Delegation as a signal: implicit communication with full cooperation

Joanna Franaszek

February 2018

Abstract

This paper examines the issue of implicit signaling of inexpressible type through delegation. I examine a communication game with perfectly aligned preferences, two-sided private information and communication frictions. The model is analyzed in the context of medical decision-making. A patient (principal) comes to a doctor's (agent's) office to choose one of two treatments that would suit his health needs. The patient perfectly knows, but cannot communicate his preference type and may acquire some informative, but imperfect and costly signal about his health. After observing the signal, he may choose the treatment or delegate the decision to the doctor, who observes the health perfectly. Even if the patient information acquisition and the signal are unobservable to the doctor, the patient's delegation choice, combined with the doctor's private information, allow the latter to extract some signal about the inexpressible preference type. I show that for a large class of parameters there exists an equilibrium, in which the doctor, basing on his information and the delegation decision can correctly understand cues about preferences and tailor the final treatment to the patient's needs. In particular the doctor's final decision (upon delegation) may be non-monotone in health.

1 Introduction

How can one party pass some information to their partner without using direct communication? One way is to signal through choice of action. However, choosing not to act – here, meaning delegation of authority – can also give the other party a signal about our private information. Such ”indecisiveness” might turn out to be beneficial even if we are relatively well informed about the decision-relevant variable.

I am especialy grateful to Paweł Gola and Matteo Foschi, who discussed early versions of the model; their feedback was invaluable and I cannot thank them enough for their time and insight. I would like to thank for valuable comments of Piero Gottardi, Andrea Mattozi, Andrea Galeotti, Wouter Dessein and the participants of the seminars in the EUI, 2017 Ce2 Workshop and Warsaw Economic Seminar.

In this paper, I examine a general principal–agent model in the context of medical decision– making. In contrast to the existing literature, I assume the agent (doctor) to be fully altruistic towards the principal (patient) – thus there is no intrinsic divergence of preferences. However, players face a problem of two-sided private information and difficulties in communicating their signal to the other party. Only the principal knows his private preference parameter, while it is the agent who perfectly observes the state of the world. The principal might privately obtain an informative but costly and imperfect signal about the state and then either decide about the action or delegate the decision to the agent. In the equilibrium, the agent correctly anticipates which principal types choose to be informed and under what conditions they would further decide to delegate authority, even though he observes nothing but the delegation decision. In particular, a positive investment in a cheap informative signal is correctly anticipated to come from a principal type that is not extreme. Moreover, the informed types delegate authority only if their acquired signal is inconclusive. This fact, along with the agent's knowledge of true x, allows him to correctly guess the most likely range of types and, in turn, adjust his choice of action to better tailor the principal needs. In other words, the principal implicitly (and imperfectly) signals his type with delegation and the sophisticated agent is able to understand the "cue". This results in a nontrivial equilibrium, in which the agent's final decision (upon delegation) is non-monotone in the state of the world.

The rest of the paper is structured as follows: first, I describe a motivating example and place the framework within a specific type of principal-agent situation, namely a relationship between a doctor and a patient. Then I briefly summarize links with the broader literature on delegation, imperfect signals and language frictions. In the next section, I introduce the model and present some early results.

Motivating example Consider a patient coming to the doctor's office to discuss two possible treatment options. The patient has some private preference about the two actions, which may be interpreted as a decision cutoff; if his state of health x is below some t, he prefers one treatment – say, a surgery – while if his state of health is above t, the other treatment (say, drugs) is preferred. The information about t is a form of tacit knowledge, which is difficult, or even impossible to express in terms of any language – an assumption that may be considered reasonable in the case of medical preferences.

The patient may obtain a costly informative signal about the state of health. Neither the investment in information nor the realization itself is observed by the doctor. After observing the signal, the patient may choose the preferred treatment himself or delegate the decision to the doctor, who, being an expert, observes x perfectly. I assume the doctor's utility coincides with the patient's and there is no inherent divergence of preferences - however, the exact value of t is known only to the patient, while the doctor has some prior $g(t)$.

Suppose the doctor ex-ante expects the patient types to be distributed uniformly, with half of the patients preferring a surgerical operation and the other half - a drug treatment. If the patient was not allowed to acquire information and delegated the choice to the doctor, the doctor would naturally choose surgery for $x < \frac{1}{2}$ and drug treatment for $x > \frac{1}{2}$. The altruistic doctor has no incentive to suggest surgery for any $x > \frac{1}{2}$. The situation changes, however, if the patient is allowed to obtain an informative signal.

Consider a doctor who is encountered by a patient with relatively low $x < \frac{1}{2}$, i.e. an indication for a surgery for an ex-ante average type t. The doctor believes the patient might have obtained a signal which was informative. He therefore infers that an average informed patient *should* know that surgery is his best choice. Yet, the patient still prefers to delegate authority. The doctor might then correctly infer that the delegation decision is more likely to come from a patient with low t, as types with $t \approx x$ find it most difficult to make an informed decision.[1]. Thus, the doctor's posterior belief about t becomes correctly "biased". Based on a belief, the doctor's strategy would change as well – he is less likely to recommend $a = 0$ whenever the decision is delegated and may suggest it only for lowest states of x, while some patients with health $x \in [\bar{x}, \frac{1}{2}]$ are prescribed $a = 1$. Since similar reasoning applies to relatively high types ($x > \frac{1}{2}$), the signaling–by– delegation leads to an equilibrium in which the doctor's choice is non-monotone in health; the patient with worse health is prescribed a less aggressive treatment than the person with better health, solely due to (correctly) signaled preferences.

We may imagine multiple situations in which the apparent "indecisiveness" of the patient does indeed provide a cue about his preferences. If a possibly well-informed patient with a bacterial infection is uncertain whether he wants to be treated with antibiotics, the doctor has a cue about the patient's general dislike for antibiotic therapy, and would take that into account when making the final decision. Similarly, if a physically healthy woman with an uncomplicated pregnancy wants to discuss with her obstetrician a possibility of cesarean section, the doctor might infer that his patient is particularly afraid of the risks of natural birth and tailor his recommendation to the woman's needs.

Related literature The model is connected to a few strands of literature. The first area is a vast literature on delegation and its association with communication, in line with the idea of Dessein (2002). However, while most of the delegation literature (see e.g. Li and Suen (2004); Alonso and Matouschek (2008); Garfagnini, Ottaviani, and Sørensen (2014) concentrate on strategic incentives with divergent preferences of the decision maker and the expert(s), I focus solely on communication frictions and make a bit unpopular assumption that the two parties share same preferences. This assumption, a bit similar to the one in Dewatripont and Tirole (2005), not only allows me to examine information transmission in isolation but also changes the players' incentives. The players want to exchange as much information as possible – both through direct communication and through signaling – but

[1] The delegation decision comes from a principal whose signal was inconclusive, like in Li and Suen (2004); Garfagnini, Ottaviani, and Sørensen (2014)

face language constraints in communication.

My setup is similar to Garfagnini, Ottaviani, and Sørensen (2014), with nondivergent preferences. The main difference – which is also the main contribution of this paper – is that Garfagnini, Ottaviani, and Sørensen (2014) consider two versions of the model, in which the principal is either uninformed or informed, while in this paper the choice to become informed is endogenous and unobservable. This, combined with an ability to signal one's information through the allocation of authority, would guarantee the existence of a sophisticated equilibrium with cues about one's type.[2]

The cues sent in the model have a flavor of signaling as in Spence (1973). However, while in classic signaling models the correlation between the unobserved type and the signal is explicit i.e. the costs of signaling are lower for types with higher productivity. In my setup, the correlation between the unobserved preference and the state, observed by the doctor arises endogenously and only through a complex mechanism of correlated informative signals and optimal decisions. Moreover, the patient's (principal's) signaling through delegation is only to a little extent driven by signaling incentives. In fact, even if the doctor was blind to signaling and chose the action only according to health and the prior belief about the patient type, there would be still room for delegation for at least some patient types. However, the model is much more interesting if the doctor can "infer beliefs from actions" (see Arieli and Mueller-Frank (2017)).

The transmission of complex medical information falls into the strand of the literature on dissemination of knowledge. As noticed e.g. by Boldrin and Levine (2005), the mere *availability* of information (in terms of e.g. results of medical tests) does not make the information *accessible* to a person, who may lack the expertise to interpret it. Communication is, therefore, costly (see also Austen-Smith (1994); Hedlund (2015); Eso and Szentes (2007); Gentzkow and Kamenica (2014) for other models of explicitly costly information transmission). Moreover, complex knowledge takes years to build and cannot be easily and costlessly transmitted. In fact, some information – such as the patient's preferences toward alternative treatments – might be impossible to verbalize. This *tacit knowledge*, as defined in Polanyi (1966), can only be transmitted through non-verbal cues, as in this model, where it is signaled in the equilibrium choices.

Finally, it is practical to compare the presented interpretation of the model to other models of doctor-patient relationship. While many models (see Xie, Dilts, and Shor (2006); Lubensky and Schmidbauer (2013); Ehses-Friedrich (2011); Johnson and Rehavi (2016)) assume some divergence of preferences, there are some which focus on altruism. Koszegi (2004) considers an altruistic doctor, who cares about the emotional well-being of the patient, which results in distorted, overly optimistic messages about the patient's health.[3]

[2] I use the word "cues", as in Dewatripont and Tirole (2005), however, the meaning of the term is very different. While in Dewatripont and Tirole (2005) sending cues is a substitute for a potential costly communication, in my model it is a way to implicitly signal one's type in the presence of language constraints

[3] The doctor-patient setup in Koszegi (2004) turned later into more general model of Koszegi (2006) in which also other applications of an altruistic agent model are proposed.

In my model, the doctor only cares about the relationship between the state of health and the optimal treatment, thus the only distortions arise in the communication process.

2 Model

To examine the communication frictions in isolation, I assume the doctor to be fully altruistic towards the patient. In other words, the doctor and the patient have the same utility function, which depends on the patient's state of health and the choice of treatment. In particular, I assume the utility is $u_t(x, a)$, where $x \in [0, 1]$ is health, $a \in \{0, 1\}$ is action, that is the choice of one of two treatments (e.g. "surgery" and "drugs"), $t \in [0, 1]$ is patient's preference type. Note that x is only observed by the doctor, and t is only observed by the patient. The two parameters can only be imperfectly transmitted to the other party.[4] The prior distribution of x and t, are, respectively $U[0, 1]$ and G_t with some continuous, full support density $g(t)$, and this is common knowledge. I assume $g(t)$ is symmetric around an axis $t = \frac{1}{2}$. This assumption is not crucial in establishing the existence of an equilibrium, but significantly helps in understanding the main contribution of the paper. I would denote $E_g(t|t \in [\frac{1}{4}, \frac{1}{2}]) = \tau$ and describe some of the results in relation to τ. I will show that there exists an equilibrium in which the patient chooses a symmetric strategy, but the understanding of the cue makes the doctor's belief (correctly) biased. For some range of signals the bias is sufficient to change the ex-ante optimal action, which results in an interesting non-monotone action profile. The family of signals becomes larger as τ gets smaller, i.e. as $g(t)$ becomes less concentrated around its mode.

I assume that both players share the same utility function.

Assumption 1. *The utility u has the following properties:*[5]

- *$u_t(x, a)$ is (weakly) increasing in x for $a = 0, 1$*
- *$\exists t$ such that $u_t(x, 0) > u_t(x, 1)$ for $x < t$ and $u(x, 0) < u(x, 1)$ for $x > t$*
- *$v(x) = u_t(x, 1) - u_t(x, 0)$ is weakly increasing in x*

The first assumption is quite straightforward: utility increases with x, as the patient enjoys more health. The second assumption allows us to interpret type t as the "private benchmark" of the patient-if his health falls below the benchmark, the patient prefers treatment $a = 0$, otherwise the patient would rather choose treatment $a = 1$. The treatments

[4] The assumption that x and t are some numbers in the unit interval is just for expositional simplicity. We could imagine a setup in which x is a point in some multidimensional abstract space. The space is separated into two compact areas and in each of them one treatment is preferred to the other. The type t would then be a boundary (e.g. a line, a plane, a manifold) between the two. However, the continuity and monotonicity conditions that are stipulated for unit interval need to be adapted to the multidimensional setting.

[5] Only the last property is crucial for establishing the results, the first two are just associated with my interpretation of the model as the utility from health. All the results would hold with those assumptions relaxed.

could be labeled according to a specific situation that we have in mind, e.g. in a bad state of health patient would agree to have surgery $(a = 0)$, while if he enjoys decent health, he would prefer to be treated with drugs $(a = 1)$. What is crucial is that all types of patient share the same ordering of actions, i.e. prefer $a = 0$ in low states and $a = 1$ in high states, while they might, of course, differ in perception of what is the cutoff between "bad" and "good" health. Finally, the monotonicity assumption means that the gains from choosing the preferred treatment are greater for more extreme states of health, i.e. the disutility from not having an operation when health is very bad is higher than when x is just below the benchmark.

Throughout this version of the paper, u is taken to be linear, i.e.

$$u(x, a) = a(x - t) \text{ for } a \in \{0, 1\}.$$

I assume that the private preference parameter t is a form of Polanyi's tacit knowledge, which cannot be explained in terms of language and is therefore impossible to communicate either via cheap-talk or any form of disclosure. The only information about t that the doctor can have comes from his beliefs regarding the patient's observed actions.

The state of health is a complex medical term which can seldom be precisely expressed in everyday language. Therefore, any information about x is necessarily imperfect. Moreover, understanding at least some information about x requires some mental or monetary cost, which could be interpreted as a cost of translating medical terms into everyday language, effort in communication, time devoted to explanations etc.

In this paper, I assume that the choice of information acquisition is binary,[6] i.e. the patient may decide to acquire an informative signal about his state of health x at a cost c or remain uninformed (equivalently: receive an uninformative signal) at no cost. The information might come from some private source (books, self-administered tests, the Internet), but a setup could also be used to analyze the case in which the patient acquires information from the doctor himself. In this interpretation, the doctor tries to communicate the state of health and the patient may exert zero or positive (namely, c) effort in understanding it. The doctor can observe neither the effort choice nor the final realization of the signal (i.e. what the patient understood from his explanation). To simplify the analysis I assume that an informative signal is binary, i.e. $s \in \{0, 1\}$ and could be interpreted as a recommendation of action, such that action 1 is recommended more often for higher states.[7] In particular, I shall assume that for any x the probability of signal $s = 1$ is $P(s = 1|x) = p(x)$, such that $p(x)$ satisfies:

Assumption 2. *The probability of signal $s = 1$, denoted by $p(x)$, satisfies:*

[6] Binary choice is sufficient to obtain the most interesting "cue" result. Preliminary results about continuous choice of investment in information suggest that one needs to be more careful when checking incentive compatibility conditions for any arbitrary investment in information. However, the basic result should still hold.

[7] Note that such a recommendation is *not* cheap-talk.

1. *$p(x)$ is increasing in x, with $p(0) = 0$ and $p(1) = 1$.*

2. *The signaling technology is symmetric around $x = \frac{1}{2}$, i.e. $p(x) = 1 - p(1-x)$. In particular, $p(\frac{1}{2}) = \frac{1}{2}$.*

3. *$p(x)$ is S-shaped with $\int_0^{1/2} p(x) < \epsilon(\tau)$.*

The first assumption is quite straightforward. The second means the signal is "fair", in a sense it treats low states and high states symmetrically. The third assumption ensures that $p(x)$ is sufficiently steep around $x = \frac{1}{2}$ – in other words, the signal discriminates well between states that are higher and smaller than the average.

The timing of the model is as follows:

1. Nature draws x (learned by the doctor) and t (learned by the patient).

2. The patient chooses to acquire an informative signal about x at cost $C = c$ (or an uninformative signal at cost $C = 0$).

3. After observing $s|x$ the Patient decides to:

   (a) retain the authority,

   (b) delegate the decision to the doctor.

4. The chosen decision-maker chooses an action $a \in \{0, 1\}$.

5. The utility $u_t(x, a) - C$ is realized.

I would show that by the choice of information and then delegation, the patient implicitly signals his type t and signal realization s. However, contrary to classic signaling models, his choices are not pure signals. In fact, the patient decides to acquire information primarily to improve his likelihood of the right choice and the doctor's strategy only slightly enhances the patient's incentives. This makes the acquisition choice particularly robust. However, the implicit signaling feature would be crucial in examining the equilibrium behavior.

2.1 The limit case

To get an intuition about the result, let us examine the limit case, in which the signal is of particularly simple form: $p_{lim}(x) = 1_{\{x \geq \frac{1}{2}\}}$. Such a signal is the "most informative" of symmetric binary signals, as it gives the precise location of x. Assume $g(t) = U[0, 1]$ and $c < \frac{1}{36}$.

I claim that the equilibrium is as follows: the patient acquires an informative signal whenever $t \in \left[\frac{1}{4}, \frac{3}{4}\right]$. Moreover, for $t \in \left(\frac{5}{12}, \frac{7}{12}\right)$ he would retain the authority, choosing an action in-line with the signal. For $t \in \left[\frac{1}{4}, \frac{5}{12}\right]$ he would delegate if the signal is $s = 0$ and for $t \in \left[\frac{7}{12}, \frac{3}{4}\right]$ he would delegate if the signal is $s = 1$. The doctor chooses $a = 1$ (upon hearing delegation) if and only if $x \in \left[\frac{1}{3}, \frac{1}{2}\right] \cup \left[\frac{2}{3}, 1\right]$, and thus, his strategy is non-monotone in health.

Let us first analyze the doctor's strategy, taking the patient's choice as given. Upon hearing delegation, the doctor anticipates $t \in \left[\frac{1}{4}, \frac{5}{12}\right] \cup \left[\frac{7}{12}, \frac{3}{4}\right]$. However, since x is observed by the doctor, he knows exactly which value of the signal must have been realized.[8] Therefore, if $x < \frac{1}{2}$ he knows delegation comes from $t \in \left[\frac{1}{4}, \frac{5}{12}\right]$ and for $x > \frac{1}{2}$ the delegating type must be $\left[\frac{7}{12}, \frac{3}{4}\right]$. His posterior belief is then $E(t|D,x) = \frac{1}{3}$ for $x < \frac{1}{2}$ and $E(t|D,x) = \frac{2}{3}$ for $x > \frac{1}{2}$. The doctor chooses $a = 1$ whenever $x \geq E(t|D,x)$, which leads exactly to the profile above. As for the patient, a formal derivation is a bit more tedious, but the intuition is simple: everyone apart from extreme types gets cheap information. Middle "unbiased" types follow the signal and retain authority; "Biased" types follow the signal if it confirms their prior choice and delegate whenever they become uncertain about the optimal action, i.e. whenever s is close to their type.

3 Equilibrium choices

The equilibrium concept is a Perfect Bayesian Equilibrium. Denote the patient's strategy as $(C(t), \sigma(s,t), a(s,t))$ with the investment in information $C : [0,1] \to \{0, c\}$, allocation of authority $\sigma(s,t) : [0,1]^2 \to \{D, P\}$ and the choice of action $a^P(s,t) \to \{0,1\}$. Denote the doctor's response after delegation as $a^D(x) : [0,1] \to \{0,1\}$ and his posterior belief as $\mu(s,t|\sigma,x)$. Since we are mainly interested in the doctor's posterior belief about the type, let us also denote $g(t|D,x) := E_s\mu(s,t|D,x)$.

I shall propose a specific form of an equilibrium and prove its existence. In the putative equilibrium, the patient strategy is symmetric around $t = \frac{1}{2}$. Extreme types of patients do not acquire information and choose the action by themselves (since they are already certain about their decision). The middle types, who are ex-ante close to indifference, either acquire information (if it is cheap) and make an informed decision or remain uninformed (if the information is expensive) and delegate the authority. Somewhat biased types chose conditional delegation i.e. they acquire an informative signal and delegate the decision only if it is "inconclusive". The doctor chooses an action, based on his belief about x, choosing $a^D(x) = 1$ if $x > E_\mu(t|D,x)$ and $a^D = 0$ if $x < E_\mu(t|D,x)$. More specifically, the doctor chooses $a = 1$ if and only if $x \in [\bar{x}, \frac{1}{2}] \cup [1-\bar{x}, 1]$, with $\bar{x} \leq \frac{1}{2}$. Observe that for $\bar{x} = \frac{1}{2}$ the strategy coincides with the trivial "ex-ante" profile; however, in the more interesting case of $\bar{x} < \frac{1}{2}$ the strategy is non-monotone in health.

[8] Kartik (2015) describe a two-dimensional information that is *muddled* into one-dimensional action, so "any observed action will generally not reveal either dimension". Here, the information is de-muddled – the doctor uses his knowledge of x to separate "low types with low signal" from"high types with high signal". Such a phenomenon would arise only imperfectly in the general model, where the probability of any signal realization is non-degenerate.

3.1 Action choice

If the patient makes the decision himself, he has no strategic interaction to consider. Therefore, he would choose $a = 1$ if $E(x|s) > t$ and $a = 1$ otherwise. Notice that for an uninformative (zero cost) signal $E(x|s) = Ex = \frac{1}{2}$. On the other hand, for informative signal the expectation $E(x|s)$ depends on $s \in \{0, 1\}$. Then $E(x|s = 1) > \frac{1}{2} > E(x|s = 0)$. However, the general rule $a = 1 \Leftrightarrow E(x|s) > t$ does not change.

If the decision was delegated to the doctor, he would choose one action over another based on the value of x. From the patient's point of view, the optimal doctor's choice of action $a^D(x)$ can be determined using the posterior belief about t, that also depends on the Doctor's information x. In the equilibrium, the patient can correctly anticipate the doctor's posterior belief $E(t|D, x)$ and takes $a^D(x)$ as given. Assume $a^D(x) = 1_{\{\bar{x} \leq x \leq \frac{1}{2}\}} + 1_{\{1-\bar{x} < x \leq 1\}}$ as described above.

Once the action choice in the last step is determined, we can proceed with backward analysis of the patient's incentives in each step. Through this section, let us denote by $V(t, C, \sigma)$ the expected value of an information investment choice C and allocation of authority σ (which may be conditional on observed s).

3.2 Allocation of authority

First, let us assume that the patient did not invest in informative signal and therefore has only prior belief on x. Such a patient would decide to delegate rather than retain authority if[9]:

$$V(t, 0, D) \geq V(t, 0, P) \Leftrightarrow \int_0^1 (x - t) a^D(x) dx \geq \max\left(\frac{1}{2} - t, 0\right). \tag{1}$$

I shall denote the types who choose uninformed (therefore, unconditional) delegation by Ω_{UD}.

Claim 1. Suppose t does not acquire information. Then $t \in \Omega_{UD}$ iff $(1 - t) \in \Omega_{UD}$ and if $\Omega_{UD} \neq \emptyset$ then $\frac{1}{2} \in \Omega_{UD}$.

A (very simple) proof of this claim and all subsequent can be found in the Appendix.

Now, consider the case with an informative signal. The patient would only invest in a signal if he is willing to use it. We can therefore exclude the case in which the patient would choose action that goes against the signal realization, as such a patient would prefer not to acquire information at all. Indeed, if a patient observes s either he would choose $a = s$ immediately or delegate the decision to the doctor. The patient prefers delegation to retainment if:

$$V(t, c, D|s) \geq V(t, c, P|s) \Leftrightarrow E(x - t|s,\ D) \geq E(x - t|s, P)$$

[9] I implicitly assume that the indifferent patient chooses delegation.

$$\begin{cases} \int_0^1 (x-t)p(x)a^D(x)dx \geq \int (x-t)p(x) & \text{if } s=1, \\ \int_0^1 (x-t)(1-p(x))a^D(x)dx \geq 0 & \text{if } s=0. \end{cases}$$

With simple algebra and an observation that $\frac{\int xp(x)1_{\{a^D=i\}}dx}{P(a^D=i)} = E(x|s, a^D = i)$, the conditions could be summarized as:

$$\begin{cases} E(x|s, a^D = 0) \leq t & \text{for } s=1, \\ E(x|s, a^D = 1) \geq t & \text{for } s=0. \end{cases} \tag{2}$$

Claim 2. For any informative signal realization s, there exists a range of patient types Ω^s_{CD}, who prefer to delegate the authority, conditionally on being informed.

To get an intuition about the result, notice that for any signal realization at least type $t = E(x|s)$ would find it strictly profitable to delegate, as conditional on his information, he is indifferent between the two actions. Thus, he may benefit by delegating to the possibly better–informed doctor. By continuity, there exists a range of types around $t = E(x|s)$ who would also prefer delegation. The intuition that relatively large (small) types delegate whenever the signal is large (small)- and thus the type is implicitly correlated with signal – would be crucial in understanding the equilibrium communication. Full proofs of all the claims can be found in the Appendix.

Claim 3. The sets Ω^0_{CD} and Ω^1_{CD} are symmetric around an axis $t = \frac{1}{2}$ and disjoint i.e. $\Omega^0_{CD} = [\underline{t}, \bar{t}]$ and $\Omega^1_{CD} = [1-\bar{t}, 1-\underline{t}]$ for some $\underline{t} < \bar{t} < \frac{1}{2}$.

The symmetry of delegation decision is a direct result of the symmetry of a^D and the signal $s|x$ around $x = \frac{1}{2}$. This implies the symmetry of the delegation decision. There is no type who delegates for both signal realizations, as such a type would prefer to deviate to not acquiring signal at all (and delegating immediately).

3.3 Information acquisition

Take the strategies in the second period described in the previous subsection as given and assume the doctor's. Going one step back, the patient needs to decide whether to invest in an informative signal or not. The expected value of the decision in the second step is the maximum of the two possible options (delegation or retainment) for both levels of investment. Therefore, the patient would choose to acquire an informative signal:

$$E_s \max\{V(t,c,D|s), V(t,c,P|s)\} - c \geq \max\left\{V(t,D,0), V(t,P,0)\right\}.$$

Lemma 1. *The set of types who acquire information Ω_c is symmetric around $t = \frac{1}{2}$. Extreme types $t = 0$ and $t = 1$ (and their neighborhood) never acquire information.*

For a given signal distribution p there exists two upper bounds ψ, ϕ, such that if $c \in (0, \psi)$ then $\Omega_c \ni \frac{1}{2}$ and there exist types who acquire information and choose according to their signal. In such a case the set Ω_{UD} is empty. If $c \in [\psi, \phi]$, then all types who acquire information delegate

Figure 1: The investment in information and delegation choice for $c < \psi$ (left) and $\psi \leq c \leq \phi$ (right).

conditionally on their signal $\Omega_c = \Omega_c \cap \left(\Omega^0_{CD} \cup \Omega^1_{CD}\right)$. *If* $c > \phi$, *nobody acquires information and the game is trivial.*

The Lemmas are simply summarized in Figure 1. Extreme types close to $t = 0, 1$ have such strong preferences towards one of the treatments that they do not feel the need to invest in information. If the cost of information is small, all "medium types" acquire information and some of them choose conditional delegation. Notice that by Claim 3, informed-and-delegating types form two disjoint intervals, therefore types close to $t = 1/2$ choose to make a decision by themselves. This is pictured in the left panel of Figure 1.

If the cost of the signal is big enough, the types close to $t = \frac{1}{2}$ are hit by a "median patient curse".[10] Notice that type $t = \frac{1}{2}$ finds it ex-ante most difficult to choose between the two actions, therefore he has an incentive to acquire information. However, since the information is (ex-ante) symmetric, type $t = \frac{1}{2}$ would expect it to be inconclusive and costly. Therefore, he would rather delegate to a perfectly informed doctor.

Note that by Lemma 1, there are no other switches in the patient strategy than those in the Figure 1.

3.4 Doctor's strategy and beliefs

Given the (known) informativeness of the signal, the doctor expects the patient to delegate whenever $t \in \Omega^0_{CD} \cup \Omega^1_{CD} \cup \Omega_{UD}$. The doctor optimally chooses actions, taking into account his expectation about the type. Formally, the doctor chooses $a^D(x) = 1$ for $x \geq E(t|D, x)$ and $a^D(x) = 0$ for $x < E(t|D, x)$. Notice however that the formula $E(t|D, x)$ is not constant and is a function of x. Denote:

$$\beta(x) = E(t|D, x) = \frac{\int tg(t|D, x)dt}{\int g(t|D, x)dt},$$

where $g(t|D, x)$ is the interim doctor's belief given the observed x and the expected patient's strategy $\{\sigma(s) = D\} \Leftrightarrow t \in \Omega_{CD} \cup \Omega_{UD}$. For a given x, the posterior belief about the type distribution $g(t|D, x)$ would typically *not* have a full support, neither will it be symmetric around $t = 1/2$. The doctor knows x, therefore, he can correctly infer what are the probabilities of acquiring a specific signal. In particular, the interval "closer" to x is more likely than the other. Thus, if doctor observes x (say, $x > \frac{1}{2}$) and delegation, then

[10] Notice that for a symmetric distribution *median* = *mean*.

he correctly infers that the most likely signals are those "close to" x, and such signals are indecisive for types t "close to" x. Therefore, his posterior belief about types is skewed towards x. Formally:

$$g(t|D,x) = \begin{cases} \frac{g(t)(1-p(x))}{A(x)} & \text{for } t \in \Omega_{CD}^0 \\ \frac{g(t)p(x)}{A(x)} & \text{for } t \in \Omega_{CD}^1 \\ \frac{g(t)}{A(x)} & \text{for } t \in \Omega_{UD} \\ 0 & \text{for all other } t \end{cases}$$

where $A(x) = (1-p(x)) \int_{\Omega_{CD}^0} g(t)dt + p(x) \int_{\Omega_{CD}^1} g(t)dt + \int_{\Omega_{UD}} g(t)dt$. Note that since g is symmetric then $\int_{\Omega_{CD}^1} g(t) = \int_{\Omega_{CD}^0} g(t)$, therefore $A(x) = \int_{\Omega_{CD}^0} g(t)dt + \int_{\Omega_{UD}} g(t)dt$.

Lemma 2. *Assume g is symmetric around an axis $t = \frac{1}{2}$. Then in any equilibrium, C, σ are symmetric around an axis $t = \frac{1}{2}$, but the doctor's cutoff function (i.e. the a posteriori expected patient type) $E(t|D,x)$ is **not** symmetric around an axis $x = \frac{1}{2}$. Instead $\beta(x) = E(t|D,x)$ is weakly increasing in x (with $\beta(\frac{1}{2}) = \frac{1}{2}$) and $\beta(x) + \beta(1-x) = 1$.*

The statement of the Lemma may not look as exciting as it really is. To fully understand it's value, notice first that if the doctor's posterior belief about the distribution of t was symmetric around $t = \frac{1}{2}$, the function $\beta(x)$ would be constant and equal to $\frac{1}{2}$, independently of x. However, the doctor's belief in the equilibrium is skewed towards the "correct" t, even though the patient's strategies are symmetric. This phenomenon can be only sustained whenever the doctor observes x, because then, given distribution $p(x)$ he can determine more likely signal realizations and, using their equilibrium association with t, infer what are more likely values of t. The doctor not only correctly anticipates the information acquisition choice, but also exploits the correlation of the doctor's and patient's signals. The correlation along with the equilibrium delegation decision allows the doctor to infer the most likely range of t, even though t is never explicitly signaled. However, the important issue is whether the "bias" in posterior beliefs is strong enough to induce the doctor to change the a priori optimal actions. For an S-shaped $p(x)$ this is exactly the phenomenon that may arise.

Denote by $\tilde{\tau} := \frac{\int_{\Omega_{CD}^0} tg(t)dt + \int_{\Omega_{UD}} tg(t)dt}{A}$, that is, the expected type of the delegating patients. Observe that if $\Omega_{UD} = \emptyset$ then $\tilde{\tau} < \tau := E(t|t \in [1/4, 1/2])$. However, in general $\tilde{\tau}$ is determined in the equilibrium – in particular, it depends on c.

Theorem 1. *If $p'(\frac{1}{2}) > \frac{1}{1-2\tilde{\tau}}$ (with a sufficient condition $p'(frac12) > \frac{1}{1-2\tau}$ if $c < \psi$) the doctor's action profile follows a non-monotone pattern, that reflects his asymmetric belief upon observing delegation:*

$$a^D(x) = \begin{cases} 1 & \textit{for } x \in \left[\bar{x}, \frac{1}{2}\right] \cup [1-\underline{x}, 1], \\ 0 & \textit{otherwise,} \end{cases} \quad \textit{for some } \bar{x} < \frac{1}{2}$$

Otherwise, the doctor's action profile in equilibrium coincides with the "naive" one:

$$a^D(x) = \begin{cases} 1 & \textit{for } x \in \left[\frac{1}{2}, 1\right], \\ 0 & \textit{otherwise.} \end{cases}$$

The doctor's choice is pictured in Figure 2. In equilibrium with cues, the doctor recommends action $a = 0$ for x small (which is intuitive), but also for relatively big $x \in \left(\frac{1}{2}, 1 - \underline{x}\right)$. The second interval is the region in which the bias in posterior beliefs plays a dominant role. In particular, even though the doctor knows x is relatively big a priori, the implicit signal coming from the recommendation makes him believe t is even bigger. Therefore action $a = 0$ is preferred. Such a nontrivial profile is only possible when the effect on the posterior beliefs induced by a signal and the delegation decision is strong enough. In particular, the marginal change in beliefs around $x = \frac{1}{2}$ must be large, namely $p'(\frac{1}{2}) > \frac{1}{1-2\tilde{\tau}}$. However, if this requirement is not satisfied, then even though the posterior belief is indeed skewed in the "right" direction, the perturbation is not strong enough to induce a switch from naive beliefs.

As a corollary from the Lemma, we can claim the following, main result of the paper:

Theorem 2. *There exists a Perfect Bayesian Equilibrium of the game with implicit signaling of type through delegation. In such an equilibrium, the patient's strategy is symmetric around $t = \frac{1}{2}$, while the doctor's strategy may be non-monotone in health state. In particular, the equilibrium choices are as follows:*

1. *If $c \leq \psi$ then only the extreme patient remain uninormed (and retain their authority). The middle types acquire inormation and retain authority, while the "somewhat biased" types delegate conditionally on the signal. If $p'(\frac{1}{2}) > \frac{1}{1-2\tilde{\tau}}$, the doctor responds with non-monotone action profile.*

2. *If $c \in [\psi, \phi]$ then only the somewhat biased types acquire inormation. The extreme types remain uninormed and retain the authority. The middle types remain uninormed and delegate authority. If $p'(\frac{1}{2}) > \frac{1}{1-2\tilde{\tau}}$ (which now is a stronger condition than above, as $\tilde{\tau}$ depends on c) , the doctor responds with non-monotone action profile.*

3. *If $c > \phi$ no patient type acquires information. Types $t \in [\frac{1}{4}, \frac{3}{4}]$ delegate and the remaining types retain authority. The doctor's strategy is trivial, as no inormation about the signal is transmitted by delegation.*

The theorem describes signal families, for which the doctor's strategy becomes non-monotone in action. The intuition is that the signaling function should indeed resemble a letter 'S' and be steep around $p(1/2)$. Notice that in the introductory example we examined a limit case with "infinite"[11] steepness and perfectly flat tails. The theorems 1 and 2 explain how this extreme signal structure could be generalized and adapted to continuous signal functions.

[11] Formally, indefinite.

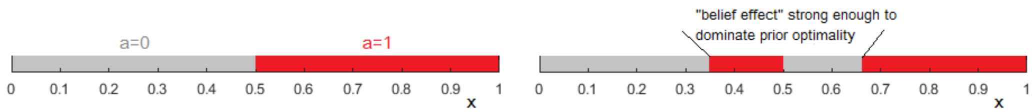

Figure 2: The action profile chosen by doctor in equilibrium if $p'(\frac{1}{2})$ is small (left) and large (right)

4 Summary

In this paper I examine a principal-agent model with two-sided private information and no conflict of interest in a context of doctor–patient communication. I assume that the decision–relevant information is two dimensional, and that each dimension is observed just by one agent. I show that if the principal has access to a costly, but informative signal about the dimension known to the agent, the agent can not only anticipate which types would find it valuable to acquire information, but also correctly infer the cue about the principal type from his decision to delegate or retain the authority.

The model describes a phenomenon of non-verbal communication through equilibrium actions and demonstrates how the joint information about the principal's strategies *and* the agent's private signal allow the latter to make nontrivial conclusions about the type of the former. Such a phenomenon is only possible because the two parties obtain potentially correlated information about the same variable and the delegation choice is an implicit information about both the observed signal and its relation to the principal's preferences. As a result, the agent's belief is correctly biased. Moreover, if the signal is S-shaped and sufficiently precise in distinguishing states $x > \frac{1}{2}$ from states $x < \frac{1}{2}$ the bias is strong enough to change the a priori optimal actions and there may arise an equilibrium, in which the agent's action profile chosen upon delegation is non-monotone in the state of the world.

I focus on communication frictions and implicit signaling to stress that when the information is cheaply available, even if the players observe nothing but the apparent ”indecisiveness” of the other party, a correct inference about their preferences and strategies can still be made. This implicit communication through delegation helps the players to coordinate on the preferred outcome.

Appendix: Proofs

Proof of Claim 1: Inequality in (1) could be rewritten as:

$$\int_0^1 xa^D(x) \geq \max\left(\frac{t}{2}, \frac{1-t}{2}\right).$$

It is now clear that if the condition holds for t, it also holds $1-t$. Also, the RHS is minimized by $t = \frac{1}{2}$, so if any t satisfies the condition, so does $t = \frac{1}{2}$.

Proof of Claim 2: Assume that a^D follows the putative pattern. Intuitively, $a^D = 1$ must be chosen more often for higher states, so:

$$E(x|s, a^D = 0) < E(x|s) < E(x|s, a^D = 1), \tag{3}$$

since

$$E(x|s) = E(x|s, a^D = 0) \cdot P(a^D = 0) + E(x|s, a^D = 1) \cdot P(a^D = 1). \tag{4}$$

Consider t varying from 0 to 1. As t increases up to $E(x|s)$, it must hit a point where $\underline{t}^s = E(x|s, a^D = 0)$ and for all $t \in [\underline{t}^s, E(x|s)]$ the first inequality of (2) is satisfied. Similarly, for $t \searrow E(x|s)$ the second inequality of (2) is satisfied, and as t increases to 1, there exists a point $\bar{t}^s$ such that for any $t > \bar{t}^s$ the delegation is no longer preferred. Therefore, for $t \in [\underline{t}^s, \bar{t}^s] =: \Omega_D^s$ delegation is preferred, while for $t \notin [\underline{t}^s, \bar{t}^s]$ the patient prefers to retain authority. It is important to notice that $[\underline{t}^s, \bar{t}^s] \ni E(x|s)$ and since the delegation decision is made after learning s, the interval differs with the realization of s.

Proof of Claim 3: Assume t satisfies condition (2) for $s = 1$ and we need to show that $1-t$ satisfies it for $s = 0$. Without loss of generality, assume t satisfies the first inequality of (2), i.e.

$$E(x|s = 1) > t \text{ and } E(x|s = 1, a^D = 0) \leq t$$

Then I need to prove that $1-t$ satisfies inequality:

$$E(x|s = 0, a^D = 1) \geq 1 - t.$$

To prove this, it is enough to show that $E(x|s = 0, a^D = 1) = 1 - E(x|s = 1, a^D = 0)$. I shall use double symmetry of all the ingredients:

$$E(x|s = 0, a^D = 1) = \frac{\int x(1-p(x))1_{x\in\{a^D=1\}}}{\int 1 - p(x)1_{x\in\{a^D=1\}}} = \frac{\int xp(1-x)1_{x\in\{a^D=1\}}}{\int p(1-x)1_{x\in\{a^D=1\}}} =$$

$$= \frac{\int (1-y)p(y)1_{y\in\{a^D=0\}}}{\int p(y)1_{y\in\{a^D=0\}}} = 1 - \frac{\int yp(y)1_{y\in\{a^D=0\}}}{\int p(y)1_{y\in\{a^D=0\}}} = 1 - E(x|s = 1, a^D = 0).$$

Therefore, if $t \in \Omega_{CD}^1$ then $1 - t \in \Omega_{CD}^0$.

We might also notice that by more general considerations, the two intervals must be disjoint. If there exists some $t \in \Omega^0_{CD} \cap \Omega^1_{CD}$ who would delegate for any signal, then he would rather not acquire the costly information at all.

Such considerations allow us to use a somewhat simpler notation. Define $[\underline{t}, \bar{t}] := \Omega^0_{CD}$ then $\Omega^1_{CD} = [1 - \bar{t}, 1 - \underline{t}]$.

Proof of Lemma 1: There exist four possible strategies: uninformed decision $(C, \sigma) = (0, P)$, informed decision (c, P), uninformed delegation $(0, D)$, and informed delegation (c, D). Recall that by Claim 3 delegation for an informed agent is always conditional on the signal and is chosen only for a signal realization closer to t, while uninformed delegation is unconditional by definition. Moreover, if the patient finds it optimal to choose informed decision, it must be the case that the chosen actions are different for different signal realizations and consistent with them, more specifically $a^P = s$.

I shall analyze the payoffs of all the above strategies and try to determine what is the optimal profile given generic t. Since the problem is symmetric around $t = \frac{1}{2}$, I shall assume explicitly that $t \geq \frac{1}{2}$. Then:

$$
\begin{aligned}
V(t,0,P) =& \max\left(\frac{1}{2} - t, 0\right) = 0 \\
V(t,0,D) =& \int_0^1 (x - t)a^D(x) = \frac{3}{8} - \left(\frac{1}{2} - \bar{x}\right)^2 - \frac{t}{2} \\
V(t,c,P) =& P(s=0) \cdot 0 + P(s=1)\frac{\int_0^1 (x-t)p(x)dx}{\int_0^1 p(x)dx} - c = \int_0^1 xp(x)dx - \frac{t}{2} - c \\
V(t,c,D) =& P(s=0) \cdot 0 + P(s=1)\frac{\int_0^1 (x-t)2p(x)a^D(x)dx}{\int_0^1 p(x)dx} - c = \int_0^1 (x-t)p(x)a^D(x)dx - c
\end{aligned}
$$

I shall analyze how the optimal strategy changes with t moving from $\frac{1}{2}$ to 1 using five observations:

1. For $t = 1$ uninformed decision $V(t, 0, P)$ dominates any other strategy.
2. Either $V(t, c, P) \geq V(t, 0, D)\ \forall t$ (for c small) or $V(t, c, P) < V(t, 0, D)\ \forall t$ (for c big).
3. $V(t, c, D) > V(t, c, P)$ for t sufficiently big.
4. $V(t, c, D) > V(t, 0, D)$ for t sufficiently big.
5. There exist t such that $V(t, c, D)$ is optimal, in particular, there are always types who acquire information.

The first observation is trivial – as $t = 1$ (or close to 1) all the payoffs become negative, apart from $V(t, 0, p)$. The second observation stems from the fact, that both formulas are of the form $-\frac{t}{2} + constant$, so in comparison, only the constant matters. In particular:

$$V(t,0,D) \leq V(t,c,P) \Leftrightarrow \frac{3}{8} - \left(\frac{1}{2} - \bar{x}\right)^2 \leq \int_0^1 xp(x)dx - c$$

$$V(t,0,D) \leq V(t,c,P) \Leftrightarrow c \leq \underbrace{\int_0^1 xp(x)dx - \frac{3}{8}}_{<0} + \left(\frac{1}{2} - \bar{x}\right)^2 =: \psi \tag{5}$$

The necessary condition being:

$$\psi = \left(\frac{1}{2} - \underline{x}\right)^2 - \frac{3}{8} + \int_0^1 xp(x)dx \geq 0 \tag{6}$$

I shall claim that whenever $\int_0^{\frac{1}{2}} \left(\frac{1}{2} - x\right) p(x)$ is sufficiently small, the inequality (6) holds. Namely, there exists a function $\epsilon(\tau)$ such that when $\int_0^{\frac{1}{2}} \left(\frac{1}{2} - x\right) p(x) < \epsilon(\tau)$, then the inequality above is guaranteed. For clarity purpose, this part of the proof is moved below in the Appendix, under Lemma 3.

The third observation is a bit more subtle. Since $a^D(x)$ is an indicator function $V(t,c,D)$ is a "truncated" version of integral in$V(t,c,P)$, in which some values of the function $(x-t)p(x)$ are replaced by 0. Such a transformation may only be beneficial if the values replaced by 0 were negative. The function $(x-t)p(x)$ is negative for $x < t$, so the bigger t is, the more attractive is $V(t,c,D)$ relative to $V(t,c,P)$. Therefore $V(t,c,D) > V(t,c,P)$ for t sufficiently big.

Similar reasoning applies to the fourth statement. $V(t,c,D)$ and $V(t,0,D)$ differ only by a weighting function and a constant c. The function $p(x)$ is between 0 and 1, and thus places relatively small weight on negative values of $(x-t)$, thus strongly diminishing their effect on the integral (while the effect of positive values is only slightly attenuated). In particular, notice that the limit of $V(t,c,D) - V(t,0,D)$ as $t \to 1$ is:

$$\int_0^1 (1-x)(1-p(x))a^D(x)dx - c > 0 \text{ for } c \text{ small.}$$

Observe, that if c satisfies inequality (5), then c also satisfies:

$$\int_0^1 (1-x)(1-p(x))a^D(x)dx - c \geq \int_0^1 (1-x)(1-p(x))a^D(x)dx - \int_0^1 xp(x)dx + \int_0^1 xa^D(x) =$$
$$= \int_0^1 p(x)(1-x)(1-a^D(x))dx > 0.$$

Therefore, as long as $c < \psi$, $V(t,0,D)$ is never chosen. In this case, types close to 1/2 chose $V(t,c,P)$, bigger types switch to $V(t,c,D)$ and extreme types switch to $V(t,0,P)$.

It is a bit more difficult to deal with a slightly bigger c, that does *not* satisfy (5). Then the middle type would choose $(0,D)$ instead of (c,P). However, there are still some types, who acquire information. More specifically, for $t_0 > \frac{1}{2}$ satisfying $\int_0^1 (x - t_0)a^D = 0$ the optimal choice is (c, D), as long as c is not too big. Indeed, as $E(x|s = 1) \leq E(x|s = 1, a^D(x) = 1$ and $E(x|a^D(x) = 1) \leq E(x|s = 1, a^D(x) = 1$, then at least for t_0 it must be that $V(c, D, t_0) + c > 0$. By continuity, there exists an upper bound ϕ such that if $c < \phi$ at least someone acquires information.

If $c > \phi$ and nobody acquires information, the equilibrium is trivial, as no significant information is transmitted through delegation.

Proof of Lemma 2

Recall the definition of the expected type (upon delegation):

$$\beta(x) = E(t|D, x) = \int_0^1 tg(t|D, x)dt.$$

With:

$$g(t|D, x) = \begin{cases} \frac{g(t)(1-p(x))}{A} & \text{for } t \in \Omega^0_{CD}, \\ \frac{g(t)p(x)}{A} & \text{for } t \in \Omega^1_{CD}, \\ \frac{g(t)}{A} & \text{for } t \in \Omega_{UD}, \end{cases}$$

where $A = \int_{\Omega^0_{CD}} g(t)dt + \int_{\Omega_{UD}} g(t)dt$, independent of x. To see that $\beta(\frac{1}{2}) = \frac{1}{2}$, observe that $g(t|D, \frac{1}{2})$ is a symmetric density function, so the expected value with respect to t is $\frac{1}{2}$. Moreover, notice that $\beta(x)$ is simply an affine transformation of $p(x)$ that preserves symmetry around a point $\left(\frac{1}{2}, \frac{1}{2}\right)$.

$$\beta(x) = \frac{1}{A}\left[p(x)\int_{\Omega^1_{CD}} tg(t)dt + (1 - p(x))\int_{\Omega^0_{CD}} tg(t)dt + \int_{\Omega_{UD}} tg(t)dt\right].$$

$$\beta(x) = \left[p(x)\frac{\left(\int_{\Omega^1_{CD}} tg(t)dt - \int_{\Omega^0_{CD}} tg(t)dt\right)}{A} + \frac{\int_{\Omega^0_{CD}} tg(t)dt + \int_{\Omega_{UD}} tg(t)dt}{A}\right].$$

Denote by $\tilde{\tau} := \frac{\int_{\Omega^0_{CD}} tg(t)dt + \int_{\Omega_{UD}} tg(t)dt}{A}$, that is, the expected type of a patient who chooses delegation. Since both sets $\Omega^0_{CD} \cup \Omega^1_{CD}$ and Ω_{UD} are symmetric around $t = \frac{1}{2}$, it is easy to show that $\frac{\left(\int_{\Omega^1_{CD}} tg(t)dt - \int_{\Omega^0_{CD}} tg(t)dt\right)}{A} = 1 - 2\tilde{\tau}$. Therefore:

$$\beta(x) = (1 - 2\tilde{\tau})p(x) + \tilde{\tau}$$

In particular, $\beta(x)$ is increasing, convex on $\left[0, \frac{1}{2}\right)$ and concave on $\left(\frac{1}{2}, 1\right]$ and symmetric around $\left(\frac{1}{2}, \frac{1}{2}\right)$ i.e. $\beta(x) = 1 - \beta(1 - x)$.

Proof of Theorem 1:

The proof is simple and follows directly from properties of $\beta(x)$ derived in the proof of Lemma 2 above. Recall that the doctor chooses $a^D(x) = 1$ if $x > \beta(x)$ and $a^D(x) = 0$ otherwise. From Lemma 2 we already know $\beta(x)$ is increasing and crosses line $id(x) = x$ at least in $x = \frac{1}{2}$. Since it's an affine transformation of $p(x)$ and $\lim_{x\to 0} \beta(x) > 0$ (which implies $\lim_{x\to 1} \beta(x) < 1$), then it crosses line $id(x) = x$ at most three times in $(0, 1)$. I will show that if $\beta'(x) > 1$ then $\beta(x) - x = 0$ has exactly three solutions in $(0, 1)$ and define

$$\underline{x} = \min\{x : \ \beta(x) = x\}. \tag{7}$$

Let us start with the limit. This is simple: observe that whatever x is, if delegation was observed, it must have come from a type $t \in \Omega_{CD} \cup \Omega_{ND}$. Then

$$\forall x \ E(t|D, x) \geq \min \Omega_{CD} \cup \Omega_{UD} = \underline{t} \ \Rightarrow \ \lim_{x\to 0} E(t|D, x) \geq \underline{t} > 0.$$

For the derivative, recall that:

$$\beta'(x) = p'(x)(1 - 2\tilde{\tau}).$$

$$\beta'(x) > 1 \Leftrightarrow p'(x) > \frac{1}{1 - 2\tilde{\tau}} =: \alpha(c). \tag{8}$$

Assume c is small, namely $c \leq \psi$, as defined in inequality (5). For such a c, no types choose to delegate conditionally. Then $\tilde{\tau} = E(t|t \in \Omega^0_{CD})$ and $\frac{1}{1-2\tilde{\tau}} < \frac{1}{1-2\tau}$, therefore as long as $p'(x) > \frac{1}{1-2\tau}$, the existence of $\underline{x} < \frac{1}{2}$ is guaranteed, regardless of c. For $c > \psi$, however, no useful upper bound exists.

Lemma 3. *If* $\int_0^{\frac{1}{2}} \left(\frac{1}{2} - x\right) p(x)dx \leq \phi(a) := \frac{(1-2a)(3-5a)-a\left(\sqrt{a^2+2(1-2a)(3-5a)}-a\right)}{16(3-5a)^2}$ *then inequality (6)holds.*

Proof. Denote $r = \left(\frac{1}{2} - \underline{x}\right)$. Recall that by definition of $\underline{x}$ in (7):

$$\underline{x} = p(\underline{x})(1 - 2\tilde{\tau}) + \tilde{\tau}.$$

Notice that $p(x)$ might be considered a cumulative distributive function for some continuous symmetric unimodal distribution Z. It is easy to determine that:

$$Var(Z) = \int_0^1 x^2 p'(x)dx - \frac{1}{4} \overset{parts}{=} \left(1 - \int_0^1 2xp(x) - \frac{1}{4}\right) = 2\left(\frac{3}{8} - \int_0^1 xp(x)dx\right).$$

On the other hand, since Z is symmetric, we can write:

$$Var(Z) = \int_0^1 \left(\frac{1}{2} - x\right)^2 p'(x)dx = 2\int_0^{\frac{1}{2}} \left(\frac{1}{2} - x\right)^2 p'(x)dx \overset{parts}{=} 4\int_0^{\frac{1}{2}} \left(\frac{1}{2} - x\right) p(x)dx.$$

Since p is concave on $\left(0, \frac{1}{2}\right)$, it can be bounded from above by piecewise linear function, using the property $p(\underline{x}) = \frac{x-\tilde{\tau}}{1-2\tilde{\tau}}$:

$$\int_0^{\frac{1}{2}} \left(\frac{1}{2} - x\right) p(x) dx = \int_0^{\underline{x}} \left(\frac{1}{2} - x\right) p(x) dx + \int_{\underline{x}}^{\frac{1}{2}} \left(\frac{1}{2} - x\right) p(x) dx \leq$$
$$\leq \int_0^{\underline{x}} \left(\frac{1}{2} - x\right) \frac{p(\underline{x})}{\underline{x}} x dx + \int_{\underline{x}}^{\frac{1}{2}} \left(\frac{1}{2} - x\right) \frac{x - \tilde{\tau}}{(1-2\tilde{\tau})} dx =$$
$$= \frac{1}{48(1-2\tilde{\tau})} \left(-8\tilde{\tau}\underline{x}^2 + 12\tilde{\tau}\underline{x} - 6\tilde{\tau} + 1\right) = \frac{1}{48(1-2\tilde{\tau})} \left(1 - 2\tilde{\tau}\left(4r^2 + 2r + 1\right)\right).$$

A sufficient condition for inequality (6) to hold is therefore:

$$\frac{1}{12(1-2\tilde{\tau})} \left(1 - 2\tilde{\tau}\left(4r^2 + 2r + 1\right)\right) \leq 2r^2$$

It is satisfied whenever $r \geq \frac{1}{4(3-5\tilde{\tau})} \left(\sqrt{\tilde{\tau}^2 + 2(1-2\tilde{\tau})(3-5\tilde{\tau})} - \tilde{\tau}\right)$ or, alternatively, if $Var(Z) < 4\epsilon(\tilde{\tau})$, where $\epsilon(\tilde{\tau}) = \frac{1}{12}\frac{(1-2\tilde{\tau}(4r^2+2r+1)}{(1-2\tilde{\tau})}$ evaluated at $r = \frac{1}{4(3-5\tilde{\tau})} \left(\sqrt{\tilde{\tau}^2 + 2(1-2\tilde{\tau})(3-5\tilde{\tau})} - \tilde{\tau}\right)$. After a bit of tiresome algebra, we get the required sufficient condition to be:

$$\int_0^{\frac{1}{2}} \left(\frac{1}{2} - x\right) p(x) \leq \frac{(1-2\tilde{\tau})(3-5\tilde{\tau}) - \tilde{\tau}\left(\sqrt{\tilde{\tau}^2 + 2(1-2\tilde{\tau})(3-5\tilde{\tau})} - \tilde{\tau}\right)}{16(3-5\tilde{\tau})} =: \epsilon(\tilde{\tau}).$$

Notice that $\tilde{\tau}$ is determined in equilibrium. However, if $\Omega_{UD} = \emptyset$ then $\tilde{\tau} = E(t|t \in \Omega^0_{CD}) < \tau$ and since ϵ is increasing, then the condition with τ instead of $\tilde{\tau}$ is stronger (and sufficient). To get some more intuition, check Figure 3 for a plot of an upper bound on $Var(Z)$ and notice that since that for an arbitrary unimodal distribution we only have $Var(Z) \leq \frac{1}{12}$, the bound is non-trivial and indeed necessary. □

References

Alonso, R., and N. Matouschek (2008): "Optimal delegation," *Review of Economic Studies*, 75(1), 259–293.

Arieli, I., and M. Mueller-Frank (2017): "Inferring beliefs from actions," *Games and Economic Behavior*, 102, 455–461.

Austen-Smith, D. (1994): "Strategic transmission of costly information," *Econometrica*, 62(4), 955–963.

Boldrin, M., and D. K. Levine (2005): "The economics of ideas and intellectual property.," *Proceedings of the National Academy of Sciences of the United States of America*, 102(4), 1252–1256.

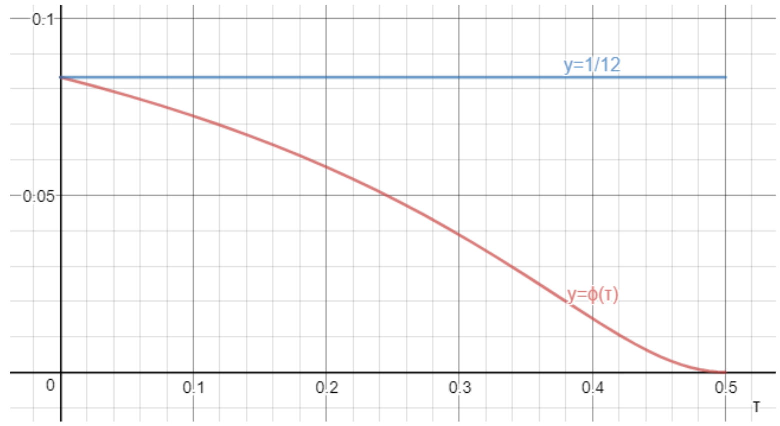

Figure 3: Upper bound on variance $\phi(\tau)$ (red) that guarantees the existence of an interesting equilibrium

Dessein, W. (2002): "Authority and Communication in Organizations," *Review of Economic Studies*, 69(4), 811–838.

Dewatripont, M., and J. Tirole (2005): "Modes of communication," *Journal of Political Economy*, 113(6), 1217–1238.

Ehses-Friedrich, C. (2011): "Expert Communication to an Informed Decision Maker," Discussion paper, Max Planck Institute of Economics.

Eso, P., and B. Szentes (2007): "The price of advice," *RAND Journal of Economics*, 38(4), 863–880.

Garfagnini, U., M. Ottaviani, and P. N. Sørensen (2014): "Accept or reject? An organizational perspective," *International Journal of Industrial Organization*, 34(1), 66–74.

Gentzkow, M., and E. Kamenica (2014): "Costly persuasion," *American Economic Review*, 104(5), 457–462.

Hedlund, J. (2015): "Persuasion with communication costs," *Games and Economic Behavior*, 92, 28–40.

Johnson, E. M., and M. M. Rehavi (2016): "Physicians Treating Physicians: Information and Incentives in Childbirth," *American Economic Journal: Economic Policy*, 8(1), 115–141.

Kartik, N. (2015): "Muddled Information," Mimeo.

Koszegi, B. (2004): "Emotional Agency: The Case of the Doctor-Patient Relationship," *Mimeo*.

——— (2006): "Emotional Agency," *The Quarterly Journal of Economics*, 1(121), 121–155.

LI, H., AND W. SUEN (2004): "Delegating decisions to experts," *Journal of Political Economy*, 112(S1), 311–335.

LUBENSKY, D., AND E. SCHMIDBAUER (2013): "Physician Overtreatment and Undertreatment with Partial Delegation," Discussion paper, Indiana University, Kelley School of Business.

POLANYI, M. (1966): *The Tacit Dimension*. The University of Chicago Press.

SPENCE, M. (1973): "Job market signaling," *The Quarterly Journal of Economics*, 87(3), 355–374.

XIE, B., D. M. DILTS, AND M. SHOR (2006): "The physician-patient relationship : The impact of patient-obtained medical information," *Health Economics*, 15(March), 813–833.