European University Institute

DEPARTMENT OF ECONOMICS

On Communication Frictions

Joanna Franaszek

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Florence, 18 May 2018

European University Institute
Department of Economics

On Communication Frictions

Joanna Franaszek

Thesis submitted for assessment with a view to obtaining the degree of Doctor of Economics of the European University Institute

Examining Board

Prof. Piero Gottardi, EUI, Supervisor
Prof. Andrea Mattozzi, EUI
Prof. Alfredo Di Tillio, Bocconi University
Prof. Salvatore Piccolo, University of Bergamo

Researcher declaration to accompany the submission of written work

I Joanna Magdalena Franaszek certify that I am the author of the work "On Communication Frictions" I have presented for examination for the PhD thesis at the European University Institute. I also certify that this is solely my own original work, other than where I have clearly indicated, in this declaration and in the thesis, that it is the work of others.

I warrant that I have obtained all the permissions required for using any material from other copyrighted publications.

I certify that this work complies with the *Code of Ethics in Academic Research* issued by the European University Institute (IUE 332/2/10 (CA 297).

The copyright of this work rests with its author. [quotation from it is permitted, provided that full acknowledgement is made.] This work may not be reproduced without my prior written consent. This authorisation does not, to the best of my knowledge, infringe the rights of any third party.

Signature and Date:

Joanna Franaszek

9/02/2018

To Bartek, who made this work possible,
To Marysia and Adaś, who made it meaningful.

Acknowledgements

It takes a village to raise a child, it takes a community to write a thesis. I would like to express my gratitude to people who supported me in this challenge. First and foremost, I must thank my supervisor Piero Gottardi for his invaluable advice and constant support during the time of struggle. I am grateful to Andrea Mattozzi and Andrea Galeotti for their comments and to all my colleagues at the European University Institute for making it an unforgettable experience.

Profound gratitude goes to my beloved husband, who helped me to maintain work–life balance with two kids and a PhD, and my parents, who always believed in my success. I owe all to you. Special thanks go to my friends: Kasia Malik-Liaci, Paweł Doligalski and Krzysztof Pytka for their logistic and emotional support.

Contents

Thesis summary **1**

1 Delegation as a signal: implicit communication with full cooperation **3**

1.1 Introduction 3
1.2 Model 7
1.3 Equilibrium choices 11
1.4 Summary 18
Appendix 19

2 When competence hurts: revelation of complex information **27**

2.1 Introduction 27
2.2 Setup 30
2.3 Equilibrium condition and notation 33
2.4 Receiver's expected gain in the game 35
2.5 Summary 40

Thesis summary

Communication is an activity that is crucial in many economic contexts. Whenever there is any asymmetry of information between two parties, there might arise a need to exchange information. Common as it is, communication is also complex and inherently imperfect. Information revelation may be mitigated by both language constraints and strategic incentives. I describe such difficulties in communication in two exemplary setups.

In ”Delegation as a signal: implicit communication with full cooperation” I examine a principal–agent model with fully aligned incentives to examine how inexpressible information can be implicitly communicated through observable choices. I assume there is two-sided private information; the principal knows his preference parameter, while the agent observes the decision-relevant state of the world. The principal may obtain a costly and imperfect signal about the state and either choose one of two possible actions or delegate the authority to the perfectly informed agent. The decision to delegate is more than just indecisiveness. In particular, an altruistic agent that observes delegation and the actual state realization can correctly anticipate the direction of the principal’s ”bias” in preferences and adjust the decision to better suit the principal’s needs. This phenomenon is a result of explicit correlation of truthful signals about the state of the world and implicit correlation of the signal and the type that arises in the equilibrium.

Chapter ”When competence hurts: revelation of complex information” examines a sender–receiver persuasion game. While the sender’s payoff depends only on the action chosen, the receiver’s payoff is affected by a state of the world, that is observed only by the sender. I assume that understanding the state of the world may require some competence from the receiver. Since noisy messages may be a result of both endogenous strategic obfuscation and exogenous constraints, they do not necessarily indicate ”bad news”. I show that being more competent could be hurtful for the receiver. Since competent receivers are more likely to correctly understand the unfavorable announcement, they would be less wary upon hearing noise. As a result, the seller is able to sustain an equilibrium in which a noisy message is a relatively good signal. With a less competent receiver the unique equilibrium is informative.

1 Delegation as a signal: implicit communication with full cooperation

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.

1.1 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 the delegating agent is 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 intermediate – not biased too much towards any o the actions. 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. Moreover, we assume that the utility is continuous, which implies that around the cutoff $x = t$ the patient is close to indifference. 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 belief $g(t)$.

Suppose the doctor ex-ante expects the patient types to be distributed uniformly, with half of the patients preferring a surgical 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 surgery whenever the decision is delegated. In fact, if the doctor believes the delegating patient to be quite well–informed, the bias might even overturn the ex-ante recommended action. In such a case, the doctor may suggest surgery only for lowest states of x, while some patients with health $x \in [\bar{x}, \frac{1}{2}]$ are prescribed drugs, even though their state of health is relatively poor. Since similar reasoning applies to relatively high types ($x > \frac{1}{2}$), the signaling–by– delegation may lead 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.[2] This might look surprising, but in fact is desirable from the patient's perspective – the delegation allows the patient to improve the final choice even more than making an informed decision.

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.

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

[2] As would be explained in the article, the implicit bias in doctor's beliefs arises in the model for *any* informative signal structure. However, only for some specific family of signals, the bias is strong enough to overturn the ex-ante decision and give rise to a non-monotonic action profile.

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 (2006), 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 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.[3]

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

[3] 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.

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.[4] 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.

1.2 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.[5] 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. 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. The parameter τ is a proxy of the concentration of $g(t)$ around the mean $Eg(t) = \frac{1}{2}$. Note that the doctor's problem of decision under delegation is more pronounced if τ is small i.e. there is a significant measure of patients with strong preferences towards one of the treatment.[6]

[4]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.

[5]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.

[6]Distributions highly concentrated around the mean would have large (i.e. close to 1/2) values of τ, while introducing some dispersion would decrease the parameter. Notice that from the doctor's perspective the extreme types $t < 1/4$ or, symmetrically, $t > 3/4$ are irrelevant, as they never delegate. Thus, the "dispersion" parameter τ is only calculated for non-extreme types.

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 1.1. The utility u has the following properties:

- $u_t(x,a)$ is continuous and (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 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 the state of 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.[7]

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

[7]While the assumption that t is inexpressible is quite strong, it serves as a useful benchmark and a departure from large family of models with perfect communication.

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, 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.[8] In particular, I shall assume that for any x the probability of signal $s = 1$ is $P(s = 1|x) = p(x)$ belongs to a very specific family of S-shaped signals.

Assumption 1.2. The probability of signal $s = 1$, denoted by $p(x)$, satisfies:

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 (see Figure 1.1). This assumption is crucial in establishing an equilibrium with doctor's non-monotone choice of action. Steepness of the signal structure ensures that the patient who chooses to acquire information is well–informed at least about the apparent direction of his health. While he lacks full information about the value of x, an informed patient acquires a quite precise signal about the state being higher or lower than the average. If given such precise information, the patient still prefers to delegate, the choice of delegation becomes a strong signal about the patient's type and allows the doctor to infer similarly strong beliefs about the possible range of t.[9]

[8] Note that such a recommendation is *not* cheap–talk.

[9] One might wonder whether the signal structure might be more complex, for example continuous. The answer is yes, but with serious restrictions. 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. Moreover, while only weak assumptions about informativeness of a signal are needed to obtain *some* bias in the doctor's belief, the bias effect is strong enough only under quite specific conditions. In particular, any signal structure that is to lead to the non-monotonic choice of action must be sufficiently informative about the direction of x.

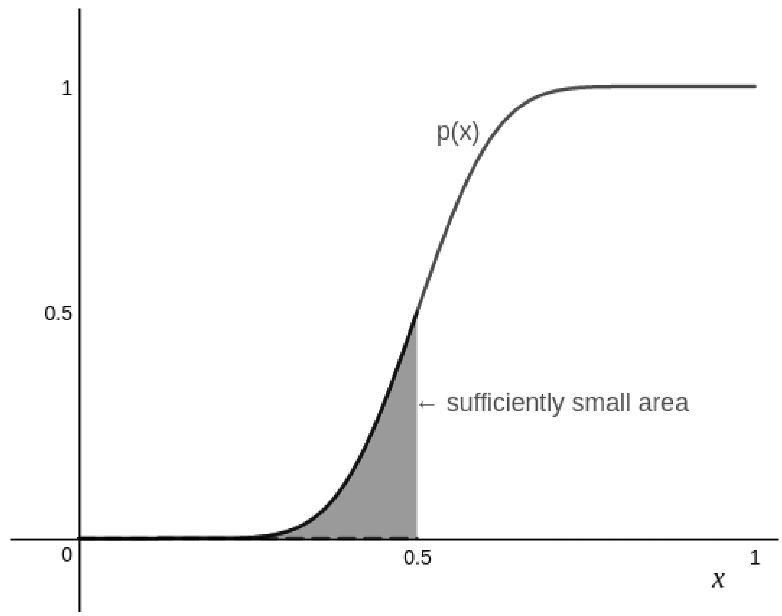


Figure 1.1: An S-shaped $p(x)$

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.

1.3 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.

1.3.1 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: $p_{lim}(x) = 1_{\{x \geq \frac{1}{2}\}}$. Notice that while it is not in fact an S-shaped distribution, it is a limit of S-shaped distributions. 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 there exists an equilibrium 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.[10] Therefore, if $x < \frac{1}{2}$ he

[10] 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

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.

The simple example in the limit case gives a flavor of the main result of the paper: if an interesting equilibrium with cues exists for the simplest signal structure $p_{lim}(x) = 1_{\{x \geq \frac{1}{2}\}}$, then it would also exist for some family of signals in the neighbourhood of p_{lim}, that I describe formally in assumption 1.2.

It must be noted that the equilibrium described above is unique in this setup. Observe that even absent the doctor, types $t \in [1/4, 3/4]$ would still find it optimal to acquire information (and a possibility to delegate would weakly enhance their incentives, whatever the doctor strategy is). Moreover, at least types $t = 3/4$ and $t = 1/4$ would delegate (as they are indifferent between retaining and delegating authority), so the set of informed-and-delegating types is non-empty. Since the signal depends on x only through $1_{\{x>1/2\}}$, $E(t|x) = E(t|s)$ is a function that takes at most two values $E(t|s = 1)$ or $E(t|s = 0)$. The doctor's optimal choice is $a^D(x) = 1_{x>1/2 \wedge x>E(t|s=1)} + 1_{x<1/2 \wedge x>E(t|s=0)}$. Suppose that $x < \frac{1}{2}$. Then $a^D(x) = 1_{x>x_0}$ (possibly for $x_0 \geq \frac{1}{2}$). In such a case, all types $t \in [1/4, \frac{x_0+\frac{1}{2}}{2}]$ would delegate, thus resulting in $E(t|D,x) = \frac{1}{8} + \frac{x_0+1/2}{4}$. Since it must also hold that $x_0 = E(t|D,x)$, then the only solution is $x_0 = 1/3$. Similar reasoning applies to $x > \frac{1}{2}$, resulting in a profile determined above.

1.3.2 Action choice

Let us now proceed with formal backward analysis for a general signal structure. Going backwards, we shall start with the 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

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.

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).

1.3.3 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[11]:

$$V(t, 0, D) \geq V(t, 0, P) \Leftrightarrow \int_0^1 (x-t)a^D(x)dx \geq \max\left(\frac{1}{2} - t, 0\right). \tag{1.1}$$

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

Claim 1.3. *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:

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

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

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:

$$\begin{cases} E(x|s, a^D = 0) \leq t & \text{for } s = 1, \\ E(x|s, a^D = 1) \geq t & \text{for } s = 0. \end{cases} \tag{1.2}$$

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

Claim 1.4. *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 1.5. *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).

1.3.4 Information acquisition

Take the strategies in the second period described in the previous subsection as given. 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:

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

Lemma 1.6. *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 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 two lemmas above are simply summarized in Figure 1.2. 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

Figure 1.2: The investment in information and delegation choice for $c < \psi$ (left) and $\psi \leq c \leq \phi$ (right).

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 1.5, 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 1.2.

If the cost of the signal is big enough, the types close to $t = \frac{1}{2}$ are hit by a "median patient curse".[12] 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 1.6, there are no other switches in the patient strategy than those in the Figure 1.2.

1.3.5 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 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:

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

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

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

Lemma 1.7. *Assume $g(t)$ 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 $\tilde{\tau}:=\frac{\int_{\Omega^0_{CD}}tg(t)dt+\int_{\Omega_{UD}}tg(t)dt}{A}$, 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.

Lemma 1.8. *If $p'(\frac{1}{2})>\frac{1}{1-2\tilde{\tau}}$ (with a sufficient condition $p'(\frac{1}{2})>\frac{1}{1-2\tau}$ 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 1.3. 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 1.8, we can claim the following, main result of the paper:

Theorem 1.9. *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"[13] steepness and perfectly flat tails. Lemma 1.8 and Theorem 1.9 explain how this extreme signal structure could be generalized and adapted to continuous signal functions.

[13]Formally, indefinite.

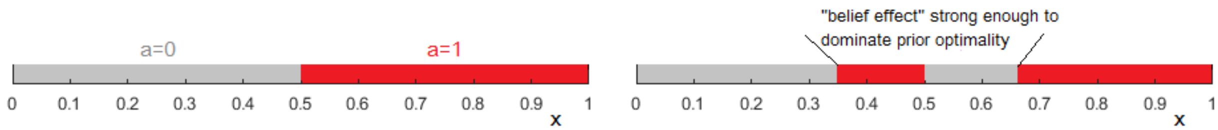


Figure 1.3: The action profile chosen by doctor in equilibrium if $p'(\frac{1}{2})$ is small (left) and large (right)

1.4 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 1.3: Inequality in (1.1) 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 1.4: Assume that a^D follows the putative pattern. Intuitively, $a^D = 1$ must be chosen more often for higher states, so:

$$E(x|s, a^D = 0) < E(x|s) < E(x|s, a^D = 1), \tag{1.3}$$

since

$$E(x|s) = E(x|s, a^D = 0) \cdot P(a^D = 0) + E(x|s, a^D = 1) \cdot P(a^D = 1). \tag{1.4}$$

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 (1.2) is satisfied. Similarly, for $t \searrow E(x|s)$ the second inequality of (1.2) 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 1.5: Assume t satisfies condition (1.2) 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 (1.2), 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 $t\in\Omega^1_{CD}$ then $1-t\in\Omega^0_{CD}$.

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 1.6: 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 1.5 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:

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

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:

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

$$V(t,0,D) \leq V(t,c,P) \Leftrightarrow c \leq \underbrace{\int_0^1 xp(x)dx - \frac{3}{8}}_{<0} + \left(\frac{1}{2} - \bar{x}\right)^2 =: \psi \tag{1.5}$$

The necessary condition being:

$$\psi = \left(\frac{1}{2} - \underline{x}\right)^2 - \frac{3}{8} + \int_0^1 xp(x)dx \geq 0 \tag{1.6}$$

I shall claim that whenever $\int_0^{\frac{1}{2}} \left(\frac{1}{2} - x\right) p(x)$ is sufficiently small, the inequality (1.6) 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 1.10.

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 (1.5), 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 (1.5). 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 1.7

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 Lemma 1.8:

The proof is simple and follows directly from properties of $\beta(x)$ derived in the proof of Lemma 1.7 above. Recall that the doctor chooses $a^D(x) = 1$ if $x > \beta(x)$ and $a^D(x) = 0$ otherwise. From Lemma 1.7 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

$$\underline{x} = \min\{x : \ \beta(x) = x\}. \tag{1.7}$$

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:

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

$$\beta'(x) > 1 \Leftrightarrow p'(x) > \frac{1}{1-2\tilde{\tau}} =: \alpha(c). \tag{1.8}$$

Assume c is small, namely $c \leq \psi$, as defined in inequality (1.5). 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 1.10. *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 (1.6)holds.*

Proof. Denote $r = \left(\frac{1}{2} - \underline{x}\right)$. Recall that by definition of $\underline{x}$ in (1.7):

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

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 $p(\underline{x}) = \frac{x-\tilde{\tau}}{1-2\tilde{\tau}}$:

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

A sufficient condition for inequality (1.6) 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 $r = \frac{1}{4(3-5\tilde{\tau})}\left(\sqrt{\tilde{\tau}^2 + 2(1-2\tilde{\tau})(3-5\tilde{\tau})} - \tilde{\tau}\right)$.

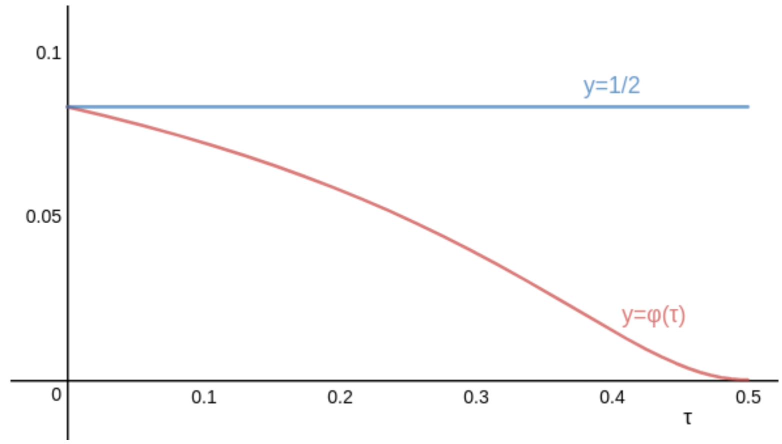


Figure 1.4: Upper bound on variance $\phi(\tau)$ (red) that guarantees the existence of an interesting equilibrium

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 1.4 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. $\square$

2 When competence hurts: revelation of complex information

Abstract

When the information might be complex and the information processing capacity of economic agents is uncertain, noisy messages do not necessarily indicate bad news. I exploit this intuition to examine a simple sender–receiver persuasion game in which the precise communication of the state of the world depends not only on sender's efforts but also on the complexity of the state and the receiver's competence. In this environment the sender–optimal equilibria maximize the amount of noise. The receiver faces a "competence curse" – a smarter type might end up with less information and lower payoff than a receiver with a somewhat smaller competence.

2.1 Introduction

It is hardly possible to imagine communication between two people that would allow for perfect exchange of any given information. Misunderstanding, misinterpretation or just imprecision might arise due to exogenous frictions, such as the sender's ability to formulate the message, the receiver's competence to absorb and correctly interpret the information content of the message or just the complexity of the matter discussed. Those competence frictions are an inherent feature of real-world communication and there is wide literature regarding such "language barriers" (a term coined by Blume and Board (2006)). However, there is also an endogenous source of frictions coming from possible divergence of interests between the two parties.

Miscommunication may be particularly bothersome if it leads to suboptimal decisions. Between 2006 and 2010, more than a million households in Poland, Croatia, Romania and other Eastern European countries took mortgage loans denominated in Swiss franc, to escape high borrowing costs in their home countries. As the franc had appreciated until 2011, and soared even further in 2015 (when the Swiss National Bank unpegged it from the euro), the Swiss–franc borrowers

I am grateful for valuable comments of Piero Gottardi, Andrea Mattozzi, Andrea Galeotti, Vitor Farinha Luz and Ran Eilat.

were left not only with monthly installments doubled, but also with mortgages worth more than the underlying properties. The dissatisfied borrowers complained about being misinformed, claiming in the European Court of Justice that the bank's presentation "was made in a biased manner, emphasizing the advantages (...), while failing to point out the potential risks or the likelihood of those risks materializing."[1] The Court accentuated that "a term under which the loan must be repaid (...) must be understood by the consumer both at the formal and grammatical level, and also in terms of its actual effects."[2]

The issue of customer's financial (il)literacy became central to the discussion on the unfortunate mortgage holders. Polish data indicates the Swiss–franc borrowers were on *average* relatively wealthy,[3] suggesting their financial literacy might have been relatively high. However, since the banks varied in their loan policies, there is a substantial concern that well-informed risk-lovers were pooled with some risk-averse victims of misinformation.

The mortgage example clearly shows that even when information transmitted between the parties must be truthful (by legal regulations), lack of congruence between the receiver and the sender might attenuate communication, if only the latter can manipulate the information content of his message. Anticipating this, the receiver would not only take the strategic incentives into consideration when interpreting the message, but might also find it worthwhile to hide his competence, in order to enhance the informativeness of the sender's message.

I show that in some situations a better–informed customer may in fact receive a worse service – in particular, a less informative message about the good he wishes to purchase. This is because lack of competence leads to skepticism and forces the seller to persuade the buyer with meaningful communication. If a client wishing to buy a car does not know much about vehicles, the seller has an incentive to persuade him by taking necessary effort in explaining the qualities of a given model. However, competent customers are better at understanding signals, which makes them less wary. This allows the seller to flood the buyer with technical details and overload him with information, that, even if unfavorable, would be harder to understand. As a result, the smarter client might end up worse off than just a bit less competent type.

Related literature The present article contributes to the growing literature on communication with limited information processing abilities. The main reference is a model by Dewatripont and Tirole (2005), henceforth DT, which inspired the current model. As in their setup, I examine a sender–receiver game, in which the former tries to persuade the latter to take some action. I modify the DT framework by adding another dimension(s) of uncertainty, which is the complexity

[1](European Court of Justice, 2017, par. 11)

[2]Ibid., par. 51.

[3]National Bank of Poland estimates franc borrowers to have 30% higher annual income and 2.5 higher level of liquid financial assets than the mortgage holders that took a loan in Polish zloty. See National Bank od Poland (2015)

of the state and the sender's message. Furthermore, I examine how the equilibrium change with the receiver's competence in understanding complex messages and what are his incentives to signal his abilities.

An idea that the receiver has some intellectual, time or attention constraints appeared in a famous model of rational inattention by Sims (2003). Glazer and Rubinstein (2004) derived optimal mechanisms of persuading a receiver that can understand only a single argument. Guembel and Rossetto (2009); Bucher-Koenen and Koenen (2015) define competence as the probability of a correct message, similarly to my approach. However, they examine a cheap-talk[4] communication, while I concentrate on truthful information revelation. A close reference is therefore Persson (2017), who builds on the DT framework to examine an issue of investment in communication with several senders competing for the attention of one receiver – or, in a similar manner – a monopolist sender communicating about several aspects of the good. In her setup, information overload arises endogenously as a result of the receiver's limited attention. If the prior is favorable – i.e. if without communication the receiver took the sender's preferred action – experts send irrelevant cues to prevent the receiver from discovering potentially unfavorable news. I describe a similar equilibrium (however, in a much simpler setup) to focus further on the effect on the receiver's competence. I discover that "smarter" types can end up with a worse outcome.

The result that competence can be harmful appeared also in Moreno de Barreda (2010); Ishida and Shimizu (2016); Rantakari (2016) cheap-talk models. They show that the receiver's access to an extra source of information – either after or before the communication takes place – decreases informativeness of the equilibrium messages. Li and Madarász (2008) – also within the cheap-talk framework – show that an extra information about the conflict of interest can decrease communication. However, the mechanism of cheap-talk games is quite different than in my model of information disclosure, where exaggeration is not allowed. Instead of lying, the sender would flood the more competent receiver with more noise. The models of Kessler (1998); Roesler and Szentes (2017) share a similar story, that is, being "too informed" may not be optimal for the receiver. Also, noiseless communication may not be optimal for welfare (see Fishman and Hagerty (1990); Goldstein and Leitner (2013); Blume, Board, and Kawamura (2007)).

Finally, this paper falls into broad literature of truthful, albeit not necessarily complete information transmission. The classic unraveling mechanism, as in Milgrom (1981); Grossman (1981) is disturbed in my model by the presence of uncertainty about the complexity of the state. The setup resembles the one in Shin (1994), who introduces uncertainty about the expert's information space. Similarly, I have uncertainty about the information complexity required to understand the state. When such complexity and the (lack of) competence in communication are introduced, "uninformative" messages may look favorable. This contrasts with the classic Milgrom (1981) result that no news is bad news.

[4]See Crawford and Sobel (1982)

The model also corresponds to a general class of Bayesian persuasion games, described in the seminal paper by Kamenica and Gentzkow (2009). As in Kamenica and Gentzkow (2009); Rayo and Segal (2010); Alonso and Câmara (2013) the sender can benefit from non-full disclosure. The main difference between my approach and Bayesian persuasion models is that I have no commitment on the sender's strategy, much like in Morgan and Stocken (2003); Dziuda (2011).

2.2 Setup

The model is an augmented version of DT setup, with an additional dimension of uncertainty.[5] There are two players, a sender and a receiver. The receiver is going to choose between a known status-quo that yields payoff (normalized to) 0 to both players, and some risky action A. Action A yields a certain payoff 1 for the Sender and an uncertain payoff that depends on the unknown state of the world ρ for the receiver. The payoff is either ρ_H in state H or ρ_L in state L, with $\rho_H > 0 > \rho_L$. The prior probability of a high state is $\alpha \in (0,1)$.

The sender has information about the state of the world $\omega \in \{H, L\}$ which he might communicate to the receiver. However, the information may be difficult to transmit. We can imagine e.g. some technical information that requires expertise to be understood. In particular, the state could be either simple to transmit, which will be denoted by complexity parameter $n = 1$ and happens with probability q or complex with $n = 2$ and prob. $1 - q$. In a simplest interpretation, n describes the number of pieces of information required to understand the state.[6] After observing the state realization and its complexity, the sender decides to send a simple or complex message to the receiver. While simple messages – if sufficient – could be understood by any receiver, complex messages require some competence. In particular, an announcement with complexity $m = 2$ could be only understood imperfectly – with probability x the receiver discovers the state, while with probability $1 - x$ the receiver regards the message as noise. This happens regardless of the underlying state complexity n, as both in a simple and in a complex state the message can be dumbfounding. I shall call x the receiver's competence and for now assume it is observable by both parties.[7]

The information communicated by the sender must be truthful, but could be noisy – in particular,

[5]The model is loosely related to the original DT setup, but closely related to the idea mentioned in footnote (32) of Dewatripont and Tirole (2005).

[6]An abstract representation of n allows for other interpretations as well. For example, if the sender is a seller of a good, $n = 1$ might indicate that the good's brand is (in)famous from marketing campaings and therefore the buyer knows the "state" (i.e. whether or not it is suitable for his needs) only after hearing the name of the brand, while $n = 2$ means that the brand some further exploration is needed.

[7]I interpret competence as e.g. financial literacy, similarly to Bucher-Koenen and Koenen (2015). Thus, more experienced financial traders simply have higher x. Another, very different idea was employed by Inderst and Ottaviani (2012), where financial literacy was associated with customers' level of strategic "sophistication". In other words, the financial novices were considered to be naive i.e. unaware of the existing conflict of interest.

the sender can exploit the receiver's (lack of) competence by issuing "too complex" message. I shall assume that if the complexity of the message is smaller than the complexity needed to understand the state (i.e. $m < n$), the message is perceived as noise.[8] But also, if a simple state is obfuscated by a complicated announcement, the receiver would only understand it with probability x. Intuitively, simple message about a simple state is always understood perfectly. It is crucial that the sender's announcement does not convey any signal about either the complexity of the state or of the message itself. In particular, if the receiver hears noise, he cannot tell whether it was because of mismatched complexities $(m < n)$ or his own small competence x. It is crucial to notice that complex states allow the sender to perfectly enforce noise (by choosing $m = 1$), while simple states can only be obfuscated imperfectly.

The timing of the model is as follows:

1. Nature chooses:
    a) state $\omega \in \{H, L\} \sim (\alpha, 1-\alpha)$
    b) complexity of the state $n \in \{1, 2\} \sim (q, 1-q)$ (known by S)
    c) competence of the receiver $x \in (0, 1)$;
2. The sender observes (ω, n) and chooses a truthful message of complexity $m \in \{1, 2\}$;
3. The receiver takes action $a \in \{\emptyset, A\}$.

It is crucial to notice that the sender chooses his message complexity m after learning the state of the world (ω, n), in other words, he does not commit to his strategy ex-ante. This is in-line with models of Morgan and Stocken (2003); Dziuda (2011) and in contrast to the important class of Bayesian persuasion models.

Figure 2.1 summarizes my assumptions regarding the information available to the receiver at each possible triple (ω, n, m). Recall that the source of noise in the message is imperceptible from the receiver's point of view.

There is a plethora of equilibria in the game. To limit myself to some more reasonable cases, I shall assume that any message is in principle cheap, but complex messages are a bit more costly to send. In particular, sending message of complexity $m = 1$ costs 0, while sending message of complexity $m = 2$ costs $c > 0$. Intuitively, c is positive, but very small,[9] just enough to induce a choice between strategies that would otherwise be equally preferred by the sender. The assumption reduces the set of equilibria to those where choosing complex messages over simple indeed has some rationale.

[8]It is still truthful, though, in a sense that the sender is not allowed to lie about the state. In case of $m = 1$, $n = 2$, the sender chooses not to provide the receiver with sufficient information, but is unable to claim that the state is high when it it low.

[9]The minimal requirement is that $c \ll \frac{1}{2}$.

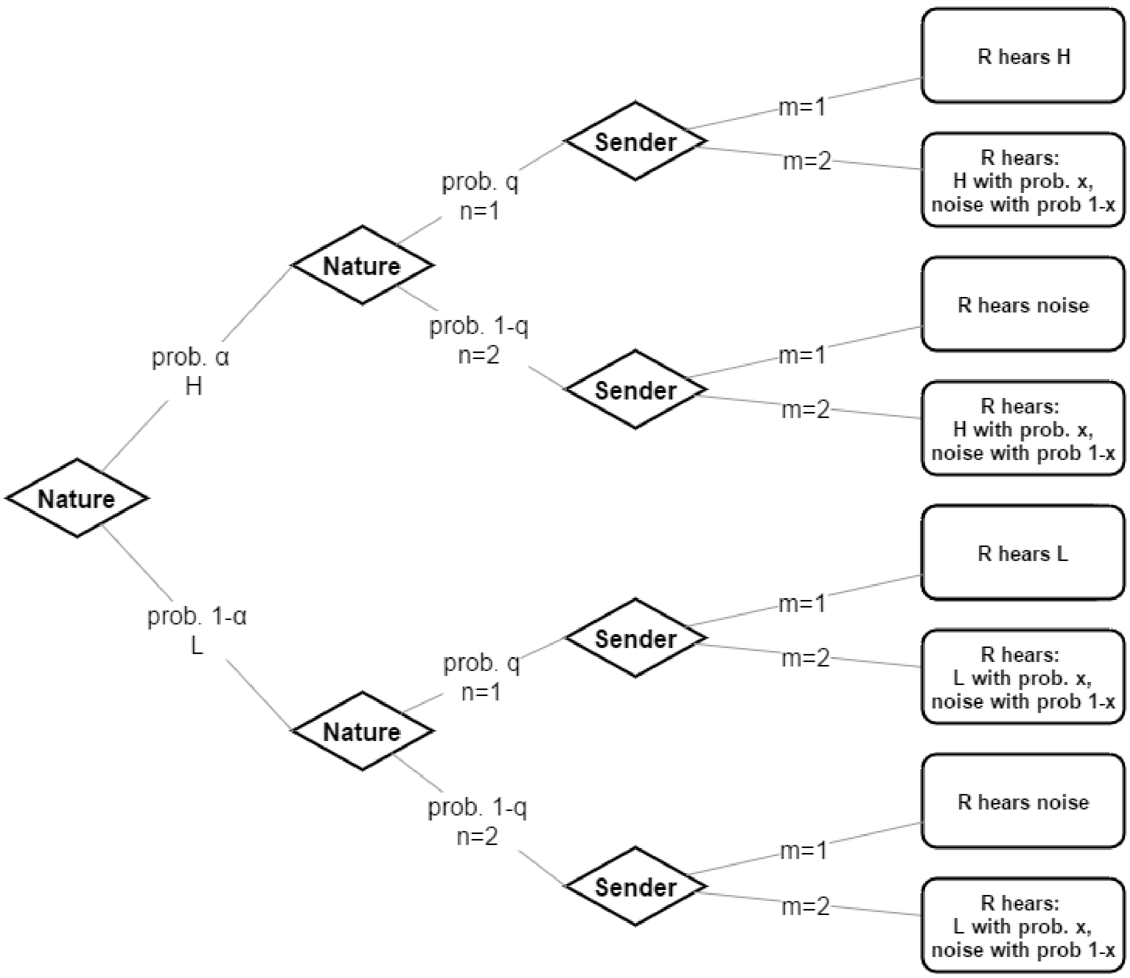


Figure 2.1: The scheme of information transmission in the game.

2.3 Equilibrium condition and notation

In order to establish the properties of a perfect Bayesian equilibrium of the game Γ_x with known competence x, one need to specify:

- The sender's message function $\sigma^S : \{H, L\} \times \{1, 2\} \to \{1, 2\}$, that for every pair (ω, n) (i.e. state and competence needed to understand it) observed by S, defines his message complexity $m \in \{1, 2\}$. For notational convenience, I shall describe the sender's strategy as a quadruple: $((H, 1), (H, 2), (L, 1), (L, 2)) \mapsto (m_{H1}, m_{H2}, m_{L1}, m_{L2})$.
- The receiver's beliefs μ in each of his information sets (which correspond to messages he understands) $\{H, L, noise\}$.
- the receiver's action function, i.e. $a : \{H, L, noise\} \to \{\emptyset, A\}$.

Lemma 2.1. *(trivial) In any perfect Bayesian equilibrium the receiver's beliefs upon hearing messages H and L are trivial, i.e. $\mu(\omega = H|H) = \mu(\omega = L|L) = 1$. Therefore, his optimal actions are $a(H) = A$, $a(L) = \emptyset$. The only nontrivial action is:*

$$a(noise) = \begin{cases} A & \text{if } \mu(H|noise)r_H + \mu(L|noise)r_L \geq 0 \\ \emptyset & \text{otherwise.} \end{cases}$$

Proof. Since the revelation is truthful, upon hearing a non-noisy message the Receiver is certain about the state, which trivially determines his optimal actions. The action upon hearing noise is a result of Bayes rule. Since for any $x < 1$ and any sender's strategy receiving noisy message has positive probability, then for every strategy profile $(m_{H1}, m_{H2}, m_{L1}, m_{L2})$ the receiver can calculate the posterior probability of the state being H when noise was heard. Then, in any perfect Bayesian equilibrium the receiver would take action A upon hearing noise only if according to his Bayesian beliefs: $\mu(H|noise)r_H + \mu(L|noise)r_L \geq 0$. □

Lemma 2.2. *In any perfect Bayesian equilibrium with costly complex messages, the sender's strategy choice must have $m_{H1} = 1$ and $m_{L2} = 1$.*

Proof. Recall that the Sender chooses his actions already knowing (ω, n). In state $(H, 1)$ the choice of sending simple message, that would surely be understood by the receiver and induce action A, strictly dominates the choice of complex message, that not only is more costly, but also leaves a possibility of misunderstanding. Similarly, in state $(L, 2)$ the complex message is more costly and in no way can it induce a better action than the simple (here meaning: noisy) message. □

Define $\gamma_\alpha = \frac{\alpha}{1-\alpha}$ as the ratio of prior probabilities of states H and L. Note that γ_α as a function of α is strictly increasing and define $\gamma_r = \frac{-r_L}{r_H}$, as the benchmark prior γ_α, that makes the receiver indifferent between action A and the null action. Note that γ_α and γ_r summarize the uncertainty about the state along two different dimensions – while γ_α is directly related to the probability distribution, γ_r depends only on the payoffs. Denote the ratio of the two parameters by $\gamma = \frac{\gamma_\alpha}{\gamma_r} = -\frac{r_H \alpha}{r_L(1-\alpha)}$. Large values of γ mean that the prior is strong (or the gain in the high state is high), while small values of γ indicate that the prior is weak (or the punishment in the low state is substantial). The parameter γ summarizes the gains or losses from uncertainty in the model. Observe that a priori – absent any communication – the receiver would take action A only if $\gamma > 1$.

Theorem 2.3. *Suppose $0 < c \leq \min(x, 1-x)$. There are four types of perfect Bayesian equilibria of the game with known competence x:*

1. *The informative equilibrium in pure strategies, that exists whenever $\gamma < \frac{1}{1-x}$:*
   a) *The sender's strategy is $((H,1),(H,2),(L,1),(L,2)) \mapsto (1,2,1,1)$,.*
   b) *The receiver's beliefs are $\mu(H|noise) = \frac{\alpha(1-x)}{1-\alpha x}$, $\mu(L|noise) = \frac{1-\alpha}{1-\alpha x}$, $\mu(H|H) = \mu(L|L) = 1$,*
   c) *The receiver's actions are $a(H) = A$, $a(L) = \emptyset$ and $a(noise) = \emptyset$;*
2. *The noisy equilibrium in pure strategies, that exists whenever $\gamma > \frac{1-qx}{1-q}$*
   a) *The sender's strategy is strategy $((H,1),(H,2),(L,1),(L,2)) \mapsto (1,1,2,1)$,*
   b) *The receiver's beliefs are $\mu(H|noise) = \frac{\alpha(1-q)}{\alpha(1-q)+(1-\alpha)(1-qx)}$, $\mu(L|noise) = \frac{(1-\alpha)(1-qx)}{\alpha(1-q)+(1-\alpha)(1-qx)}$, $\mu(H|H) = \mu(L|L) = 1$,*
   c) *The receiver's actions are $a(H) = A$, $a(L) = \emptyset$ and $a(noise) = A$;*
3. *The mixed semi-informative equilibrium for $\gamma \in \left(\frac{1}{1-x}, \frac{1-qx}{(1-q)(1-x}\right)$*
   a) *The sender's strategy is $(1,2,(1-r,r),1)$ with $r = \frac{(1-q)(\gamma(1-x)-1)}{q(1-x)}$,*
   b) *The receiver's actions are $a(H) = A$, $a(L) = \emptyset$ and $a(noise) = (A, \emptyset)$ with probabilities $(1-\frac{c}{x}, \frac{c}{x})$;*
4. *The mixed semi-noisy equilibrium for $\gamma \in \left(\frac{1-qx}{1-q}, \frac{1-qx}{(1-q)(1-x}\right)$*
   a) *The sender's strategy is $(1,(p,1-p),2,1)$ with $p = \frac{(1-\gamma(1-q)(1-x)-qx)}{\gamma x(1-q)}$,*
   b) *The receiver's actions are $a(H) = A$, $a(L) = \emptyset$ and $a(noise) = (A, \emptyset)$ with probabilities $(\frac{c}{1-x}, 1-\frac{c}{1-x})$;*

Corollary 2.4. *If $c > \max(x, 1-x)$ then for $\gamma < 1$ the only equilibrium is $\{(1,1,1,1), \emptyset\}$. For $\gamma > 1$ and $x < c < (1-x)$ the equilibrium is a mixed–strategy profile $\{(1,1,(1-r,r),1),(b,1-b)\}$ with possible corner solution $r = 1$ and $b = 1$ that is sustained for $\gamma > \frac{1-qx}{1-q}$. If $\gamma > 1$ and $1-x < c < x$ the only equilibrium is a mixed–strategy profile $\{(1,(p,1-p),1,1),(b,1-b)\}$ with possible $p = 0$ and $b = 0$ whenever $\gamma > \frac{1}{1-x}$.*[10]

[10]The case $c > \max(x, 1-x)$ is ruled out, since the minimal requirement is $c < 1/2$.

Proofs of the theorem and corollary are moved to the Appendix. In line with the intuition about small, almost negligible cost, I shall concentrate in the rest of the paper solely on the case of $c < \min(x, 1-x)$. The equilibria in such a case are pictured in the left panel of Figure 2.2. Notice that even with costly messages, multiple equilibria still persists. For a given pair (γ, x) either the equilibrium is unique, or there are three equilibria (two pure, one mixed or one pure and two mixed).

To make the analysis of the competence effect more explicit, I shall impose an equilibrium selection rule, to enable comparative statics between unique outcomes. Since in my model it is the sender, who possesses the information and therefore has more "initiative", I shall follow the approach of Bayesian persuasion models and concentrate on sender-optimal equilibria.[11] It is not difficult to verify that in fact, the sender-best equilibrium is the one maximizing the amount of noise.

Theorem 2.5. *The sender-best equilibrium is:*

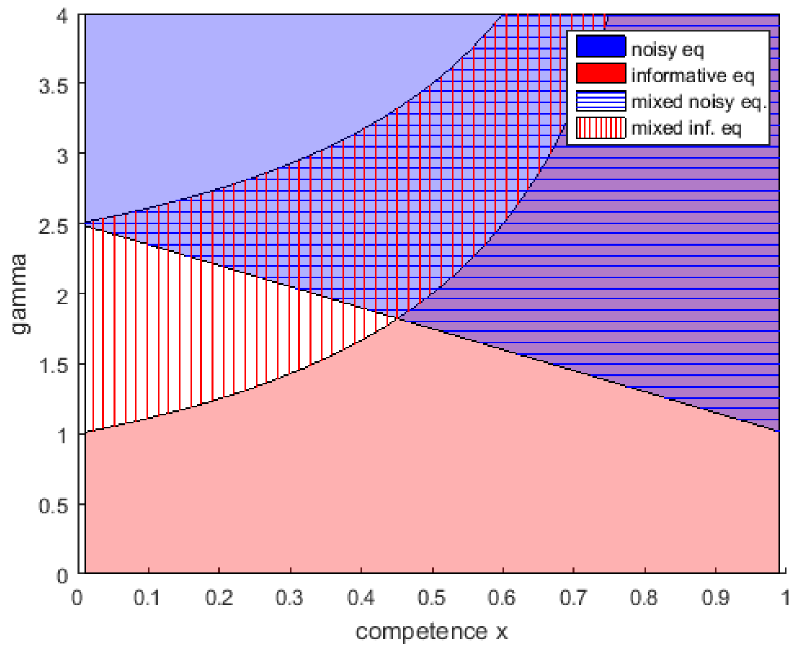
- $\{(1,1,2,1), a(noise) = A\}$ *for* $\gamma > \frac{1-qx}{1-q}$,
- $\{(1,2,1,1), a(noise) = \emptyset\}$ *for* $\gamma < \min\left(\frac{1-qx}{1-q}, \frac{1}{1-x}\right)$,
- $\{(1,2,(1-r,r),1), a(noise) = (1-\frac{c}{x}, \frac{c}{x})\}$ *for* $\frac{1}{1-x} < \gamma < \frac{1-qx}{1-q}$.

The noisy equilibrium $\{(1,1,2,1), A\}$, is preferred by the sender, whenever it can be supported by the receiver's beliefs. If the noisy equilibrium fails to exist, the existing informative or mixed semi-informative equilibrium is unique, therefore, it is sender-best. Sender-best equilibria are pictured in the right panel of Figure 2.2.

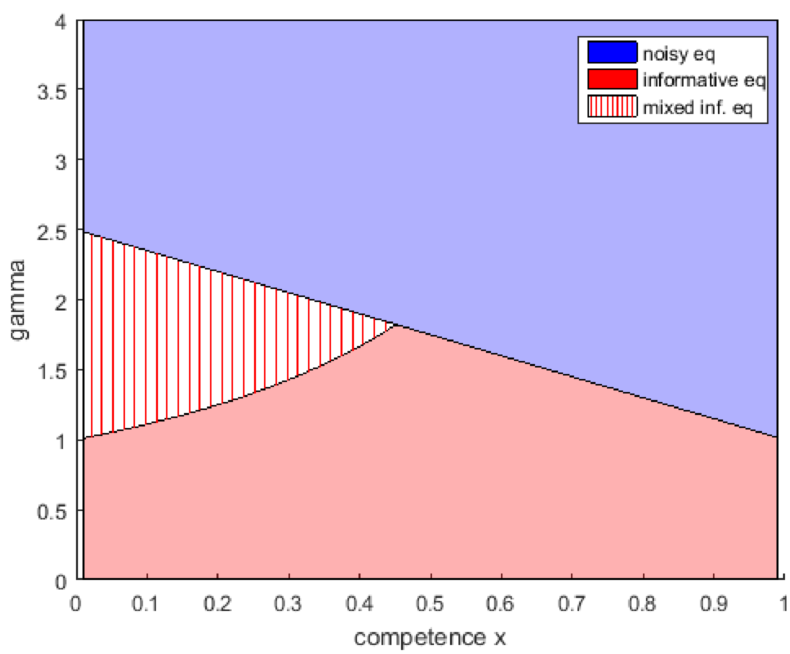
2.4 Receiver's expected gain in the game

We may now analyze the receiver's gain from the game and how it depends on his language competence x. Assume $1 < \gamma < \frac{1}{1-q}$, which is the most interesting range, as depending on x, there is a possibility of up to three different (sender-best) equilibria. Notice that for a given γ, a higher competence may be a burden for the receiver, as it might result in a noisy equilibrium $\{(1,1,2,1), A\}$, while for somewhat lower values of x the unique equilibrium is either informative $\{(1,2,1,1), \emptyset\}$ or semi-informative $\left\{(1,2,(1-r,r),1), \left(\frac{c}{1-x}, 1-\frac{c}{1-x}\right)\right\}$.

[11] Glazer and Rubinstein (2012) propose a different approach, where the receiver commits to a "persuasion codex". With such an assumption, the selected equilibria would be those optimal for the receiver.



(a) All equilibria



(b) Sender-best equilibria

Figure 2.2: Pure (solid fill) and mixed (pattern fill) equilibria in the game for a given γ and x.

Effect of competence

To see why higher types face a "competence curse", observe that for more competent receivers, noise might be a favorable message. Assume that the sender chooses a strategy $(1,1,2,1)$ and the receiver anticipates it. In a state $(L,1)$, the receiver hears a signal that he correctly understands as L with probability x. High values of x are a sign of competence, therefore the receiver is "relatively good" in identifying low state correctly with certainty. As a result he believes noise to be less likely to arise in the low state. As $m(L|noise)$ is low, $m(H|noise)$ must be relatively high, thus, the noise becomes a signal of a good state. Competence becomes a curse – because of the receiver's good understanding of low states, the sender is able to "sell" the noisy message as a favorable signal and maintain the equilibrium in which very little information is transmitted.

More formally, let us analyze the receiver's expected payoff:

$$\begin{aligned} E^R_{\text{info eq}}\rho(x) &= (\alpha q + \alpha(1-q)x)\,\rho_H \\ E^R_{\text{noisy eq}}\rho(x) &= \alpha\rho_H + \rho_L(1-\alpha)(1-qx) \end{aligned} \tag{2.1}$$

For a given equilibrium profile the receiver always benefits larger competence – as both $E^R_{\text{info eq}}\rho(x)$ and $E^R_{\text{noisy eq}}\rho(x)$ are increasing in x. However, this is not the case when a change in x would induce a change in the equilibrium profile.

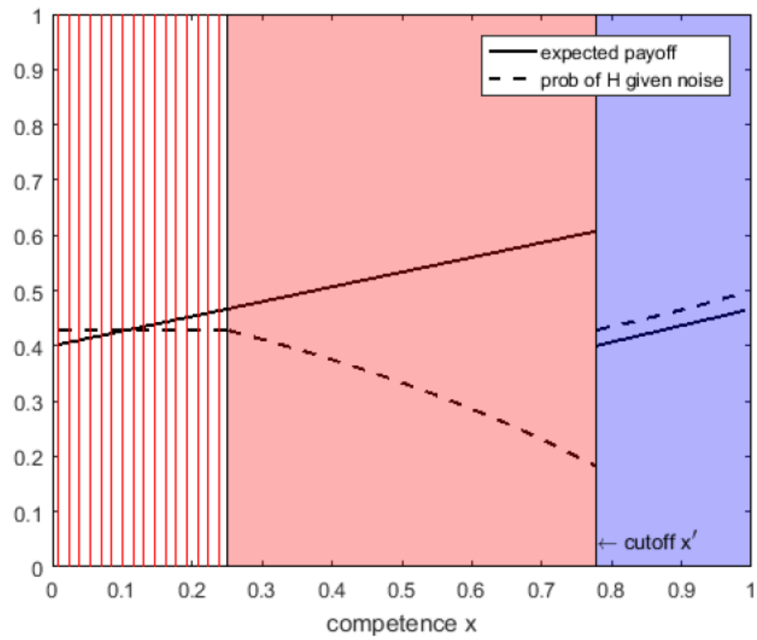
Assume (x,γ) result in a noisy equilibrium and consider a downward change in x. As competence decreases, $m(H|noise)$ – which is an increasing function of x in the noisy equilibrium – also plummets, up to a point where it is no longer profitable for the receiver to choose A upon hearing noise and he would rather take the $\emptyset$ action instead. The sender is then forced to switch to an informative strategy $(1,2,1,1)$ and a new equilibrium arises. It must be noted that a switch from a noisy to informative equilibrium – as x decreases – not only increases informativeness, but brings a discontinuous jump in receiver's utility (see Figure 2.3).

Take $x' = \frac{1}{q}(1-\gamma(1-q))$ and let us analyze the expected gain/loss for an equilibrium switch when x decreasing around a neighborhood of x':

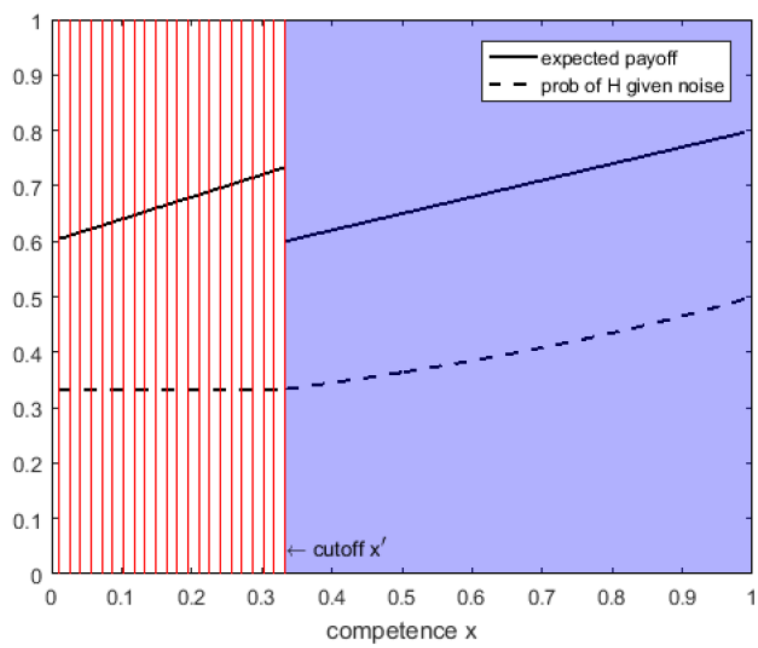
$$Er(x \searrow x') - Er(x \nearrow x') = -\alpha\rho_H(1-q)x' < 0$$

A marginal change in x around the threshold bring a discrete decrease in utility. Therefore, for a small ϵ, two receiver types $x'+\epsilon$ and $x'-\epsilon$ not only end up in different equilibria, but also the more competent receiver is strictly worse off. If it was possible, he would rather decrease his competence to $x'-\epsilon$ to induce an informative equilibrium than remain more competent, but less informed.

Similar reasoning applies to a switch from a noisy to mixed semi-informative equilibrium, whenever



(a) $\rho_H = 4/3,\ \rho_L = -1,\ q = 0.6,\ \alpha = 0.5$



(b) $\rho_H = 2,\ \rho_L = -1,\ q = 0.6,\ \alpha = 0.5$

Figure 2.3: Discrete change in the receiver's expected utility and $m(H|noise)$ when the equilibrium changes from the informative (red) or semi-informative (red stripes) to noisy (blue) profile.

γ is so high that the informative equilibrium no longer exists. Observe that by a definition of a mixed equilibrium, the receiver is indifferent between his two choices conditioned on noise, therefore:

$$E^{R}_{\text{mixed}(p,r,\beta)}\rho = \alpha\rho_H\left(q + (1-q)x\right),$$

which is of the same functional form as $E^{R}_{\text{info eq}}\rho$. Therefore, also a switch from a noisy to a semi-informative mixed equilibrium brings a discontinuous loss in the receiver's utility and the receiver type close to the threshold would prefer to compromise some competence in order to get a better outcome.

However, when the equilibrium is switched from an informative to mixed semi-informative, the change in utility is continuous and the utility is increasing in x. In fact, since the receiver's utility in both equilibria share the same functional form, the semi-informative equilibrium is a natural "alternative" to an informative equilibrium, whenever the latter cannot be sustained. Therefore, the receiver of type $x \in [0, x']$ has no incentive to reduce his competence, even if it was possible.

Change in a prior signal

Similar reasoning applies to changes in γ. Just like with competence, having higher initial prior does not necessarily benefit the receiver. In particular, if the receiver faces an increase in γ, he might end up in worse equilibrium. This is quite intuitive, as more favorable prior makes the receiver more likely to choose A, thus decreasing the sender's incentives to transmit information.

The prior information represented by γ and the communication competence x are substitutes. It would be interesting – but beyond the scope of this paper – to examine a model in which the two types of communication skills are substantially different; while one dimension represents the stock of knowledge, the other describes the ability to absorb new knowledge. In reality, those two dimensions of informational literacy are distinct skills.

Private information about competence

I have shown that the receiver might face a "competence curse" – in particular, if his competence is so high that it induces an uninformative equilibrium, the receiver might be worse off that with somewhat lower x. However, reducing x is hardly possible.

Assume that competence becomes the receiver's private information. To simplify, let us consider a case in which competence may be either x_L with probability π or x_H with probability $1-\pi$ and denote the equilibrium probability of type i choosing A upon hearing noise as b_i. To make things interesting, assume $x_L < \bar{x} < x_H$, where $\bar{x}$ satisfies $\frac{1}{1-\bar{x}} = \frac{1-q\bar{x}}{1-q}$[12] and the probabilities π

[12] $\bar{x}$ is a crossing point of the indifference curves that define equilibrium conditions

or $1-\pi$ are sufficiently separated from 0 and 1 – so that the two–types case does not trivially collapse to the one–type setup.

It can be shown (see: Appendix) that even though there are multiple equilibria in the setting, as long the noisy profile $\{(1,1,2,1), b_L = 1, b_H = 1\}$ can be sustained – that is for $\gamma > \frac{1-qx_L}{1-q}$ – it remains the sender–best equilibrium. If $\gamma < \frac{1}{1-x_L}$ the unique equilibrium is the informative one in which neither receiver type chooses A unless he is certain about the state. The only interesting case is therefore $\frac{1}{1-x_L} < \gamma < \frac{1-qx_L}{1-q}$. Indeed, in this range the high type might benefit from the uncertainty. The only (therefore, sender-best) equilibrium is $\{(1,2,(1-r,r),1), b_L = \frac{c}{\pi(1-xH)}, b_H = 0\}$ with $r \in [0, \bar{r}(x_L)]$.[13] Notice that the high type is not only better off than without uncertainty about x but also his outcome is higher than the low type's payoff. If the sender attempts to persuade the low type, he must send at least a semi-informative message. More competent type "freerides" and can now enjoy the more favorable outcome. The low type, on the other hand, enjoys the same outcome as if he played single-handedly.

This outcome would persist if the receiver could send a cheap-talk message. Notice that the high type would always want to send the same message as the low type, as it is in his best interest to be pooled. If instead the types would be able to credibly certify their types at no cost, the low type is able to separate, but has no strong incentives to do so – in fact, he is indifferent between being certified or not.

2.5 Summary

In this paper, I examine a persuasion game in which the state of the world might be difficult to transmit. The sender perfectly observes the state of the world and its complexity, i.e. how difficult it is to understand the state. The sender then chooses a simple or a complex message, with the latter bearing a small cost. His goal is to persuade the receiver to take some action A that yields the latter an uncertain payoff. As an alternative, the receiver can take an outside option $\emptyset$ with a payoff 0.

I show that when there is uncertainty about the complexity of information, noise is no longer perceived purely as "bad news". This is because noise might come from two different sources: exogenous complexity of information required for successful communication or endogenous sender's incentive to obfuscate the unfavorable state.

I concentrate on sender-best equilibria and show that there are three types of equilibrium profiles. If the prior γ is high, the receiver is willing to choose action A upon hearing noise and the sender can sustain his mostly preferred noisy equilibrium in which little information is transmitted. If

[13] More specifically, $\bar{r}(x_L) = \frac{q - x_L - qx_L(1-x_L)}{q(1-xL)}$.

γ is small, the receiver is more wary and the sender's best option is to send as much information as possible and refrain from issuing extra noise. Depending on (γ, x), this leads to either an informative or a semi-informative equilibrium. The surprising result, however, is that for a given γ more competent receiver types might end up in the worse, noisy equilibrium than somewhat smaller types, who are guaranteed to end up in the informative outcome. The competence becomes a curse.

To understand the result, suppose the sender tries to 'sell' noise as a good signal and issue intentionally complex messages. Since information revelation must be truthful, the announcements are sometimes correctly understood – more often, if the receiver is more competent. As a result, upon hearing noise the high type would attach less likelihood to the state being low than high. Thus, noise becomes a favorable message and the sender has no incentive to transmit any information. This equilibrium cannot be sustained for a less competent receiver, precisely because of his little understanding of the complex messages. The incompetent type is warier and unwilling to choose A upon hearing noise. Therefore, the sender has no choice, but to persuade him with an informative announcement.

In a comparative statics exercise I show that the utility loss associated with an equilibrium change is discrete and negative – in other words, the smart receiver would have a strong incentive to "play dumb". While in the standard setup, this is not possible, I also examine a game in which the sender is uncertain about the receiver's competence, which might be either high or low. I show that for a relevant range of γ the competent receiver strictly benefits from extra uncertainty, as he now ends up in the unique semi-informative outcome. The low type's outcome is the same, so he has no strict incentive to disturb the pooling equilibrium.

Appendix

Proof of theorem 2.3

Proof. I shall analyze which strategy profiles might arise in an equilibrium. By Lemma 2.2, there are only four feasible strategies of the sender: $(1,1,1,1)$, $(1,2,1,1)$,$(1,1,2,1)$, $(1,2,2,1)$. Notice also that in any perfect Bayesian equilibrium the receiver's beliefs must be consistent with the sender's strategy. Let us assume $c < \min(x, 1-x)$, which means there is at least some incentive to invest in costly message. Examine four cases:

1. The sender uses strategy $(1,1,1,1)$. Such strategy is consistent with the receiver's beliefs $\mu(H|noise) = \alpha$ and $\mu(L|noise) = 1-\alpha$. Assume that the receiver chooses $a(noise) = A$. The sender might benefit from deviating to $(1,1,2,1)$, generating noise with positive probability Assume the receiver takes an action $\emptyset$ when hearing noise. The sender can benefit from deviating to $(1,2,1,1)$, i.e. more informative message. Thus strategy $(1,1,1,1)$ is never optimal.

2. The sender uses strategy $(1,2,1,1)$. This strategy is consistent with the receiver's beliefs $\mu(H|noise) = \frac{\alpha(1-x)}{1-\alpha x}$ and $\mu(L|noise) = \frac{1-\alpha}{1-\alpha x}$. Assume the receiver takes A when hearing noise. Then the sender has an incentive to deviate to a more noisy message $(1,2,2,1)$. In the other case, when the receiver takes $\emptyset$ upon hearing noise, there is no incentive to deviate. The appropriate beliefs imply $\gamma < \frac{1}{(1-x)}$ and in such a case the strategies $\{(1,2,1,1),\ a(noise) = \emptyset\}$ constitute an equilibrium.

3. The sender uses strategy $(1,1,2,1)$. Upon hearing noise the receiver would take action A if $\gamma \geq \frac{1-qx}{1-q}$ and $\emptyset$ otherwise. In the latter case, i.e. $a(noise) = \emptyset$, the sender has an incentive to deviate from costly $(1,1,2,1)$ to less costly $(1,1,1,1)$. If $a(noise) = A$, there is no incentive to deviate and the profile $\{(1,1,2,1), a(noise) = A\}$ constitutes an equilibrium when $\gamma \geq \frac{1-qx}{1-q}$.

4. The sender uses strategy $(1,2,2,1)$. Upon hearing noise the receiver would take action A if $\gamma \geq \frac{1-qx}{(1-q)(1-x)}$ and $\emptyset$ otherwise. If $a(noise) = A$ the sender has an incentive to deviate from more costly $(1,2,2,1)$ to less costly $(1,1,2,1)$. In the second case, when $a(noise) = \emptyset$, the sender has an incentive to deviate from costly $(1,2,2,1)$ to less costly $(1,2,1,1)$. Thus, $(1,2,2,1)$ is not used in any equilibrium.

Notice that the analysis above could be also performed taking purely interim point of view, i.e. analyzing just the actual choice in critical states $(H,2)$ and $(L,1)$. This approach would be used to examine mixed strategies. Since the choice is made after the state is realized, the choices of m_{H2} and m_{L1} are interdependent only through the beliefs they induce in the equilibrium. The mixing could be an arbitrary $(1,(p,1-p),(1-r,r),1)$. In any mixed equilibrium in which at least one of p, r is interior, the receiver must be indifferent between choosing A and $\emptyset$, therefore p, r must satisfy:

$$\gamma = \frac{qr(1-x)+(1-q)}{(1-q)(1-x+px)}. \tag{2.2}$$

The receiver's response is $(b, 1-b)$ where $b = P(a(noise) = A)$.

Assume the receiver plays according to a strategy $(b, 1-b)$ with $b \in (0,1)$. Consider the state $(H, 2)$ and the sender's choice of $(p, 1-p)$ that costs $c(1-p)$. Notice that the sender's payoff is linear in p.

$$E(\text{payoff in}(H,2)) = (-x(1-b)+c)p + \alpha(1-q)(x+(1-x)b-c) \tag{2.3}$$

If $b = 1 - \frac{c}{x}$ then the sender's choice of p could be arbitrary, as the payoff is constant in p. If, $b < 1 - \frac{c}{x}$ then the sender finds it optimal to choose $p = 0$ and if $b > 1 - \frac{c}{x}$ then the optimal choice is $p = 1$.

A similar reasoning applies to changes in r, when the state is $(L, 1)$ The strategy $(1-r, r)$ costs cr and any incremental change in r results in a change in utility:

$$E(\text{payoff in}(L,1)) = ((1-x)b - c)r \tag{2.4}$$

Notice that generically for a given pair (c, x) it cannot simultaneously hold that $b = 1 - \frac{c}{x}$ and $b = \frac{c}{1-x}$ as long as $c \neq x(1-x)$. Therefore, at most one of (2.3) and (2.4) can be independent of p or r and allow for an interior choice of the parameter. Therefore mixing would be performed only in one of the critical states $(H, 2)$ and $(L, 1)$. In the other state, the incentives would drive the receiver to choose a corner solution from a set $\{0, 1\}$. This is quite clear if we observe that the sender's decision is indeed a linear programming problem.

The first type of mixed equilibrium is of the form $\{(1,(p,1-p),2,1),(b_1,1-b_1)\}$ with $b_1 = 1 - \frac{c}{x}$ and exists whenever $\frac{1-qx}{1-q} \leq \gamma \leq \frac{1-qx}{(1-q)(1-x)}$. Notice that for small c, the probability of the receiver taking action upon hearing noise is close to 1, therefore this equilibrium is relatively noisy We shall call it a semi-noisy mixed equilibrium.

The second type of mixed equilibrium is of the form $\{(1,2,(1-r,r),1),(b_2,1-b_2)\}$ with $b_2 = \frac{c}{1-x}$ and exists whenever $\frac{1}{(1-x)} \leq \gamma \leq \frac{1-qx}{(1-q)(1-x)}$. Whenever c is small, b_2 is close to zero. Therefore this equilibrium would be labeled as a semi-informative mixed equilibrium.

In any mixed equilibrium, the condition (2.2) must be satisfied, thus the mixed equilibria can be sustained only within some subset of the (x, γ)-space.

In the unlikely case of $c = x(1-x)$, the equilibrium is $\{(1,(p,1-p),(1-r,r),1),(x,1-x)\}$. In this equilibrium the probability of the receiver accepting the action A is exactly equal to his competence.

□

Proof of corollary 2.4

Proof. The result in a corollary follows naturally from the proof above. For $c > 1 - x$ the sender does never have an incentive to choose any r but 0, as is clear from his payoff characterization in (2.4). A strategy profile $(1, (p, 1-p), 1, 1)$ could be supported as long as γ satisfies (2.2) for some $p \in [0, 1\,]$, which implies $1 \leq \gamma \leq \frac{1}{(1-x+px)}$. For $\gamma < 1$ the equilibrium is $\{(1,2,1,1), A\}$ and for $\gamma > 1/(1-x)$ it must be $\{(1,1,1,1), \emptyset\}$. For $c > x$ the sender with incentives like in (2.3) must choose $p = 1$. A mixed strategy $(1, 1, (1-r, r), 1)$ is feasible as long as $1 \leq \gamma \leq \frac{1-qx}{1-q}$. For $\gamma < 1$ the equilibrium is $\{(1,1,1,1), \emptyset\}$ and for $\gamma > (1-qx)/(1-q)$ it must be $\{(1,1,2,1), A\}$. □

Proof of theorem 2.5

Proof. This result is quite intuitive, but I will prove it formally Notice that the sender's payoff from an arbitrary pure or mixed strategy of the general form $\{(1, (p, 1-p), (1-r, r), 1)\ , (b, 1-b)\}$ is:

$$\begin{aligned} Eu^S(\text{eq. profile}) = \alpha q + \alpha(1-q)[(1-p)(x+(1-x)b) + pb] + \\ + (1-\alpha)qr(1-x)b + (1-\alpha)(1-q)b - c(\alpha(1-q)(1-p) + r(1-\alpha)q). \end{aligned}$$

The mixed strategy payoff is quite easy to derive. For mixed equilibria, recall that by the definition of equilibrium b, the payoff must be independent of p in a semi-noisy equilibrium and of r in the semi-informative equilibrium.

$$\begin{aligned} Eu^S(\text{info eq.}) &= \alpha q + \alpha(1-q)(x-c), \\ Eu^S(\text{noisy eq.}) &= \alpha + (1-\alpha)((1-qx) - cq), \\ Eu^S(\text{semi-info eq.}) &= \alpha q + \alpha(1-q)[x + (1-x)b - c] + (1-\alpha)(1-q)b, \\ Eu^S(\text{semi-noisy eq.}) &= \alpha q + \alpha(1-q)[x + (1-x)b - c] + (1-\alpha)[q((1-x)b - c) + (1-q)b]. \end{aligned}$$

It is clear that $Eu^S(\text{noisy eq.}) > Eu^S(\text{info eq.})$ as $x - c < x + c < 1$ for $c < \min(x, 1-x)$. Notice also that the mixed profiles are increasing in b, therefore:

$$\begin{aligned} Eu^S(\text{semi-noisy eq.}) < \alpha q + \alpha(1-q)(1-c) + (1-\alpha)(1-qx-qc) < Eu^S(\text{noisy eq}) \\ Eu^S(\text{semi-info eq.}) < \alpha q + \alpha(1-q)(1-c) + (1-\alpha)(1-q) < Eu^S(\text{noisy eq}) \end{aligned}$$

Therefore Eu^S(noisy eq.) dominates all other payoffs. $\square$

Equilibria in the game with two receiver types

As an additional comment, I shall describe the equilibria in the game of one sender playing against a receiver of uncertain competence that is either x_L with probability π or x_H with probability $1-\pi$.

The receiver of type i chooses a strategy $(b_i, 1-b_i)$ with $b_i \in [0, 1]$. The sender plays a profile $(1,(p,1-p),(1-r,r),1)$ that also incorporates pure strategies.

The sender chooses his actions after learning the state, so again, we will consider his choices in the crucial states $(H,2)$ and $(L,1)$. In the high and complex state, he would prefer to send a complex message than a simple if:

$$\pi x_L(1-b_L) + (1-\pi)x_H(1-b_H) \geq c \tag{2.5}$$

In $(L,1)$ the complex message is preferred as long as:

$$\pi(1-x_L)b_L + (1-\pi)(1-x_H)b_H \geq c \tag{2.6}$$

Denote $f(p,r,x) = \frac{qr(1-x_i)+1-q}{(1-q)(1-x_i+px_i)}$. The receiver chooses $b_i = 0$ if $\gamma < f(p,r,x_i)$, $b_i = 1$ if $\gamma > f(p,r,x_i)$ and $b_i \in (0,1)$ if $\gamma = f(p,r,x_i)$. Notice that $f(p,r,x)$ is generically not constant, so typically at most one x_i may satisfy $f(p,r,x_i) = \gamma$, and thus have $b_i \in (0,1)$. This leaves us with eight possible cases:

- $b_L = 0, b_H = 0$. The sender's strategy must be $(1,2,1,1)$ and $\gamma < \frac{1}{1-x_H}$;
- $b_L = 0, b_H \in (0,1)$. Then $b_H = \frac{c}{(1-\pi)(1-x_H)}$ and the sender responds with $(1,2,(1-r,r),1)$ which can be sustained if $\gamma = \frac{qr(1-x_H)+1-q}{(1-q)(1-x_H)}$. But $f(p,r,x)$ increasing, so it can't be that $\gamma \geq f(p,r,x_L)$, that is required for $b_L = 0$. Such an equilibrium does not exist.
- $b_L = 0, b_H = 1$. Then $(1,2,2,1)$, but we can't have $\gamma \leq \frac{1-qx_H}{(1-q)(1-x_H)}$ and $\gamma \geq \frac{1-qx_H}{(1-q)(1-x_H)}$, as $f(p,r,x)$ increasing. Again, not feasible.
- $b_L \in (0,1), b_H = 0$. Then $b_L = \frac{C}{\pi(1-x_H)}$, the sender responds with $(1,2,(1-r,r),1)$ and $\gamma = \frac{qr(1-x_L)+(1-q)}{(1-q)(1-x_L)} < \frac{qr(1-x_H)+(1-q)}{(1-q)(1-x_H)}$.
- $b_L \in (0,1), b_H = 1$. Then $b_L = 1 - \frac{c}{\pi x_L}$, the sender responds with $(1,(p,1-p),2,1)$ and $\gamma = \frac{1-qx_L}{(1-q)(1-x_L+px_L)} > \frac{1-qx_H}{(1-q)(1-x_H+px_H)}$, which might be sustained only if $p > 1-q$;
- $b_L = 1, b_H = 0$. Then the sender responds with $(1,2,2,1)$ and now it's possible that $\frac{1-qx_H}{(1-q)(1-x_H)} > \gamma > \frac{1-qx_L}{(1-q)(1-x_L)}$.

- $b_L = 1, b_H \in (0,1)$. Then $b_H = 1 - \frac{c}{(1-\pi)x_H}$ and $(1,\ (p,\ 1-p), 2, 1) \frac{1-qx_H}{(1-q)(1-x_H+px_H)} = \gamma > \frac{1-qx_L}{(1-q)(1-x_L+px_L)}$ and must be $p < 1-q$;
- $b_L = 1, b_H = 1$. The sender chooses $(1,\ 1, 2,\ 1)$ and $\gamma > \frac{1-qx_L}{(1-q)} > \frac{1-qx_H}{(1-q)}$;

For a given γ, a few equilibria might coexist. However, as is clear from the proof of theorem 2.5 above, whenever the noisy equilibrium exists, it dominated (from the sender's point of view) all other possibilities. This leaves us with only three sender-best equilibria:

- $\{(1,1,2,1),\ b_L = 1,\ b_H = 1\}$ whenever $\gamma > \frac{1-qx_L}{(1-q)}$;
- $\{(1,2,\ (1-r,r),\ 1),\ b_L = \frac{c}{\pi(1-x_H)},\ b_H = 0\}$ whenever $\frac{1}{1-x_L} < \gamma < \frac{1-qx_L}{(1-q)}$ with $r = \frac{(1-q)(\gamma(1-x)-1)}{q(1-x)}$;
- $\{(1,2,1,1),\ b_L = 0,\ b_H = 0\}$ whenever $\gamma < \frac{1}{1-x_L}$;

Notice that those equilibria are the same *as if* the sender has played only against a low-type receiver. The low-type receiver obtains exactly the same outcome as in a game without the presence of a high type. However, x_H's payoff is substantially different. By being pooled with the low type in an informative (or semi-informative) equilibrium, the competent re- ceiver is able to avoid being lured into a noisy equilibrium.

Bibliography

Alonso, R., and O. Câmara (2013): "Persuading skeptics and reaffirming believers," Discussion paper, University of Southern California, Marshall School of Business.

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.

Blume, A., and O. J. Board (2006): "Language barriers," *Econometrica*, 12(13), 16–17.

Blume, A., O. J. Board, and K. Kawamura (2007): "Noisy talk," *Theoretical Economics*, 2, 395–440.

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.

Bucher-Koenen, T., and J. Koenen (2015): "Do Seemingly Smarter Consumers Get Better Advice?," Discussion paper.

Crawford, V. P., and J. Sobel (1982): "Strategic information transmission," *Econometrica*, 50(6), 1431–1451.

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

Dewatripont, M. (2006): "Costly Communication and Incentives," *Journal of the European Economic Association*, 4(2/3), 253–268.

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

Dziuda, W. (2011): "Strategic argumentation," *Journal of Economic Theory*, 146(4), 1362–1397.

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.

European Court of Justice (2017): "Ruxandra Paula Andriciuc and others v Banca Românească SA," Judgement of 20 September 2017, C-186/16, EU:C:2017:703.

Fishman, M. J., and K. M. Hagerty (1990): "The optimal amount of discretion to allow in disclosure," *The Quarterly Journal of Economics*, 105(2), 427–444.

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.

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

——— (2012): "A Model of Persuasion with Boundedly Rational Agents," *Journal of Political Economy*, 120(6), 1057–1082.

Goldstein, I., and Y. Leitner (2013): "Stress Tests and Information Disclosure," Discussion paper.

Grossman, S. J. (1981): "The Informational Role of Warranties and Private Disclosure about Product Quality," *The Journal of Law and Economics*, 24(3), 461.

Guembel, A., and S. Rossetto (2009): "Reputational cheap talk with misunderstanding," *Games and Economic Behavior*, 67, 736–744.

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

Inderst, R., and M. Ottaviani (2012): "Financial advice," *Journal of Economic Literature*, 50(2), 494–512.

Ishida, J., and T. Shimizu (2016): "Cheap talk with an informed receiver," *Economic Theory Bulletin*, 4(1), 61–72.

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.

Kamenica, E., and M. Gentzkow (2009): "Bayesian Persuasion," *American Economic Review*, 101(October), 2590–2615.

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

Kessler, A. S. (1998): "The Value of Ignorance," *The RAND Journal of Economics*, 29(2), 339–354.

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.

Li, M., and K. Madarász (2008): "When mandatory disclosure hurts: Expert advice and conflicting interests," *Journal of Economic Theory*, 139(1), 47–74.

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

Milgrom, P. (1981): "Good news and bad news : representation theorems and applications," *The Bell Journal of Economics*, 12(2), 380–391.

Moreno de Barreda, I. (2010): "Cheap talk with two-sided private information," Discussion paper, London School of Economics.

Morgan, J., and P. C. Stocken (2003): "An analysis of stock recommendations," *RAND Journal of Economics*, 34(1), 183–203.

National Bank od Poland (2015): "Zasobność gospodarstw domowych w Polsce," .

Persson, P. (2017): "Attention Manipulation and Information Overload," *NBER Working Papers Series*, 23823.

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

Rantakari, H. (2016): "Soliciting Advice: Active versus Passive Principals," *Journal of Law, Economics, and Organization*, 32(4), 719–761.

Rayo, L., and I. Segal (2010): "Optimal Information Disclosure," *Journal of Political Economy*, 118(5), 949–987.

Roesler, A. K., and B. Szentes (2017): "Buyer-optimal learning and monopoly pricing," *American Economic Review*, 107(7), 2072–2080.

Shin, H. S. (1994): "The Burden of Proof in a Game of Persuasion," *Journal of Economic Theory*, 64(1), 253–264.

Sims, C. A. (2003): "Implications of rational inattention," *Journal of Monetary Economics*, 50(3), 665–690.

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.