overtreated, such consumers may want to reveal some—or all—of their expertise to the seller.

The goal of this paper is to study buyers' optimal information disclosure in credence goods markets. We particularly look at the case where buyers are heterogenous with respect to the information about the services they need, i.e., they have different levels of "expertise." Our question is twofold. First, we describe optimal behavior from the viewpoint of each buyer type. We distinguish between fully informed, partially informed, and completely uninformed buyers. Second, we study how these optimal behaviors influence the reaction of the sellers. In particular, we study under which conditions optimal information disclosure on the part of informed or partly informed buyers can benefit the uninformed buyers.

We get two main results. Firstly, in equilibrium, less-than-fully expert buyers can be strictly better off withholding (rather than disclosing) some of their evidence from the seller. Thus, they prefer to appear less informed than they are (to "feign ignorance"). This result implies that revealing that one has some expertise can be detrimental. Indeed, when a buyer with less-than-full expertise discloses "bad" news, he communicates to the seller that he is more willing to accept unnecessary treatment than a buyer with no expertise whatsoever.[2] As a response, the seller is more likely to overtreat a buyer who has disclosed such evidence than one who has not. Anticipating this, the buyer prefers to feign lack of expertise.

[1] Overtreatment plagues both private and public medical care systems: \$200bn is wasted annually on unnecessary medical treatment in the US (Johnson and Rehavi 2016; Lyu et al. 2017), while in the UK the NHS is losing £687 million a year on unnecessary prescriptions and X-rays alone (Academy of Medical Royal Colleges 2014; Malhotra et al. 2015). The extent of buyer exploitation is even more dramatic in financial services: British consumers were repaid £36.4bn between 2011-2019 for unnecessary add-on insurance (FSCS 2019). Finally, unnecessary automobile repairs cost British motorists £2.3bn a year (Green Flag Limited 2018), while two-thirds of American drivers do not trust car mechanics, citing overtreatment as the main reason for their lack of confidence (The American Automobile Association 2016).

[2] A buyer's information can be classified as "good" or "bad" by comparing his interim to his ex-ante expected expenditure for receiving just necessary treatment.

Secondly, buyer populations with sufficient expertise can become fully protected from seller exploitation through the use of pooling strategies (essentially concealing their level of expertise from sellers). Importantly, this full protection does not require the presence of fully-expert consumers. A straightforward benefit of providing a consumer with expertise is that he becomes able to detect and avoid some types of overtreatment. Our second result shows that in addition to that, there is a positive externality that benefits other, non-expert consumers: when the seller does not know the buyer's expertise level, she is reluctant to offer unnecessary treatments as it is likely that her attempt to overtreat will be detected. So, consumer exploitation is eliminated without requiring that all consumers have expertise.

Our main contribution to the literature is to provide the first formal model of buyer-seller communication in a credence goods environment. Such markets have been analyzed in the past (see Dulleck and Kerschbamer (2006) for an organized literature survey). Despite that, customers' strategic decision to disclose or not their expertise—and if so, to what extent—has not been studied before, to the best of our knowledge.[3] Our analysis gives rise to new intuitions and proposes a new channel through which seller exploitation can be prevented.

The model, in short, is the following. A Customer (he) requires the services of an Expert (she). There are two distinct kinds of services that the Customer may need: service 1 and service 2. He requires either or both. The Customer has one of four types: the "clueless" type has no information about his needs; partially knowledgeable (or partially informed) type 1 can observe whether or not he needs service 1, but receives no information about service 2; partially knowledgeable type 2 gets informed about service 2 but not service 1; the knowledgeable type is perfectly informed. When approached by the Customer, the Expert perfectly observes the Customer's needs and makes a take-it-or-leave-it offer to (verifiably) provide either or both services. Importantly, before receiving the Expert's offer, we allow the Customer to selectively disclose verifiable evidence about his information. The Expert is liable and so she has to at least provide the service that the Customer needs. Therefore, the Expert can attempt to offer unnecessary services. If the Expert's offer is rejected, the Customer has his problem treated by an "honest expert" (outside option) who always offers just the necessary services (never cheats) but at a higher price than the Expert.[4]

The standard credence goods model (Pitchik and Schotter 1987; Wolinsky 1993) assumes that there are two possible problems from which the Customer may be suffering: a minor one (which is cheap) and a serious one (which is more expensive). In contrast, Cus-

[3] The only papers in which some communciation occurs between buyers and sellers are Fong (2005) where consumers can choose to disclose or not that they suffer from a symptom and Hyndman and Ozerturk (2011) where some customer types are more likely to suffer from a severe problem and may, thus, want to hide that. We discuss these papers in Section 5.

[4] In the repairs and health care applications, the "honest expert" could be the official car dealership and a second medical opinion, respectively.

tomers in our model may or may not be in need of each of different kinds of services, rather than being in need of either a high or a low quantity of one service. Our multiple-services structure naturally generates a partial order of Customer types based on their expertise (which services they are aware whether or not they need). This heterogeneity provides content to Customers' disclosure messages and allows us to study communication.

We also contribute to the literature on the disclosure of hard evidence which typically uses message structures similar to ours (e.g. Hart, Kremer, and Perry (2017) who take a mechanism design approach or Glazer and Rubinstein (2004, 2006) who study persuasion). Unlike these models, ours has two-sided uncertainty: the Expert is uncertain about the Customer's type and at the same time some Customer types are uncertain about the problem from which they suffer. In this sense, we have a case of two-sided persuasion: the Customer tries to persuade the Expert not to overtreat him, while the Expert tries to persuade the Customer to accept her offer. Moreover, our paper provides one more real-world application of hard evidence disclosure (like, for example, Eső and Wallace 2018).

Ely and Välimäki (2003) also study situations where sellers can exploit consumers and focus on how reputation can help—or, in fact, be detrimental to—efficiency. As in our model, they also consider an Expert who treats Customers with treatments being either appropriate or not for the Customer's problem at hand. Exploitation in Ely and Välimäki (2003) is the act of a "bad" behavioral Expert type to offer inappropriate treatment to the customer. By contrast, in our model decisions to overtreat result from the Expert's revenue maximization. The Expert does not have a "type," she is just trying to make as much money as she possibly can. Real-world Experts are likely to have reputational concerns while at the same time facing Customers with heterogeneous expertise. We, therefore, argue that our approach is complementary to Ely and Välimäki (2003)'s in addressing real-world markets.

The particular problem structure we use is natural for several credence goods markets like car repairs and healthcare, as consumers are likely to vary in their level of expertise in such issues and treatment is usually provided in discrete units (e.g. fixing of particular parts or prescription of particular medication or tests). This structure, coupled with consumers' ability to actively disclose or withhold information, gives rise to our first novel result: the possibility of pooling "active protection" equilibria where some (or all) non-knowledgeable consumer types are protected from overtreatment.

Consequently, providing some consumers with expertise can have positive spillovers to the rest of the consumer population. In fact, we highlight that consumer populations with sufficient expertise can achieve equilibria in which they are fully protected from overtreatment, even without having fully-expert (knowledgeable) consumers among them. Our results broaden the spectrum of policy instruments aimed at improving consumer welfare on credence goods markets. They show that careful targeting of public information campaigns can reduce consumer exploitation and point out which pieces of expertise should be targeted. Interestingly, our model suggests that such campaigns can be effective even if some buyers do not fully understand their content (through affecting sellers' beliefs).

When motorists visit mechanics, they may want to attempt to show expertise by talking about parts they know are faulty (more than just describing symptoms). Our model provides an explanation for such behavior. It is plausible that in car repairs, Customer knowledge about one part is usually highly correlated with knowledge about other parts of the car. The strategy "only talk about broken parts" allows individuals with less knowledge (those who observe the particular broken parts) to pool with others with more knowledge (those who also observe some other non-faulty parts). Since the latter are relatively more than the former (because of the high correlation of expertise across parts), mechanics avoid overtreating them. In the language of Section 4.3.1, this is a partial active protection equilibrium, as several but not all types are protected from overtreatment (in particular, motorists without any expertise are not protected).

Additionally, we show 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. This non-disclosure result echoes the analyses in Glazer and Rubinstein (2006), Sher (2011) and Dye (1985). It is also consistent with recent experimental findings, as we will argue in the next paragraph.

Kerschbamer, Neururer, and Sutter (2019) conducted field experiments to investigate the effect of self-diagnosis websites and review aggregators on the price charged for a credence good. They find that disclosing a self-diagnosis does not protect consumers from seller exploitation in the computer repair industry. In particular, when customers disclosed self-diagnoses stating that most likely their problem was caused by an incorrect part (and, thus, revealed that their knowledge about the problem is imprecise) they got charged significantly more than their peers who did not provide any assessment (or whose self-diagnosis identified the correct problem). Such seller behavior is consistent with the passive protection equilibrium of Section 4.3.2 where the Expert does not overtreat the clueless type but overtreats partially knowledgeable Customers when she identifies them. These results are in line with our theoretically established incentive to hide information when it is imprecise.

The rest of the paper is organized 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. Section 5 discusses some closely related literature and Section 6 concludes with some paths for future research.

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, lowercase Greek characters are 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.[5] 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.

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 type and is always able to identify the exact problem he suffers from. Types t_{10} and t_{01} are partially knowledgeable (or 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.

2.3 Information sets

According to the above typology, given his type and the state of the world, the Customer can find himself at 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

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

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}	$m_{11}, m_{01}, m_{10}, m_{00}$

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 at which they can find themselves is given in Table 2. We also use the notation $I(w, t)$ to denote the information set at 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 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.

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.*[6]

[6]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.

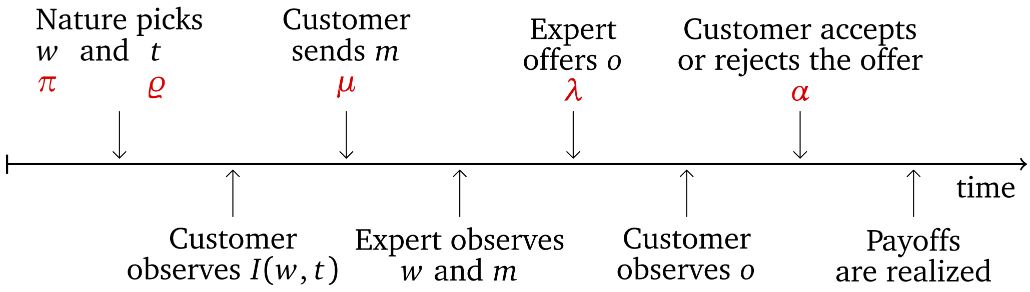


Figure 1: Timing of events and decisions. Variables in Greek characters indicate the distributions that the respective variables follow.

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 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.[7] 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

[7]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.

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.

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

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

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. 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 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 (11))

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

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.,

$$\lambda(w_{10}, m_{01}) = \lambda(w_{01}, m_{10}) = 0. \tag{5}$$

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 in Appendix B. □

The intuition behind the above result is the following. In a state of the world where only one of the two issues needs 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 knowledgeable 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 knowledgeable 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 knowledgeable type to pool with the clueless type to decrease the probability of being exploited.

The result of Proposition 1 is stark: in a world without knowledgeable Customers (which is most likely the case in the real world), partially knowledgeable 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 exploitation.

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 decrease 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 knowledgeable 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.[8] The next proposition provides a necessary and sufficient condition for the existence of such equilibria.

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

$$\pi_{11}h_{11} \geq \pi_{10}A_{10} + \pi_{01}A_{01}. \tag{6}$$

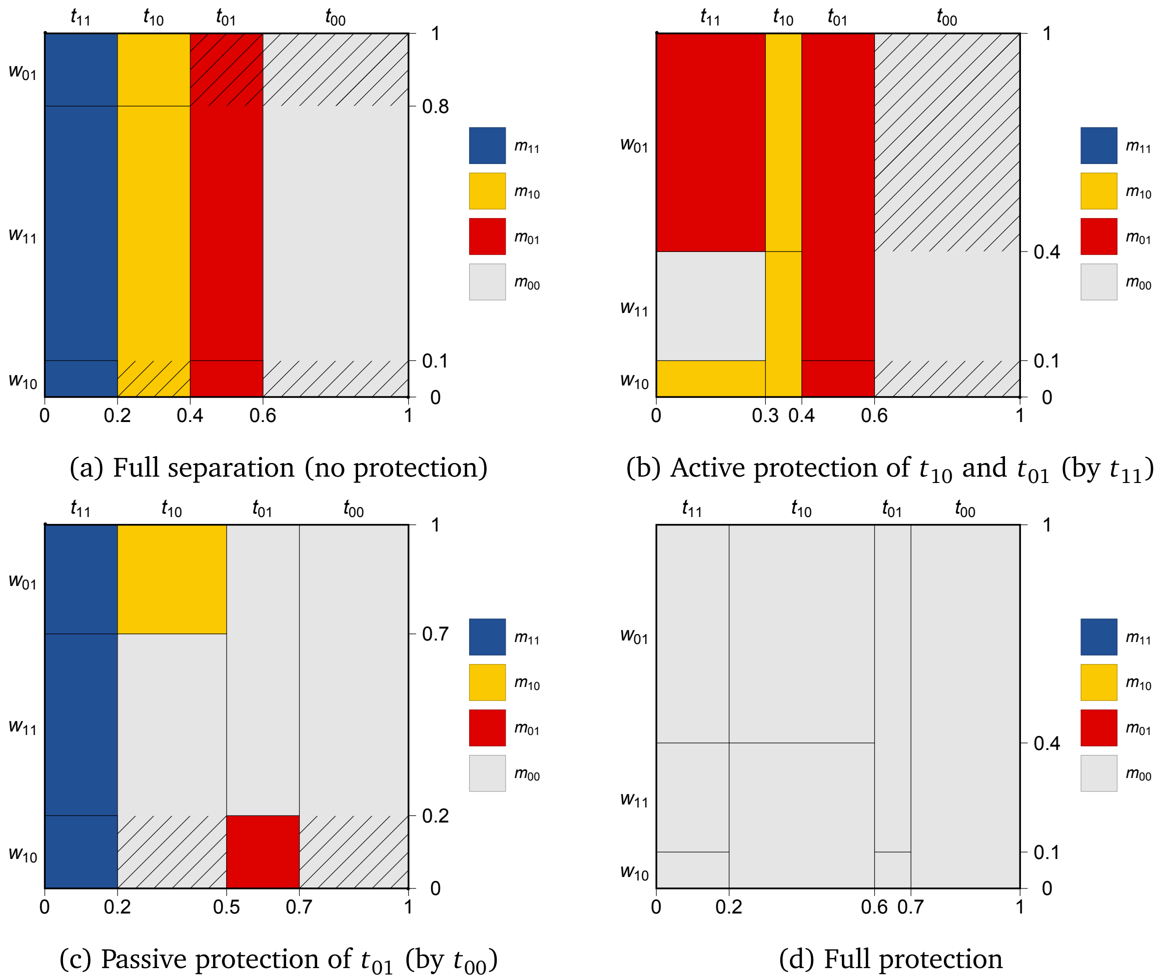
So, separating equilibria exist only if w_{11} occurs too often or h_{11} is too high. In particular, condition (6) says that the ex-ante expected benefit from accepting treatment o_{11} when it is truthful (and, therefore, avoiding paying the extra h_{11}) outweighs the expected loss of accepting a 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

[8] 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 $i \neq j$ $(i, j \in \{01, 10, 11\}, w \in \Omega)$.



(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: $p_{11} = 1$, $p_{10} = 0.5$, $p_{01} = 0.8$, $h_{11} = h_{10} = h_{01} = 0.1$.

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.[9] Protection happens by t' hiding some information that he has, essentially not disclosing the issue that he knows does not need fixing. 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).[10]

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

$$\frac{\varrho(t_{1x})}{\varrho(t_{1x}) + \varrho(t_{0x})} \geq \frac{p_{11} - p_{10}}{p_{11}}, \tag{7}$$

[9] 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).

[10] 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}.

where $x = *$ if both types in t_{0*} are protected and $x = 1$ if only the partially knowledgeable 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}).[11] 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*

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

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 protection equilibrium is one where all types send message m_{00} in all states of the world (see Figure 2d).[12]

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-knowledgeable 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.

[11] 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 (7) also holds.

[12] 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.

Proposition 4. *Let $G = (p, h, \pi, \varrho)$ be an expertise signaling game and let π and ϱ have full support. Let also $\frac{\varrho(t_{01})}{\varrho(t_{00})+\varrho(t_{01})} \leq \frac{p_{01}}{p_{11}}$ and $\pi_{11}h_{11} \leq \pi_{10}A_{10}$. 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) $\frac{\varrho(t_{10})}{\varrho(t_{00})+\varrho(t_{10})} \geq \frac{p_{10}}{p_{11}}$.

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

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 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.[13][14] In particular, if 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 outcome in which the Expert never cheats on the equilibrium path.*

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.

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

[14] 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 relevant part.

5 Discussion

Two studies that relate closely to ours are Fong (2005) and Hyndman and Ozerturk (2011). Both studies document cases similar to our result that some consumers with partial information are better off not disclosing this information to the expert. In all three studies (the two mentioned earlier and the present one) this happens because such consumers are more willing to accept a fraudulent offer than uninformed consumers and would, therefore, not let the expert know that.

Hyndman and Ozerturk (2011) enrich the standard two-state model of Pitchik and Schotter (1987) by allowing for some consumers to receive a noisy signal about the problem from which they suffer. Based on the signal's realization, the are "high-" and "low-risk" informed consumers. They find that "high-risk" consumers are more likely to be cheated by the expert than uninformed ones and conclude that such consumers should, therefore, "seek out ways to hide this [...] or to otherwise mimic the type who is less likely to be cheated" (Hyndman and Ozerturk 2011, p.635).

The mechanism that generates this result in both Hyndman and Ozerturk (2011) and this paper is that consumers who assign high probability to be suffering from the serious problem (require both services in our case) are more willing to accept a serious treatment—even if fraudulent—than uninformed consumers. Therefore, they are more likely to be cheated by the expert and, as a result, they would rather withhold their information from the expert. The difference is in the way in which consumers arrive at these beliefs. In Hyndman and Ozerturk (2011), different types arrive at different posteriors through noisy observation of the problem. In our setting, they come naturally from a set of expertise-bearing types who have accurate information. Importantly, we show that in our multiple-issue setting with expertise types, partially knowledgeable consumers who hide their knowledge can even get fully protected from fraud.

Fong (2005) studies how consumer heterogeneity can affect equilibrium outcomes in markets for credence goods. He considers situations in which there are two serious problems which the expert can only imperfectly diagnose whereas she can perfectly diagnose the problem if it is minor. Simultaneously, he considers consumer types who possess a "signal" that can help the expert accurately diagnose a serious problem. He finds that there are cases where the consumers may not want to share this signal as this may make them more susceptible to overtreatment.

The mechanism that drives this result is the following. If a consumer who has disclosed the signal is, indeed, suffering from one of the two serious problems, the expert will always offer him the appropriate serious treatment. Therefore, the expected value of a serious treatment to such a consumer is increased, which leads to him being more willing to accept a major treatment compared to a consumer without the signal. Notice that in Fong (2005)'s setting, the consumer cannot make use of his signal in order to accept or reject the offer that he receives. The signal he possesses does not inform him but only the expert.

In contrast to the aforementioned studies, we view consumers as active market participants that are able to *selectively* disclose their knowledge, rather than the expert being immediately able to identify them. Moreover, in our setting information is actually useful to the individual consumer who is now able to identify and reject some of the Expert's fraudulent offers. This makes us uniquely able to examine the resulting implications of consumers' strategic choice of information disclosure.

6 Concluding Remark

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 type- and state-space sizes 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.

As we showed, protection equilibria survive the introduction of signaling costs (Observation 1 and the analysis surrounding it). So, robustness is not a concerning issue. The question is not so much whether protection can be sustained but more about how to get to a protection equilibrium once a different equilibrium (where overtreatment occurs) has been established. 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 cost-effective to train partially informed customers to become knowledgeable or uninformed ones to become partially knowledgeable? The answer can vary from industry to industry and depend on whether training costs are super- or sub-additive. Our model can be used to inform policy on the provision of training programmes to fight exploitation by sellers, and to adjust them to 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)

$$\pi''_{11} h_{11} \geq \pi''_{10} A_{10} + \pi''_{01} A_{01}. \tag{8}$$

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

$$\sum_{t \in T} \varrho'(t|w_i, m)\alpha(I(w_i, t), m)p_{11} \geq p_i. \tag{9}$$

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

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

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

$$P(I,m) \leq P(I,m') \qquad \text{all } m' \in M(t). \tag{11}$$

B Omitted Proofs

B.1 Proof of Proposition 1

Consider a population with $\varrho(t_{11}) = 0$ and the subgame that begins with the problem w_{10}.

From his indifference condition, a Customer at info set I_{1*} who has sent message m accepts an offer of o with positive probability only if

$$\lambda(w_{10}, m) \leq \frac{\pi_{11}h_{11}}{\pi_{10}A_{10}}.$$

So, the Expert's best response to a Customer t_{10} who has revealed his type (has sent m_{10}) is to cheat with probability

$$\lambda^*(w_{10}, m_{10}) = \min\left\{\frac{\pi_{11}h_{11}}{\pi_{10}A_{10}}, 1\right\}, \tag{12}$$

which is always positive.

There are three potential cases for the Expert's equilibrium cheating probability $\lambda(w_{10}, m_{00})$:

1. If $\lambda(w_{10}, m_{00}) = \lambda^*(w_{10}, m_{10})$, then the Customer at I_{1*} is indifferent between disclosing his type and not. Therefore the statement holds with equality.

2. If $\lambda(w_{10}, m_{00}) < \lambda^*(w_{10}, m_{10})$, then the Customer at I_{1*} strictly prefers not disclosing his type, as he is being cheated less under m_{00} than under m_{10}. Therefore the statement holds with strong inequality.

3. We will now show that it cannot be that $\lambda(w_{10}, m_{00}) > \lambda^*(w_{10}, m_{10})$.

   By contradiction, assume that $\lambda(w_{10}, m_{00}) > \lambda^*(w_{10}, m_{10})$ in some PBE. Then, the Customer at information set I_{1*} is strictly worse-off under m_{00} than under m_{10}. Therefore, in equilibrium, he is sending m_{10} for sure ($\mu(m_{10}|I_{1*}) = 1$). Similarly, the Customer at information set I_{*0} will be sending m_{01} for sure in such an equilibrium, since this would guarantee him the fair offer of o_{10} (see equation 5). So, the only remaining type that could be sending m_{00} in this PBE is t_{00}.

Since $\lambda(w_{10}, m_{00}) > 0$, the Expert is cheating with positive probability, so she is at least indifferent between cheating and not. This means that when she is (fraudulently) offering o_{11}, her offer is accepted with positive probability. So, type t_{00} (the only remaining type) is at least indifferent between accepting and rejecting w_{11}. From type t_{00}'s indifference condition we get that he is accepting with positive probability only if

$$\pi_{11}h_{11} \geq \pi_{10}\lambda(w_{10}, m_{00})A_{10} + \pi_{01}\lambda(w_{01}, m_{00})A_{01}$$

which means that

$$\lambda(w_{10}, m_{00}) \leq \frac{\pi_{11}h_{11} - \pi_{01}\lambda(w_{01}, m_{00})A_{01}}{\pi_{10}A_{10}} \leq \lambda^*(w_{10}, m_{10}),$$

a contradiction. A symmetric case holds the Customer at information set I_{*1}. □

B.2 Proof of Proposition 2

"⇒" direction:

The only message profile that leads to a fully separating PBE is the one where t_i sends m_i for $i \in \{11, 10, 01, 00\}$. Given the discussion of sections 3.1 and 3.2, and the fact that clueless Customers can only send m_{00}, in order to establish a fully separating equilibrium all that is required is to make sure that the incentive of Customers at information sets I_{1*} and I_{*1} to send m_{00} (see Proposition 1) is weak.

So, it must be that (see equation 12)

$$\begin{cases} \lambda(w_{10}, m_{00}) = \lambda^*(w_{10}, m_{10}) = \min\left\{\frac{\pi_{11}h_{11}}{\pi_{10}A_{10}}, 1\right\} & \text{and} \\ \lambda(w_{01}, m_{00}) = \lambda^*(w_{01}, m_{01}) = \min\left\{\frac{\pi_{11}h_{11}}{\pi_{01}A_{01}}, 1\right\}. \end{cases} \tag{13}$$

In a separating equilibrium, upon receiving message m_{00}, the Expert will cheat with positive probability under both problems w_{10} and w_{01}. Otherwise, condition (13) would be violated. So, under either w_{10} or w_{01}, when the Expert makes a fraudulent offer, she should have it accepted with positive probability by t_{00}, i.e.,

$$\pi_{11}h_{11} \geq \pi_{10}\lambda(w_{10}, m_{00})A_{10} + \pi_{01}\lambda(w_{01}, m_{00})A_{01}. \tag{14}$$

Say $\pi_{01}A_{01} > \pi_{11}h_{11}$. Then the above equation, together with (13) reads

$$\pi_{11}h_{11} \geq \pi_{11}h_{11} + \pi_{01}\lambda(w_{01}, m_{00})A_{01} \Rightarrow \lambda(w_{01}, m_{00}) \leq 0$$

which contradicts (13). Similarly for $\pi_{01}A_{01} > \pi_{11}h_{11}$. So, it has to be that

$$\pi_{11}h_{11} \geq \max\{\pi_{10}A_{10}, \pi_{01}A_{01}\}.$$

Therefore, $\lambda(w_{10}, m_{00}) = \lambda(w_{01}, m_{00}) = 1$ and, because of (14), we get

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

"⇐" direction:

Assume that $\pi_{11} h_{11} \geq \pi_{10} A_{10} + \pi_{01} A_{01}$. We will construct a fully separating equilibrium. Assume that the Expert uses strategy:

$$\begin{aligned} \lambda(w_{01}, m_{00}) = \lambda(w_{10}, m_{00}) = \lambda(w_{01}, m_{01}) = \lambda(w_{10}, m_{10}) = 1 \\ \lambda(w_{01}, m_{11}) = \lambda(w_{10}, m_{11}) = \lambda(w_{01}, m_{10}) = \lambda(w_{10}, m_{01}) = 0. \end{aligned} \tag{15}$$

The clueless Customer prefers accepting to rejecting o_{11} (at least weakly) for any possible strategy of the Expert (equation (14)). Let him accept for sure $\alpha(I_{**}, m_{00}) = 1$. Moreover, the Customer at information set I_{1*} (I_{*1}) also accepts o_{11} for sure after sending any of the two messages available to him, as

$$\begin{aligned} \lambda(w_{10}, m_{10}) = \lambda(w_{10}, m_{00}) = 1 < \frac{\pi_{11} h_{11}}{\pi_{10} A_{10}} \\ (\lambda(w_{01}, m_{01}) = \lambda(w_{01}, m_{00}) = 1 < \frac{\pi_{11} h_{11}}{\pi_{01} A_{01}}). \end{aligned}$$

Given that $\alpha(I_{**}, m_{00}) = \alpha(I_{1*}, m_{00}) = \alpha(I_{*1}, m_{00}) = \alpha(I_{1*}, m_{10}) = \alpha(I_{*1}, m_{01}) = 1$, it is a best response for the Expert to use (15). Finally, as Customers at each of the information sets I_{1*} and I_{*1} face the same cheating probability independently of what message they send, they are indifferent between the two messages available to them. Therefore, it is a best response for the Customer at I_{1*} to send m_{10} and for the Customer at I_{*1} to send m_{01}. □

B.3 Proof of Proposition 3

"⇒" direction:

Consider the subgame that begins with the problem w_{01}. In an active protection equilibrium, the Expert is never cheating after receiving a message m that can be sent by both types in t_{1x} and in t_{0x} ($m = m_{10}$ if only t_{01} is protected, whereas $m = m_{00}$ if both types in t_{0*} are protected).

Since the Expert is not cheating after receiving message m, when best-responding all Customers with types in t_{0x} will accept an offer of o_{11} after having sent m (even a fraudulent one off the equilibrium path). Moreover, all Customers in t_{1x} will send m for sure, as they are being protected.

On the other hand, Customers in t_{1x} will only accept offers which they know to be honest, i.e., w_{01} but reject w_{11}. Moreover, they will send m only if they are sure that the Expert will not cheat under m. Let μ be the proportion of Customers in t_{1x} that send message m.

Given the above, it must be a best response for the Expert to play $\lambda(w_{01}, m) = 0$. So, from the Expert's best response:

$$\varrho'(t_{0x}|w_{01}, m)p_{11} \leq p_{01} \quad \Rightarrow \quad \frac{\varrho(t_{0x})}{\mu\varrho(t_{1x}) + \varrho(t_{0x})} \leq \frac{p_{01}}{p_{11}} \quad \Rightarrow \quad \mu \geq \frac{\varrho(t_{0x})}{\varrho(t_{1x})}\left(\frac{p_{11}}{p_{01}} - 1\right)$$

and, since $\mu \leq 1$, we get that

$$\frac{\varrho(t_{1x})}{\varrho(t_{1x}) + \varrho(t_{0x})} \geq \frac{p_{11} - p_{01}}{p_{11}}. \tag{16}$$

"⇐" direction:

Assume that $\frac{\varrho(t_{1x})}{\varrho(t_{1x})+\varrho(t_{0x})} \geq \frac{p_{11}-p_{01}}{p_{11}}$ and consider the signaling profile in which each Customer type t in $t_{1x} \cup t_{0x}$ sends message m in his respective information set $I(w_{01}, t)$. As seen above, the Expert's best response to that is to play $\lambda(w_{01}, m) = 0$, which establishes protection under w_{01}.

Note that it is important that when $x = 0$, no type other than t_{11} and t_{10} can send $m = m_{10}$. If they could, they would do so and that would make the Expert's "incentive compatibility" constraint (16) harder to satisfy. □

B.4 Proof of Proposition 4

In a passive protection equilibrium we want t_{01} to be protected under w_{01} by sending message m_{00} without any "help" from types t_{11} and t_{10} (who can identify problem w_{01}). Therefore, types t_{11} and t_{10} will not send m_{00} when the problem is w_{01}, as this would imply active protection (at least partially). Moreover we assume that type t_{11} sends a message $m \in \{m_{11}, m_{10}\}$ under w_{01} and a message $m' \in \{m_{11}, m_{01}\}$ under w_{10} as, if he sends m_{01} under w_{01} or m_{10} under w_{10}, he is actively protecting type t_{01} and/or t_{01}, respectively.

"⇒" direction (The conditions on the parameters are put in $\boxed{\text{boxes}}$ as derived):

In a passive protection equilibrium, the Expert does not cheat under w_{01} after receiving message m_{00}. From the Expert's best response, this means that

$$\frac{\varrho(t_{00})\alpha(I_{**}, m_{00}) + \varrho(t_{01})\alpha(I_{*1}, m_{00})}{\varrho(t_{00}) + \varrho(t_{01})} \leq \frac{p_{01}}{p_{11}} \tag{17}$$

which, since $\alpha(I_{*1}, m_{00}) = 1$ in t_{01}'s best response, leads to

$$\alpha(I_{**}, m_{00}) \leq \frac{p_{01}}{p_{11}}\left(1 + \frac{\varrho(t_{01})}{\varrho(t_{00})}\right) - \frac{\varrho(t_{01})}{\varrho(t_{00})}. \tag{18}$$

Since it must be that $\alpha(I_{**}, m_{00}) \geq 0$, we get

$$\boxed{\frac{\varrho(t_{01})}{\varrho(t_{01}) + \varrho(t_{00})} \leq \frac{p_{01}}{p_{11}}.} \tag{19}$$

Lemma 1. *Under w_{10} and after receiving message m_{00}, the Expert cheats with probability*

$$\lambda(w_{10}, m_{00}) = \boxed{\frac{\pi_{11}h_{11}}{\pi_{10}A_{10}} \in (0,1].} \tag{20}$$

Proof. Notice that since $\lambda(w_{01}, m_{00}) = 0$, t_{00} and t_{10} have the same threshold for accepting o_{11} (at information sets I_{**} and I_{1*}, respectively): $\lambda^* = \min\left\{\frac{\pi_{11}h_{11}}{\pi_{10}A_{10}}, 1\right\}$ (see (12)). Moreover, the Expert uses $\lambda(w_{10}, m_{10}) = \lambda^*$, as type t_{10} is not actively protected by type t_{11}. We examine all possible cases:

1. If $\lambda(w_{10}, m_{00}) \in [0, \lambda^*)$, then type t_{10} sends message m_{00} at info set I_{1*}, as he is cheated on with smaller probability than if he sends m_{10}. Both t_{00} and t_{10} always accept offer o_{11} under m_{00} ($\alpha(I_{**}, m_{00}) = \alpha(I_{1*}, m_{00}) = 1$). So, the Expert's best response to that is to cheat with probability $\lambda(w_{10}, m_{00}) = 1 > \lambda^*$, a contradiction.

2. If $\lambda(w_{10}, m_{00}) \in (\lambda^*, 1]$, then type t_{10} sends m_{10} in equilibrium and t_{00} always rejects an offer of o_{11}. Also, the Expert is indifferent between cheating and not cheating (she is playing a mixed strategy). So from her indifference condition, and since $\mu(m_{00}|I_{1*}) = 0$, we get that $\alpha(I_{**}, m_{00}) = \frac{p_{10}}{p_{11}} \Rightarrow \frac{p_{10}}{p_{11}} = 0$, which contradicts our assumption.

3. If $\pi_{11}h_{11} > \pi_{10}A_{10}$, then $\alpha(I_{**}, m_{00}) = 1$ for any strategy of the Expert and condition (17) yields $p_{11} \leq p_{01}$, which contradicts our assumption.

So, it has to be that $\pi_{11}h_{11} \leq \pi_{10}A_{10}$ and that $\lambda(w_{10}, m_{00}) = \lambda^* = \frac{\pi_{11}h_{11}}{\pi_{10}A_{10}}$. □

So, in this passive protection equilibrium the Expert is always going to be indifferent between cheating and not cheating under w_{10} and after receiving m_{00}. Furthermore, the type t_{10} Customer is indifferent between sending m_{10} and m_{00} under I_{1*}, so assume that he sends m_{00} with probability μ. From the Expert's indifference condition, we get:

$$\frac{\varrho(t_{00})\alpha(I_{**}, m_{00}) + \mu\varrho(t_{10})\alpha(I_{1*}, m_{00})}{\varrho(t_{00}) + \varrho(t_{10})} = \frac{p_{10}}{p_{11}} \tag{21}$$

and since $\alpha(I_{1*}, m_{00}) \in [0,1]$, we have that

$$\alpha(I_{**}, m_{00}) \in \left[\frac{p_{10}}{p_{11}}\left(1 + \frac{\varrho(t_{10})}{\varrho(t_{00})}\right) - \frac{\mu\varrho(t_{10})}{\varrho(t_{00})}, \frac{p_{10}}{p_{11}}\left(1 + \frac{\varrho(t_{10})}{\varrho(t_{00})}\right)\right] \tag{22}$$

Combining (22) together with (18) and the natural condition $\alpha(I_{**}, m_{00}) \in [0,1]$, we get:

$$\alpha(I_{**}, m_{00}) \in \left[\max\left\{\frac{p_{10}}{p_{11}}\left(1 + \frac{\varrho(t_{10})}{\varrho(t_{00})}\right) - \frac{\mu\varrho(t_{10})}{\varrho(t_{00})}, 0\right\}\right., \\ \left.\min\left\{\frac{p_{10}}{p_{11}}\left(1 + \frac{\varrho(t_{10})}{\varrho(t_{00})}\right), \frac{p_{01}}{p_{11}}\left(1 + \frac{\varrho(t_{01})}{\varrho(t_{00})}\right) - \frac{\varrho(t_{01})}{\varrho(t_{00})}, 1\right\}\right]$$

The only additional, non-trivial restriction that the above condition imposes is:

$$\frac{p_{10}}{p_{11}}\left(1+\frac{\varrho(t_{10})}{\varrho(t_{00})}\right)-\frac{\mu\varrho(t_{10})}{\varrho(t_{00})}\leq\frac{p_{01}}{p_{11}}\left(1+\frac{\varrho(t_{01})}{\varrho(t_{00})}\right)-\frac{\varrho(t_{01})}{\varrho(t_{00})}\Rightarrow$$
$$\mu\varrho(t_{10})\geq\varrho(t_{10})\frac{p_{11}}{p_{11}}+\varrho(t_{00})\frac{p_{10}-p_{01}}{p_{11}}+\varrho(t_{01})\frac{p_{11}-p_{01}}{p_{11}}$$

which along with $\mu\leq 1$, yields

$$\boxed{\frac{p_{10}}{p_{11}}\leq\frac{p_{01}}{p_{11}}\frac{\varrho(t_{00})+\varrho(t_{01})}{\varrho(t_{00})+\varrho(t_{10})}+\frac{\varrho(t_{10})-\varrho(t_{01})}{\varrho(t_{00})+\varrho(t_{10})}.}\tag{23}$$

Now, if $\frac{p_{10}}{p_{11}}\geq\frac{\varrho(t_{10})}{\varrho(t_{00})+\varrho(t_{10})}$, condition (23) is trivially satisfied when (19) is satisfied, which means that (23) imposes additional restrictions only when $\frac{p_{10}}{p_{11}}<\frac{\varrho(t_{10})}{\varrho(t_{00})+\varrho(t_{10})}$.

"$\Leftarrow$" direction:

From the above analysis, if the parametric conditions mentioned hold, then the strategy profile that has:

1. the following signaling strategy:
   - t_{11} sends m_{11} always,
   - t_{01} sends m_{00} at I_{*1} and m_{01} at I_{*0}
   - t_{10} sends m_{00} at I_{1*} and m_{10} at I_{0*}
   - t_{00} sends m_{00} always
2. the following Expert's behavior
   - Never cheat under any $w\in\Omega$ if m_{11} is received
   - Never cheat under w_{10} if m_{01} is received
   - Cheat with probability $\frac{\pi_{11}h_{11}}{\pi_{10}A_{10}}$ under w_{10} if m_{10} or m_{00} is received
   - Cheat with probability $\min\left\{\frac{\pi_{11}h_{11}}{\pi_{01}A_{01}},1\right\}$ under w_{01} if m_{01} is received
   - Never cheat under w_{01} if m_{10} or m_{00} is received
3. the following (non-trivial) acceptance probabilities[15]
   - $\alpha(I_{*1},m_{01})=\begin{cases}\frac{p_{01}}{p_{11}} & \text{if } \pi_{11}h_{11}\leq\pi_{01}A_{01},\\ 1 & \text{otherwise}\end{cases}$
   - $\alpha(I_{*1},m_{00})=1$

[15]Where $[x]_+=\begin{cases}x & \text{if } x>0\\ 0 & \text{otherwise}\end{cases}$.

- $\alpha(I_{1*}, m_{10}) = \frac{p_{10}}{p_{11}}$
- $\alpha(I_{**}, m_{00}) = \left[\frac{p_{10}}{p_{11}}\left(1 + \frac{\varrho(t_{10})}{\varrho(t_{00})}\right) - \frac{\varrho(t_{10})}{\varrho(t_{00})}\right]_+$
- $\alpha(I_{1*}, m_{00}) = \frac{p_{10}}{p_{11}}\left(1 + \frac{\varrho(t_{00})}{\varrho(t_{10})}\right) - \frac{\varrho(t_{00})}{\varrho(t_{10})}\alpha(I_{**}, m_{00})$

and

4. the beliefs that result from Bayes's rule on the equilibrium path and off-equilibrium-path beliefs

- $\varrho'(t_{01}|w_{01}, m_{01}) = 1$
- $\varrho'(t_{10}|w_{10}, m_{10}) = 1$

is a passive protection equilibrium. All the above analysis holds symmetrically by making the index changes: $01 \longleftrightarrow 10$, $*1 \longleftrightarrow 1*$, and $*0 \longleftrightarrow 0*$. □

B.5 Proof of Observation 1

Notice that a Consumer type who knows that one of the two issues—say issue 1—does not need fixing (i.e. he is at info set I_{0*} or I_{01}), in the worst case scenario (in which he is offered o_{11} by the Expert and rejects) pays $p_{01} + h_{01}$ to the honest expert. In the best case scenario for him, he pays p_{01} to the Expert. Symmetrically for a Consumer type at info set I_{*0} or I_{10}. Therefore, the largest amount that disclosing one's expertise can save is $\max\{h_{10}, h_{01}\}$.

Now, if $c > \max\{h_{10}, h_{01}\}$, then it is strictly dominant for all Consumers to send message m_{00}. Consumers who know that some issue does not need fixing will still reject fraudulent offers of o_{11}. Since the conditions of Corollary 1 are satisfied, the Expert will never cheat after receiving m_{00}, which establishes full protection as the unique equilibrium outcome. □

B.6 Proof of Proposition 5

In the construction of the separating equilibrium of Proposition 2, it was noted that partially informed Consumers at information set I_{1*} (I_{*1}) are *indifferent* between sending m_{10} (m_{01}) and m_{00}. Therefore, sending m_{00}, i.e., pooling with type t_{00} Customers is also a best response for them. Keeping all other signaling choices (and, thus, payoffs on the equilibrium path) the same as in the separating equilibrium establishes the pooling equilibrium described in the statement of the Proposition. □

References

Academy of Medical Royal Colleges (2014). *Protecting resources, promoting value: a doctor's guide to cutting waste in clinical care*. Tech. rep. Academy of Medical Royal Colleges.

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.

Dulleck, U., R. Kerschbamer, and M. Sutter (2011). "The economics of credence goods: An experiment on the role of liability, verifiability, reputation, and competition". *American Economic Review* 101 (2), pp. 526–55.

Dye, R. A. (1985). "Disclosure of nonproprietary information". *Journal of accounting research*, pp. 123–145.

Ely, J. C. and J. Välimäki (2003). "Bad reputation". *The Quarterly Journal of Economics* 118 (3), pp. 785–814.

Eső, P. and C. Wallace (2018). "Bargaining with Hard Evidence". *The Economic Journal* 0 (ja). DOI: 10.1111/ecoj.12632.

FSCS (2019). *Monthly PPI refunds and compensation*. Tech. rep. Financial Services Compensation Scheme.

Fong, Y.-F. (2005). "When Do Experts Cheat and Whom Do They Target?" *The RAND Journal of Economics* 36 (1), pp. 113–130.

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

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

Green Flag Limited (2018). *Stop unnecessary work on your car. Find out what jobs really need doing*. Tech. rep. Green Flag Limited.

Hagenbach, J. and F. Koessler (2017). "Simple versus rich language in disclosure games". *Review of Economic Design* 21 (3), pp. 163–175. DOI: 10.1007/s10058-017-0203-y.

Hart, S., I. Kremer, and M. Perry (2017). "Evidence games: truth and commitment". *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.

Johnson, E. M. and M. M. Rehavi (2016). "Physicians treating physicians: Information and incentives in childbirth". *American Economic Journal: Economic Policy* 8 (1), pp. 115–41.

Kerschbamer, R., D. Neururer, and M. Sutter (2019). "Credence goods markets and the informational value of new media: A natural field experiment". *MPI Collective Goods Discussion Paper* (2019/3).

Lyu, H., T. Xu, D. Brotman, B. Mayer-Blackwell, M. Cooper, M. Daniel, E. C. Wick, V. Saini, S. Brownlee, and M. A. Makary (2017). "Overtreatment in the United States". *PLoS One* 12 (9), e0181970.

Malhotra, A, D Maughan, J Ansell, R Lehman, A Henderson, M Gray, T Stephenson, and S Bailey (2015). "Choosing Wisely in the UK: the Academy of Medical Royal Colleges' initiative to reduce the harms of too much medicine". *British Medical Journal* 350, h2308.

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

Sher, I. (2011). "Credibility and determinism in a game of persuasion". *Games and Economic Behavior* 71 (2), pp. 409 –419. DOI: http://dx.doi.org/10.1016/j.geb.2010.05.008.

The American Automobile Association (2016). *Most U.S. Drivers Leery of Auto Repair Shops*. Tech. rep. The American Automobile Association.

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