University of Cyprus
Department of Economics

Working Paper 11-2018

Noisy persuasion

Elias Tsakas and Nikolas Tsakas

Department of Economics, University of Cyprus, P.O. Box 20537, 1678 Nicosia, Cyprus
Tel.: +357-22893700, Fax: +357-22895028, Web site: http://www.ucy.ac.cy/econ/en

Noisy persuasion*

Elias Tsakas[†] Nikolas Tsakas[‡]

Maastricht University *University of Cyprus*

First draft: March, 2017

This draft: August, 2018

Abstract

We study the effect of noise due to exogenous information distortions in the context of Bayesian persuasion. In particular, we ask whether more noise is always harmful for the information designer (viz., the sender). We show that in general this is not the case. That is, more noise is often beneficial for the sender. However, when we compare noisy channels with "similar basic structures", more noise cannot be beneficial for the sender. We apply our theory to applications from the literatures on voting and cognitive biases.

Keywords: Bayesian persuasion; data distortions; optimal signal; garbling.

JEL codes: C72, D72, D82, D83, K40, M31.

1. Introduction

Information distortions is one of the most common and widely-studied phenomena in many areas within economics. The interest in the subject stems primarily from the fact that the noise induced by such distortions often leads to inefficiencies. In this paper we study the effect of information distortions in the context of the recently surging literature on Bayesian persuasion.

Persuasion games are sender-receiver games with commitment (Rayo and Segal, 2010; Kamenica and Gentzkow, 2011). In particular, an information designer (viz., the female sender) chooses an

*We are indebted to Matthew Gentzkow and Stefan Terstiege for valuable comments and discussions. Elias thanks the Department of Economics at UC Davis for its hospitality while working on this project.

[†]Department of Economics (AE1), Maastricht University, P.O. Box 616, 6200 MD, Maastricht, The Netherlands; Homepage: http://www.elias-tsakas.com/home.html; E-mail: e.tsakas@maastrichtuniversity.nl

[‡]Department of Economics, University of Cyprus, P.O. Box 20537, 1678, Nicosia, Cyprus; Homepage: https://sites.google.com/site/nikolastsakas/; E-mail: tsakas.nikolaos@ucy.ac.cy

experiment (viz., a signal) which is commonly known; the decision-maker (viz., the male receiver) observes an outcome of the experiment (viz., a message) and subsequently takes an action that affects both agents. Now, the caveat introduced in this paper is that the data that is observed by the receiver is often different from the actual realization of the experiment. Such distortions are typically attributed to errors in gathering, processing and transmitting information.

For instance, consider the problem of voters' persuasion by a politician who designs a political experiment in order to convince them to approve a proposal in an election (Alonso and Câmara, 2017). In this case, distortions occur when voters misinterpret the politician's messages, either due to their own attention limitations, or due to mistakes by the politician's campaign in the implementation of the experiment, or due to (intentional or unintentional) misreporting by the media. Then, we naturally ask: *Which is the effect of noise on the politician's (optimal) choice of an experiment, and consequently on her expected utility in equilibrium?*

The obvious restriction that noise poses to the sender is that it restricts the set of signals from which she can effectively choose. For instance, assume that the receiver observes the actual message only with probability $1 - \varepsilon$ and every other message with small positive probability. Then, clearly there are signals that the sender cannot choose, e.g., the perfectly informative signal that reveals the true state with certainty is not feasible anymore. Hence, the sender maximizes her expected utility over a restricted set, and therefore her expected utility in equilibrium will decrease, implying that noise is (weakly) harmful.[1] However, the previous argument applies only when we compare the noiseless case with a noisy one. Then, we ask whether the argument extends to cases where we compare any two types data distortions, with one being noisier than the other. Formally, *is the sender's expected utility in equilibrium increasing with respect to the (Blackwell) informativeness of the noisy channel that describes the respective distortions?*

Surprisingly, the answer is in general negative. Indeed, there are pairs of noisy channels, one being a garbling of the other (Blackwell, 1951, 1953), such that the the sender's expected utility under the garbled channel (viz., the more noisy one) is strictly larger than the expected utility under the original channel (viz., the less noisy of the two). Thus, *more noise can be beneficial for the sender.*

The underlying idea is that the set of signals that can be effectively chosen by the sender does not always shrink monotonically with respect to the Blackwell informativeness ordering. This is because garbling does not just increase the "amount of noise", but also often changes the "noise structure". In our previous example, assume that there are two possible sources of distortions, e.g., the media

[1] This observation provides an interesting comparison with Blume et al. (2007), who show that noise can be beneficial in standard cheap talk games.

misreport the outcome of the political experiment with some probability, and at the same time the firm that is hired by the campaign to run the experiment is incompetent, thus making mistakes in the implementation of the political experiment. The politician cannot control the media, but can nevertheless replace the firm with one that does not make such errors. The question then becomes, whether it would be at her best interest to remove one of the two sources of distortions by indeed replacing the firm. Our results suggest that this is not necessarily the case, i.e., she may prefer to maintain the firm that confuses the voters, as such confusion (combined with media misreporting) gives her access to experiments that are not feasible when mistakes are exclusively due to media misreporting. In other words, *more noise can be beneficial for the sender when we combine different sources of data distortions*.

Our previous analysis indicates that in order to understand the effect of noise on the sender's expected utility, we must first disentangle the effects of garbling on the "amount of noise" from the effects on the "structure of noise". In particular, we ask whether noise is always harmful when we compare channels of similar structure. To do so, we first identify two basic classes of noisy channels, henceforth referred as *canonical* and *partitional* channels, respectively. A channel is canonical if the error probability is relatively small, and moreover the probability of confusing message s with message t is equal to the probability of confusing t with s, e.g., this is the case when the message space is endowed with a distance function describing the closeness between any two messages. Canonical channels are interesting because they often appear in applications where data distortions are present, especially when such distortions are due to mistakes in data gathering (viz., measurement errors), data processing (viz., errors in storage/retrieval of data) and data transmission (viz., communication errors). On the other hand, a channel is partitional when messages are partitioned in equivalence classes, such that two messages within the same class are completely indistinguishable from one another. The reason why partitional channels are appealing is twofold: on the one hand, they are also often used in applications, especially in cases where noise is due to mistakes in data processing (viz., errors due to the coarseness of the language); on the other hand, as we show, partitional channels constitute the basic building block for nearly every noisy channel.

Our main result (Theorem 2) shows that, within each of our two basic classes, monotonicity of the sender's expected utility with respect to the channel's informativeness is restored. In other words, if two canonical (resp., partitional) channels are comparable in Blackwell's sense, then the sender's expected utility in equilibrium is larger under the more informative channel, i.e., *more noise is always harmful*. With the previous machinery and results at hand, we study applications of persuasion games with noise.

First, we focus on the problem of persuading voters that we have already introduced above

(Section 7.1). If the politician cares only about winning the election, it is well-known that absent of any distortions, she will always target approval by simple majority. However, in the presence of distortions, this is not always the case, i.e., she may adopt a more aggressive political experiment, targeting approval by unanimity. Surprisingly, her optimal experiment under noise depends, not only on the underlying noisy channel, but also on the source of said distortions. That is, if noise is due to mistakes in the campaign (in which case, the same effective message is drawn for all voters) she still targets majority, whereas if noise is due to mistakes in the voters' understanding (in which case, an independent message is drawn for each voter) she will target approval by unanimity (Propositions 3 and 4). The same intuition applies to all noisy persuasion games with multiple receivers.

Second, we focus on cognitive biases, with particular attention to *conservatism bias* (Section 7.2). Accordingly, people appear to anchor their beliefs to their priors and update information only partially; a phenomenon which is typically attributed to heuristic treatments. Nevertheless, following recent literature in psychology (Costello and Watts, 2014; Hilbert, 2012), we show that conservatism bias naturally emerges in a simple persuasion game with noise. In particular, combining two partitional structures leads to two types of messages: perfectly informative and perfectly uninformative ones. Hence, although in equilibrium the sender chooses the same optimal signal as in the noiseless case, the receiver will only update his belief if he observes one of the informative messages, and will stick to his prior otherwise. We confidently conjecture that other well-known cognitive biases can be explained within our framework of noisy Bayesian persuasion.

Our work lies on the intersection of two streams of literature: Bayesian persuasion and information distortions.

While there are several earlier (e.g., Glazer and Rubinstein, 2004; Milgrom and Roberts, 1986) as well as contemporary (e.g., Rayo and Segal, 2010) influential papers on persuasion, we view Kamenica and Gentzkow (2011) as the natural predecessor of our work. This choice is primarily based on the fact that their model has been the benchmark setting for most recent papers in this literature. Since Kamenica and Gentzkow's (2011) paper appeared, the persuasion literature has developed in several different directions. Alonso and Câmara (2016) have extended the benchmark model to one with heterogeneous priors. Alonso and Câmara (2017), Arieli and Babichenko (2016), Bergemann and Morris (2016a,b), Taneva (2014) and Wang (2013) have studied Bayesian persuasion with multiple receivers, while Laclau and Renou (2017) have simultaneously considered multiple receivers and heterogeneous priors. Bizzotto et al. (2017), Brocas and Carillo (2007) and Piermont (2016) study dynamic versions of the persuasion game. Perez-Richet (2014) and Hedlund (2017) consider a persuasion game with a privately informed sender, whereas Kolotilin et al. (2018) study persuasion with a privately informed receiver. None of the aforementioned papers considers data distortions.

The literature on information distortions has mostly focused on cheap talk games. The closest predecessor to our paper within this literature is by Blume et al. (2007), who introduce noisy communication to a standard cheap talk game á la Crawford and Sobel (1982) in an analogous way to our variant of Kamenica and Gentzkow's (2011) persuasion game. They are primarily interested in aggregate welfare, whereas our analysis focuses on the sender's expected utility and discusses separately the effect of noise on the receiver's expected utility. Noisy communication devices in cheap talk games have also been studied by Blume (2012), Blume and Board (2014), Goltsman et al. (2009) and Guembel and Rossetto (2009). Other related papers include Hernández and von Stengel (2014) on Nash codes, Koessler (2001) on consensus via communication and Landeras and Pérez de Villareal (2005) on screening. Sobel (2013) provides an overview of related experimental designs. We should mention the seminal example of Myerson (1991) that provided inspiration for many of the aforementioned contributions.

Overall, the only other paper in the literature that studies Bayesian persuasion in the presence of noise is Le Treust and Tomala (2018). In their work, they also consider information distortions similarly to our work, but they allow for multiple experiments that are conducted sequentially. Then, they study the effect of noise on the sender's expected utility as the number of experiments increases.

The paper is structured as follows: Section 2 provides a leading example that will be used throughout the paper. Section 3 introduces our model and some preliminary results. In Section 4 we study the sender's optimal signal, emphasizing that the standard (concavification) technique that is widely used in the literature is only partially applicable to our case. In Section 5 we present our main monotonicity results. Section 6 revisits our motivating example. In Section 7 we present various applications of our general model. Section 8 contains a concluding discussion. All proofs are relegated to the Appendices.

2. Motivating example

The following example is an adaptation of the one used by Kamenica and Gentzkow (2011). Consider a (female) prosecutor who conducts an investigation whose outcome should be fully reported to a (male) judge, who in turn has to decide whether to convict or acquit a defendant. There are two states of nature, namely the defendant can be either innocent (I) or guilty (G). The prosecutor always prefers the conviction of the defendant, whereas the judge prefers to convict the defendant if he is guilty and acquit otherwise. Both agents are Bayesian expected utility maximizers and share a common prior that attaches probability $\mu_0 \in (0,1)$ to G.

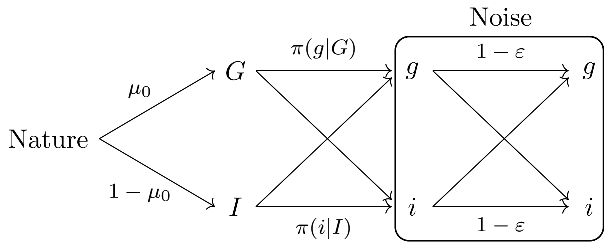


Figure 1: A noisy version of Kamenica and Gentzkow's (2011) prosecutor-judge example.

The prosecutor commits to an investigation (*signal*), which is represented by a pair of distributions $\pi(\cdot|G)$ and $\pi(\cdot|I)$. The judge is assumed to know the investigation that the prosecutor runs. The realized outcome of the investigation (*message*) is observed by the prosecutor, who is legally obliged to truthfully report it to the judge. However communication is noisy, implying that the judge may misunderstand (with probability ε) the prosecutor's message, e.g., due to non-appropriate use of words by the prosecutor, or due to misinterpretation of the arguments by the judge etc.

Due to the fact that communication is noisy, the judge observes the outcome of some effective investigation σ_π, which is a noisy version of π. For instance, when the prosecutor has chosen the investigation π, the probability of the judge hearing g given that the defendant is actually guilty is equal to $\sigma_\pi(g|G) = (1-\varepsilon)\pi(g|G) + \varepsilon\pi(i|G)$ rather than $\pi(g|G)$. Then, using Bayes rule, the judge will form posterior belief $\mu_g \in (0,1)$ with probability $\tau_\pi(\mu_g) = \mu_0\sigma_\pi(g|G) + (1-\mu_0)\sigma_\pi(g|I)$ and posterior belief $\mu_i \in (0,1)$ with the remaining probability $\tau_\pi(\mu_i) = 1 - \tau_\pi(\mu_g)$. Thus, the prosecutor anticipates the judge's action for each of the two posteriors and calculates her expected utility $\hat{v}(\pi)$ as a function of the signal π. Hence, the prosecutor's problem boils down to choosing an investigation that maximizes $\hat{v}(\pi)$. Obviously, the presence of noise will typically alter the optimal choice of the prosecutor, and a fortiori her expected utility. This is because the distributions of posteriors that the judge can form depends on the structure of the noisy channel. The general question is then: *which are the effects of noise on the prosecutor's optimal signal and on her expected utility?*

3. Persuasion game with noise

3.1. Our model

Let $\Omega = \{\omega_1, \ldots, \omega_N\}$ be a (finite) set of states and A be a compact action space. There are two agents, a (female) sender and a (male) receiver, with a common full-support prior $\mu_0 \in \Delta(\Omega)$ and continuous utility functions, $v : A \times \Omega \to \mathbb{R}$ and $u : A \times \Omega \to \mathbb{R}$ respectively.

Let $S = \{s_1, \ldots, s_K\}$ be a finite set of messages that can be encoded with the available technology.

We assume that K is large. In particular, we assume that $K \geq N$ (see Section 8.1). Noise is modelled with a *channel* $p : S \to \Delta(S)$.[2] A *p-signalling structure* (or simply a *signalling structure*) consists of a pair of signals, $\pi : \Omega \to \Delta(S)$ and $\sigma_\pi : \Omega \to \Delta(S)$ such that

$$\sigma_\pi(s|\omega) = \sum_{t \in S} p(s|t)\pi(t|\omega) \tag{1}$$

for every $\omega \in \Omega$ and every $s \in S$, i.e., σ is a garbling of π via the noisy channel p. For a signalling structure, (π, σ_π), we call π the *actual signal* and σ_π the *effective signal.* Obviously, given a channel p, each signalling structure is identified by the actual signal. Under certain conditions the converse is also true, i.e., the signalling structure can be also identified by the effective signal. The set of actual signals is denoted by $\Pi := (\Delta(S))^\Omega$, whereas the set of effective signals is denoted by $\Sigma_p := \{\sigma_\pi | \pi \in \Pi\} \subseteq \Pi$. Whenever it is clear which is the channel, we omit the subscript, thus simply writing Σ.

A signalling structure can be interpreted in a number of ways. All interpretations have the common features that (i) the sender commits to an experimental design (viz., the actual signal) which is known to the receiver, and (ii) the receiver observes some – possibly distorted – data (viz., a realization of the effective signal) before choosing an action. Interpretations differ in how the data is gathered, processed and transmitted. Thus, we can model various different forms of data distortions (viz., noisy channels), usually appearing in some of the following stages:

- Data gathering: The experiment is run by an agent, henceforth called the data collector, who could in principle be the sender or the receiver or even a third party. The data collector observes a noisy version of the actual data, due to *measurement errors* in the underlying the experiment.

- Data processing: The raw data is gathered by the data collector and is processed before being used by the receiver. Processing can take the form of storage (either in the collector's memory or in some external device) and retrieval at a later time, in which case noise is attributed to *memory constraints.* Alternatively, processing errors can be due to the collector's *lack of expertise* which precludes him/her from correctly encoding or interpreting the actual message.

- Data transmission: The data collector is some agent other than the receiver who (truthfully) communicates the observed data to the receiver. The receiver observes a noisy version of the transmitted message, due to *communication errors* or *language barriers* that lead to misunderstanding of the communicated data.

[2] As we later discuss (Section 3.3.2) and formally prove (Proposition E1), our theory extends to nearly all cases where the channel's input and output message space differ from each other.

The noisy channel p often combines multiple types of data distortions (see Section 7 for applications).

It is important to stress that *messages do not have a particular meaning within our model.* Instead, meaning is acquired via the signalling structure. Nevertheless, we often consider noisy channels that do not treat all messages in the same way. At first glance, this may seem as if we implicitly attach semantics to our messages. However, this is not the case. Such differences are only meant to capture the idea that some messages are more difficult to gather/process/transmit, thus leading to possibly "more mistakes", *irrespective of the meaning that they carry.* We further elaborate on this issue in Section 8.1.

3.2. Equilibrium existence

After the sender having chosen some signal $\pi \in \Pi$ and the receiver having heard some $s \in \operatorname{supp}\big(\sigma_\pi(\cdot|\omega)\big)$, the receiver forms a posterior belief $\mu_s \in \Delta(\Omega)$ via Bayes rule, viz., for each $\omega \in \Omega$,

$$\mu_s(\omega) = \frac{\mu_0(\omega)\sigma_\pi(s|\omega)}{\mathbb{E}_0[\sigma_\pi(s|\cdot)]}, \tag{2}$$

where $\mathbb{E}_0[x] := \langle \mu_0, x\rangle$ for each $x \in \mathbb{R}^\Omega$. For each $s \in S$ we define $M_s \subseteq \Delta(\Omega)$ as the set of posteriors μ_s induced by some $\sigma \in \Sigma$ with $\sigma(s|\omega) > 0$ for some $\omega \in \Omega$. Let $M := \bigcup_{s\in S} M_s$ denote the set of all posteriors that the receiver could possibly form. Then, once the receiver has formed his posterior $\mu \in M$, he chooses an action that maximizes his expected utility, $u_\mu(a) = \langle \mu, u(a,\cdot)\rangle$. Since u_μ is continuous over the compact set A, a maximum always exists. If there are multiple maxima, the receiver chooses the one that maximizes the sender's expected utility (given μ). If there are multiple sender-preferred maxima, the receiver picks an arbitrary one. We denote the receiver's optimal action, given the posterior $\mu \in M$, by $\hat{a}(\mu)$.

Now, given the signal $\pi \in \Pi$, the sender forms a distribution (viz., second order belief) $\tau_\pi \in \Delta(M)$ over the receiver's posteriors, i.e., for each $\mu \in M$,

$$\tau_\pi(\mu) = \sum_{s\in S} p(\{t \in S : \mu_t = \mu\}|s)\mathbb{E}_0[\pi(s|\cdot)]. \tag{3}$$

The sender's expected utility from $\pi \in \Pi$ (given that the receiver chooses optimally) is equal to

$$\hat{v}(\pi) := \mathbb{E}_{\tau_\pi}[\hat{v}_0], \tag{4}$$

where $\hat{v}_0(\mu) := \mathbb{E}_0[v(\hat{a}(\mu),\cdot)]$ for each $\mu \in M$. An *optimal signal* for the sender is one from $\arg\max_{\pi\in\Pi} \hat{v}(\pi)$. We denote the *(sender's) value* of her optimal signal by

$$\hat{v}^*_p := \max_{\pi\in\Pi} \hat{v}(\pi). \tag{5}$$

Then we prove that a (sender-preferred) subgame perfect equilibrium exists.

Proposition 1. *An optimal signal for the sender exists.*

The proof does not rely on a concavification argument (Kamenica and Gentzkow, 2011), which is typically used to characterize the optimal signal and a fortiori to prove existence in persuasion games. In Section 4 we review the concavification argument and illustrate why it cannot be used to prove our existence theorem. Instead, our proof proceeds by showing that the set of distributions over posterior beliefs that can be achieved by some signal is compact. Of course our proof can also be used in the standard noiseless Bayesian persuasion game, thus providing an alternative non-constructive existence proof, more general than the one in Kamenica and Gentzkow (2011).[3] Our complete proof is relegated to Appendix A.

3.3. Channel structure

According to the standard representation, a channel p is identified by a $K \times K$ stochastic matrix P with typical entry $P_{k,\ell} := p(s_\ell | s_k)$. Throughout the paper we arbitrarily interchange p and P. Our framework does not restrict the possible structures of noise. Nevertheless, some families of channels are adequate for modelling basic specific sources of noise, and for this reason we analyze them separately.

We first restrict attention to noisy channels that satisfy the following property:

(A_1) Properly noisy: There are at most $N-1$ messages in S that can be fully distinguished from each other, i.e., if there is a partition $\mathcal{T}$ of S such that $p(T|s) = 1$ for every $s \in T$ and every $T \in \mathcal{T}$, then $|\mathcal{T}| < N$.

The previous condition guarantees that the sender cannot circumvent the restrictions imposed by the noisy channel, and therefore the sender's problem does not trivially degenerate to the standard (noiseless) model of Bayesian persuasion, which has already been extensively studied in the literature. Note that this condition is assumed to hold despite the fact that the set of messages is sufficiently rich ($K \geq N$), i.e., roughly speaking, although there are many actual messages, most of them can be confused with each other.

Definition 1. Two channels p and q are said to be *isometric* whenever one can be obtained by sequential permutations of rows and columns of the other. $\triangleleft$

An isometry class is a family of channels that is closed with respect to row and column permutations. The most obvious isometry class is the class of noiseless channels, which contains all permutation matrices. It is not difficult to verify that all channels in a isometry class induce the

[3] The term "more general" refers to the fact that our proof works both in the noisy and the noiseless case.

same set of distributions over posterior beliefs – as we merely relabel the messages – and therefore for our purposes they are deemed equivalent. Thus, the following result follows directly.

Proposition 2. *If p and q are isometric, then $\hat{v}_p^* = \hat{v}_q^*$.*

In the remainder of this section we introduce two basic classes of noisy channels. Such channels appear in various applications. Moreover, they often form the basis of more complex noise structures. In fact, as we show later in the paper, all channels of interest can be obtained by combining such basic channels. In this sense, studying these classes is of particular importance.

3.3.1. Canonical channels

Consider the following two properties of noisy channels:

(A_2) Diagonally dominant: An arbitrary entry in the main diagonal of P is larger than the sum of the remaining entries in the same row, i.e., $p(s|s) > 1/2$ for all $s \in S$.

(A_3) Symmetric: The matrix P is symmetric, i.e., $p(s|t) = p(t|s)$ for all $s, t \in S$.[4]

Condition (A_2) says that errors occur with relatively low probability, viz., the probability of the receiver hearing the actual message is larger than $1/2$. Condition (A_3) postulates that the probability of confusing s with t is equal to the probability of confusing t with s. For instance, assume that there is an underlying metric d in S and the error probabilities are monotonic in the distance between the messages, i.e., the further away two messages are from each other, the less likely it is that they get mixed up. Then, from the symmetry of the metric follows the symmetry of the transition matrix.

Definition 2. We say that P is *canonical* whenever it is isometric to some diagonally dominant channel, and isometric to some (perhaps other) symmetric channel. ◁

Well-known examples of canonical channels contain different versions of noisy typewriters (Cover and Thomas, 2006) and different versions of circulant matrices, which constitute a special case of Latin squares (Marshall et al., 2011).[5] For instance, consider the case where the receiver hears the true message with probability $1 - \varepsilon$ and every other message with (small) equal probability $\varepsilon/(K-1)$. Such matrices are called *strongly symmetric*. Whenever $|S| = 2$, all previous examples collapse to

[4] In information theory the term "symmetric channel" is reserved for some P such that every two rows and every two columns are permutations of each other respectively (Cover and Thomas, 2006). Obviously the two notions of symmetry differ from each other.

[5] Latin square is called an $N \times N$ square matrix with N different elements, each appearing exactly once in each row and exactly once in each column.

the well-known binary symmetric channel, such as the one described in the motivating example of Section 2.

Canonical channels are often used to model basic forms of *noisy data gathering* (e.g., the data collector observes false data with some small probability due to measurements errors related to the experimental technology) or *noisy data processing* (e.g., the data are stored in some memory and when retrieved mistakes occur with small probability) or *noisy data transmission* (e.g., the actual message is accurately observed by the data collector and truthfully communicated to the receiver who in turn hears a different message with some small probability).

3.3.2. Partitional channels

Consider the following property of noisy channels:

(A_4) Idempotent doubly stochastic: There is a partition $\mathcal{T}$ of S, such that for each $T \in \mathcal{T}$, if $s, t \in T$ then $p(t|s) = 1/|T|$.

The previous condition says that messages within the same $T \in \mathcal{T}$ are completely indistinguishable to the data collector or the receiver (Aumann, 1976). Thus, whenever $s \in T$ is sent, $T \in \mathcal{T}$ is heard, suggesting that at most $|\mathcal{T}|$ messages can be effectively used, and consequently reducing the dimension of the effective signal.

Definition 3. We say that P is *partitional* whenever it is isometric to some idempotent doubly stochastic channel. ◁

It is easy to verify that every idempotent doubly stochastic channel is also symmetric, but not necessarily diagonally dominant. In fact the only partitional channels that are also canonical are the noiseless channels, implying that the two classes are essentially disjoint.

Partitional channels are often used to model basic forms of *noisy data processing* (e.g., the data collector has limited perception, thus failing to understand the difference between data points that the experiment may yield) or *noisy data transmission* (e.g., the receiver has a coarser language than the data collector).

Interestingly, every noisy channel with rational transition probabilities (including those with different input and output message spaces) can be obtained as a combination of two partitional channels with the same message space (see Proposition E1).[6] Notably the latter is true for canonical channels (with $\varepsilon \in \mathbb{Q}$) too, and in this sense we slightly abuse terminology by classifying canonical channels as basic. In either case, the aforementioned result suggests that *partitional channels are the building blocks of any form of noise.*

[6]The formal definition of the combination of two channels is introduced after we define garbling (see Section 5).

4. Optimal signal

In their seminal paper, Kamenica and Gentzkow (2011) computed the exact value of the sender's optimal signal for the noiseless case, using a standard concavification technique which was originally introduced in the literature of repeated games by Aumann and Maschler (1995).[7] Then, we naturally ask: *can their characterization be extended to the noisy case?*

Let us first briefly revisit their approach. Whenever p is noiseless we arbitrarily interchange notation between $\Delta(\Omega)$ and M. A distribution $\tau \in \Delta(\Delta(\Omega))$ is said to be Bayes-plausible whenever $\int_{\Delta(\Omega)} \mu d\tau(\mu) = \mu_0$. Obviously, the sender's (second-order) belief τ_π is Bayes-plausible for every $\pi \in \Pi$. They show that the sender's problem reduces from maximizing over the set of signals to maximizing over the set of Bayes-plausible distributions:

Lemma 1 (Kamenica and Gentzkow, 2011)**.** *Let p be a noiseless channel. Then, for every Bayes-plausible $\tau \in \Delta(M)$ there exists some $\pi \in \Pi$ such that $\mathbb{E}_\tau[\hat{v}_0] = \hat{v}(\pi)$.*

The result is proven in two steps. First, it follows from Caratheodory Theorem that, for every Bayes-plausible $\tau \in \Delta(\Delta(\Omega))$ there exists some Bayes-plausible $\tau^* \in \Delta(\Delta(\Omega))$ (possibly different from τ) such that $|\operatorname{supp}(\tau^*)| \leq N+1$ and $\mathbb{E}_{\tau^*}[\hat{v}_0] = \mathbb{E}_\tau[\hat{v}_0]$. Second, they prove that there exists a signal $\pi \in \Pi$ such that $\tau_\pi = \tau^*$. This transformation allows not only to prove that an optimal signal exists, but more importantly to explicitly calculate its value. In particular, define the concave closure of $\hat{v}_0$ as follows:

$$V_M(\mu) := \sup \Big\{ \ z \in \mathbb{R} : (\mu, z) \in \operatorname{conv}\big(\{(\lambda, \hat{v}_0(\lambda)) \mid \lambda \in M\}\big) \ \Big\} \tag{6}$$

for each $\mu \in M$. Given this definition, the authors show that the value of the optimal signal is given by $\hat{v}^* = V_{\Delta(\Omega)}(\mu_0)$, implying that persuasion is beneficial for the sender if and only if $V_{\Delta(\Omega)}(\mu_0) > \hat{v}_0(\mu_0)$, i.e. they obtain a necessary and sufficient condition for persuasion. In fact, if $\hat{v}_0$ is convex but not concave then persuasion is always beneficial, whereas if $\hat{v}_0$ is concave then persuasion is never beneficial.

Now let us switch focus to the noisy case, asking whether a similar characterization result can be established. Formally, we ask: *for an arbitrary channel p, is it the case that $\hat{v}^*_p = V_M(\mu_0)$?* As it turns out this is not the case in general.

Example 1. Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and $\mu_0 \in \Delta(\Omega)$ be uniformly distributed. Moreover let $A := \Delta(\Omega)$ and assume that the two agents have aligned preferences, with their expected utility being given by the strictly convex function $\|\mu - \mu_0\|^2$, which is maximized at the extreme points of $\Delta(\Omega)$ with the

[7] A similar approach has been taken in subsequent papers on Bayesian persuasion (e.g., see Alonso and Câmara, 2016; Bizzotto et al., 2017; Hedlund, 2017; Kolotilin et al., 2018; Laclau and Renou, 2017; Wang, 2013).

corresponding expected utility being equal to 2/3, i.e., the receiver reports his posterior belief and they are both paid by a proper scoring rule with the highest expected utility being achieved when the receiver learns the true state.[8] Assume that noise is modelled by a partitional channel with only two effective messages, i.e., S is partitioned into $\mathcal{T} = \{T_1, T_2\}$ so that $p(t|s) = 1/|T|$ for each $s, t \in T$ and each $T \in \mathcal{T}$. Notice that every posterior can be achieved, i.e., $M = \Delta(\Omega)$, implying that $V_M(\mu_0) = 2/3$. Nevertheless the sender can only use signals that yield at most two messages with positive probability, i.e., she can only choose Bayes-plausible $\tau \in \Delta(M)$ such that $\tau(\{\mu_1, \mu_2\}) = 1$ for some $\mu_1, \mu_2 \in M$ which are collinear with μ_0. Thus, if one of these posteriors is an extreme point, the other one is not. Hence, $\hat{v}_p^* < 2/3$, implying that the value of the optimal signal is not given by the concave closure of $\hat{v}_0$. ◁

The reason why concavification fails is that whenever p is noisy Lemma 1 does not necessarily hold. In particular, although for every $\tau \in \Delta(M)$ there exists some $\tau^* \in \Delta(M)$ such that $|\operatorname{supp}(\tau^*)| \leq N+1$ and $\mathbb{E}_{\tau^*}[\hat{v}_0] = \mathbb{E}_\tau[\hat{v}_0]$, it is not necessarily the case that there exists some $\pi \in \Pi$ such that $\tau_\pi = \tau^*$. Indeed, in the previous example, there is no signal yielding a uniform distribution over the three extreme points of $\Delta(\Omega)$, although all three of them belong to M.

While this approach does not pin down the value of the optimal signal, it still provides an upper bound. Furthermore, it provides a necessary condition for persuasion to be beneficial for the sender.

Theorem 1. *For an arbitrary channel p, the following hold:*

(i) The sender's value is bounded, viz., $\hat{v}_p^* \leq V_M(\mu_0)$.

(ii) Persuasion is not beneficial if $\hat{v}_0(\mu_0) = V_M(\mu_0)$.

The proof of the previous theorem follows almost immediately from Kamenica and Gentzkow's (2011) original result. As we have already mentioned, the reason why concavification does not provide a sufficient condition for persuasion to be beneficial, is that not all Bayes-plausible distributions with support in M can be induced by some signal. This is true even in cases where the channel has a nice structure, e.g., canonical (Section 6) or partitional (Example 1). Nevertheless, in certain special cases, we can provide sufficient conditions for persuasion, similar to the noiseless case, viz., when $\hat{v}_0$ is strictly convex, persuasion is beneficial, unless of course p is completely uninformative.

The previous result also directly implies that noise is always harmful for the sender, as $M \subseteq \Delta(\Omega)$ implies $V_M(\mu_0) \leq V_{\Delta(\Omega)}(\mu_0)$, and therefore $\hat{v}_p^* \leq V_{\Delta(\Omega)}(\mu_0) = \hat{v}^*$. This implies that Blume et al.'s (2007) result for signalling games does not extend to persuasion games. That is, *commitment makes noise always harmful for the sender.*

[8] Recall that under a proper scoring rule it is strictly dominant for the receiver to report his true posterior belief.

5. Monotonicity

Is more noise always harmful for the sender? In order to answer this question, we first recall Blackwell's (incomplete) informativeness relation over the set of communication channels (Blackwell, 1951, 1953). We say that q is a *garbling* of p (viz., q is *more noisy* than p) whenever there is a channel $r : S \to \Delta(S)$ such that

$$q(t|s) = \sum_{u \in S} p(u|s) r(t|u) \tag{7}$$

for each $s, t \in S$. In this case we write $p \succeq q$. The channel r models the additional noise that we combine with p to obtain the noisier channel q. Whenever r is not noiseless, we write $p \succ q$. Then, we ask: *does $p \succeq q$ imply $\hat{v}^*_p \geq \hat{v}^*_q$?*

First, in the most obvious special case where p is noiseless, as we have already shown, noise is harmful for the sender. Second, consider the case where the sender and receiver have aligned preferences (i.e., $u = v$). It is again quite clear that more noise is harmful for the sender (Proposition E2). This follows directly from Blackwell's well-known theorem (Blackwell, 1951, 1953). Third, when two diagonally dominant channels are Blackwell equally-informative (i.e., $p \succeq q$ and $q \succeq p$), they always induce the same value (Proposition E3).

However, it turns out that the sender's value is not always increasing in the channel's informativeness, i.e., more noise may be beneficial for the sender, as illustrated by the following example.

Example 2. Let $\Omega = \{\omega_1, \omega_2, \omega_3\}$ and $\mu_0 \in \Delta(\Omega)$ be uniformly distributed. Moreover, let $S = \{s_1, \ldots, s_{20}\}$, and consider the partitional channel p with $\mathcal{T}_p = \{T_1, T_2\}$ such that $T_1 = \{s_1, \ldots, s_{10}\}$ and $T_2 = \{s_{11}, \ldots, s_{20}\}$. Suppose that additional noise is introduced (on top of p) via another partitional channel r with $\mathcal{T}_r = \{T'_1, T'_2, T'_3\}$ where $T'_1 = \{s_1, s_2\} \cup \{s_{11}, s_{12}, s_{13}\}$, $T'_2 = \{s_3, s_4, s_5\} \cup \{s_{14}, \ldots, s_{18}\}$ and $T'_3 = \{s_6, \ldots, s_{10}\} \cup \{s_{19}, s_{20}\}$. Thus, we obtain a new channel q, which is a garbling of p (via r). Note that q is not partitional, even though both p and r are. Now, take a signal π such that $\pi(T_1|\omega_1) = 1$, $\pi(T_2|\omega_2) = 1$ and $\pi(T_1|\omega_3) = \pi(T_2|\omega_3) = 1/2$. In this case, the distribution of posteriors τ^q_π puts positive probability to three beliefs in $\Delta(\Omega)$, (viz., $\mu^q_1 = (\frac{4}{21}, \frac{7}{21}, \frac{10}{21})$, $\mu^q_2 = (\frac{2}{8}, \frac{3}{8}, \frac{3}{8})$ and $\mu^q_3 = (\frac{10}{21}, \frac{7}{21}, \frac{4}{21})$), which are not collinear. Importantly this distribution cannot be obtained under p which induces at most two posterior beliefs, collinear with the prior, as there are only two effective messages (i.e., $|\mathcal{T}_p| = 2$). Finally, assume that the set of actions is $A := \Delta(\Omega)$, and the receiver's expected utility is given by $\|\mu - \mu_0\|^2$ (i.e., similarly to Example 1 the sender is asked to report his posterior belief and is paid according to a proper scoring rule). The sender's utility is equal to 1 if the receiver reports some $\mu \in \{\mu^q_1, \mu^q_2, \mu^q_3\}$ and 0 otherwise, i.e., intuitively, she has bet on the event that the receiver will report a posterior belief in $\{\mu^q_1, \mu^q_2, \mu^q_3\}$. The latter directly implies that the sender's value is equal to $\hat{v}^*_q = \hat{v}(\pi) = 1$ under the channel q. However, since the three posteriors are not collinear, the sender can achieve at most one of them under the channel p, thus implying $\hat{v}_p^* < 1$. Hence, $\hat{v}_p^* < \hat{v}_q^*$ even though $p \succ q$. A more detailed presentation of this example can be found in Appendix E.2.1. ◁

Crucially the previous example relies on the fact that p and q have inherently different structures, viz., p is partitional whereas q is not. In fact, q is essentially a channel with different input and output effective message spaces, viz., $\mathcal{T}_p$ and $\mathcal{T}_r$ respectively. This explains the reversal in the sender's preferences for the two channels. This type of structural differences usually arise when different sources of data distortion coexist. In the previous example for instance, assume that p models noisy data processing (e.g., due lack of expertise of the data collector) and r models noisy data transmission (e.g., due to language barriers in the communication between the data collector and the receiver). Finally, let us stress once again that in our previous example monotonicity is violated even though both p and r have a nice structure (i.e., they are both partitional), implying that the structural differences between p and q are not an artefact of combining "exotic" channels.

In what follows we focus on the basic channels that we discussed earlier. We show that, if both channels (p and q) are canonical or both are partitional, monotonicity is restored.

Theorem 2. *If $p \succeq q$ then $\hat{v}_p^* \geq \hat{v}_q^*$, when one of the following conditions holds:*

(i) p and q are canonical channels.

(ii) p and q are partitional channels.

For both cases, the main idea behind the proof is to show that the more informative channel yields a larger set of feasible distributions of posteriors, for every signal and every preference profile of the two agents. Hence, it induces a higher value for the sender. To prove the first part, we show that every row of Q (the stochastic matrix of q) can be written as a convex combination of the rows of P (the stochastic matrix of p), thus implying that the extreme points of Σ_q all belong to Σ_p. Hence, there are more feasible distributions of posteriors under p than under q. To prove the second part, we show that for partitional channels, it is the case that $p \succeq q$ if and only if p has more effective messages than q. The latter also implies that $\succeq$ is complete over the set of partitional channels, implying that in this case the converse holds too, i.e., if $\hat{v}_p^* \geq \hat{v}_q^*$ then $p \succeq q$. In other words, $\succeq$ naturally induces a complete preference relation (for the sender) over the set of partitional channels. This is not necessarily the case when the channels are canonical, as in that case $\succeq$ is incomplete (see Section 8.2).

Within each of the previous two classes of channels, Blackwell's informativeness relation is equivalent to another well-known relation, viz., *(matrix) majorization*, which is used extensively as a measure of dispersion in the theory of income inequality (e.g., Lorentz curves or Dalton transfers). Formally, we say that P majorizes Q and we write $P \trianglerighteq Q$, whenever $Q = PR$ for some doubly stochastic matrix R (Marshall et al., 2011). Then, it can be shown that for every pair of canonical or partitional matrices, $p \succeq q$ if and only if $P \trianglerighteq Q$ (Lemma C3).

An example of Blackwell-ordered canonical channels ($p \succeq q$) is one where p and q are strongly symmetric with q yielding a larger error probability. An example of Blackwell-ordered partitional channels ($p \succeq q$) occurs when q is obtained by combining p with another partitional channel r such that $\mathcal{T}_p$ refines $\mathcal{T}_r$. Note that the latter is violated in Example 2 where $\mathcal{T}_r$ is not a coarsening of $\mathcal{T}_p$.

The general conclusion of this section – and to an extent of the entire paper – is that for basic Blackwell-ordered channels of similar structure, more noise is harmful for the sender. Such similarities are observed when the source of data distortions is the same in the two channels, and therefore in a sense the two channels are directly comparable. However, *when there are multiple sources of data distortions (as it is commonly the case in applications), such similarities often disappear and consequently more noise can benefit the sender.*

6. Motivating example revisited

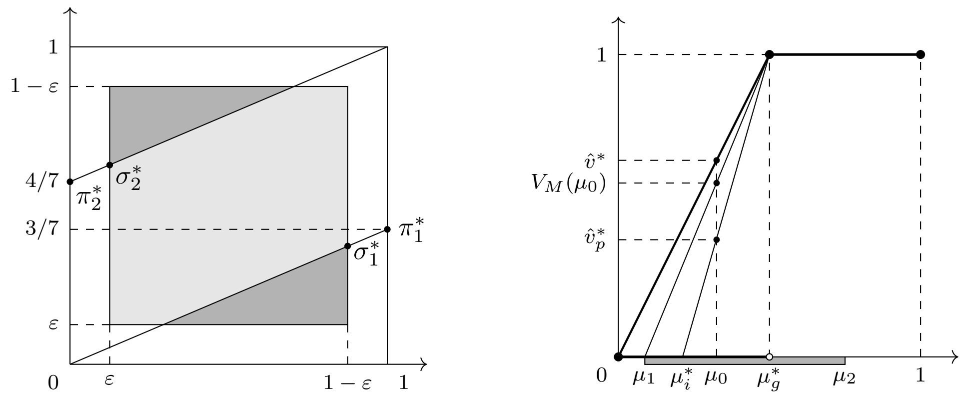
Recall the prosecution example with noise from Section 2, supposing that the prior probability of guilt is $\mu_0 = 0.3$ (like in Kamenica and Gentzkow, 2011) and the probability of misinterpretation under the strongly symmetric channel p is $\varepsilon < 0.3$. *Which is the optimal signal?*

The prosecutor would like to maximize the probability of conviction, viz., she would like to maximize the probability that the judge's posterior attaches probability at least 0.5 to the defendant being guilty. Obviously, by $\mu_0 < 0.5$, every signal would induce at most one such posterior, i.e., it cannot be the case that both $\mu_g \geq 0.5$ and $\mu_i \geq 0.5$. Thus, the prosecutor focuses on signals that yield either $\mu_g \geq 0.5$ or $\mu_i \geq 0.5$ with positive probability. These are the only candidates for an optimal signal and every other signal will induce the same expected utility as the prior, namely 0. The signals that yield $\mu_g \geq 0.5$ are the ones in the lower shaded triangle of Figure 2.a, while the signals that yield $\mu_i \geq 0.5$ are those in the upper shaded triangle. Let us focus on the signals that induce $\mu_g \geq 0.5$, and by symmetry the analysis is identical for those inducing $\mu_i \geq 0.5$. For an arbitrary σ in the lower triangle, the probability that the prosecutor's distribution attaches to μ_g is equal to $\tau_\sigma(\mu_g) = 0.3\sigma(g|G) + 0.7\sigma(g|I)$. The prosecutor then faces a constrained linear optimization problem, viz., she wants to maximize $\tau_\sigma(\mu_g)$ subject to σ belonging to the lower triangle. Obviously, the optimal solution to this problem is σ_1^*, where $\sigma_1^*(g|G) = 1 - \varepsilon$ and $\sigma_1^*(g|I) = 3(1-\varepsilon)/7$. Notice that by symmetry, σ_2^*, defined by $\sigma_2^*(g|\cdot) = 1 - \sigma_1^*(g|\cdot)$, is also an optimal signal. In fact, σ_1^* and σ_2^*

are the only optimal signals. Finally notice that

$$\hat{v}_p^* = 0.6(1-\varepsilon).$$

Observe that, consistently with Theorem 2, the sender's value is decreasing in the error probability.



(a) Judge's effective investigation.

(b) Prosecutor's expected utility.

Figure 2: Optimal investigation in the prosecution example with noise.

Now, let us switch attention to Figure 2.b. It is not difficult to verify that the set of posteriors μ_s that the judge could form conditional on each of the two messages $s \in \{i, g\}$ is $M = \{\mu \in [0,1] : \mu_1 \leq \mu \leq \mu_2\}$, where $\mu_1 = 3\varepsilon/(3\varepsilon + 7(1-\varepsilon))$ and $\mu_2 = 3(1-\varepsilon)/(3(1-\varepsilon)+7\varepsilon)$. Observe that the two posteriors that occur with positive probability when σ_1^* is chosen are such that $\mu_i^* = 3\varepsilon/(4+6\varepsilon)$ and $\mu_g^* = 0.5$. Then, it is straightforward to verify that $\mu_1 < \mu_i^*$, implying that $\hat{v}_p^* < V_M(\mu_0)$. Notice that consistently with Theorem 1, the concave closure of $\hat{v}_0$ merely provides an upper bound rather than a characterization of the prosecutor's value.

7. Applications

7.1. Persuading voters

Building on Alonso and Câmara (2017), consider a politician (the sender) who proposes a new agenda a_1 to replace the status quo policy a_0. There are three voters (the receivers) $I := \{1,2,3\}$ and two states $\Omega = \{\omega_1, \omega_2\}$. A (prior or posterior) belief in $\Delta(\Omega)$ is identified by the probability it assigns to ω_1. All agents are assumed to share a common prior, which for simplicity and without loss of generality is assumed to be $\mu_0 = 0.5$. For each $i \in I$ the preferences are represented by the random variable $\delta_i : \Omega \to \mathbb{R}$, where $\delta_i(\omega) := u_i(a_1, \omega) - u_i(a_0, \omega)$ is i's net utility from adopting the new agenda at ω. We assume that $\delta_i^* := \delta_i(\omega_1) > 0$ and $\delta_i(\omega_2) = -1$ for all $i \in I$, implying that all voters agree that the new agenda is good at ω_1 and it is bad at ω_2. Let $\mu_i^* := 1/(1+\delta_i^*)$ be the threshold belief that leads i to vote in favor of a_1, i.e., voter i prefers a_1 to a_0 under a belief $\mu \in [0,1]$ if and only if $\mu \geq \mu_i^*$. We assume that $\mu_0 < \mu_1^* < \mu_2^* < \mu_3^*$, i.e., no voter approves the new policy under the prior belief. In general, the politician cares about how many votes she gets, viz., $v(\kappa)$ is the politician's utility from getting κ votes. In particular, throughout this section, we let $0 = v(0) = v(1) < v(2) \leq v(3)$, i.e., in order for the new agenda to be approved a simple majority is needed, but the politician may still prefer unanimous support over simple majority.

The politician can influence the voters by means of a binary political experiment π (the signal), with possible outcomes $S := \{s_1, s_2\}$. There are distortions in the messages that the voters receive, modelled by a strongly symmetric channel p with error probability $\varepsilon \in [0, 1/2)$, i.e., $p(s|s) = 1 - \varepsilon$ and $p(t|s) = \varepsilon$ for $t \neq s$. For instance, a voter could misinterpret the politician's message due to campaign mistakes, or due to the fact that political messages are communicated via the media which may intentionally or unintentionally distort their meaning, or even due to his own limited ability to understand the message. Depending on the source of the distortions we consider two cases:

- A *common effective message* is drawn for all voters, e.g., when noise is due to campaign errors or media distortions.
- A *different effective message* is independently drawn for each voter, e.g., when noise is due to voters' own misunderstanding.

Note that in both cases, the actual message is the same for all voters, implying that differences in the effective messages – and the subsequent belief heterogeneity – can only be attributed to data distortions.

Each binary signal π is identified by the pair of posteriors in the support of the distribution τ_π. A pair of posteriors (μ_0^-, μ_0^+) is feasible if and only if $\varepsilon\mu_0^+/(2\mu_0^+ - (1-\varepsilon)) \leq \mu_0^- \leq (1-\varepsilon)\mu_0^+/(2\mu_0^+ - \varepsilon)$ (Le Treust and Tomala, 2018, Lem. 3.4). In this case we refer to μ_0^- as the "bad posterior" and to μ_0^+ as the "good posterior". The intuition is that no voter will be persuaded to approve the new agenda under the bad posterior, and may or may not under the good posterior depending on how large μ_0^+ is. For a fixed good posterior $\mu_0^+ \geq \mu_2^*$, the politician wants the bad posterior μ_0^- to be as small as possible (Alonso and Câmara, 2017, Prop. 1). It is not difficult to verify that the optimal signal always belongs to $\{\pi_2^\varepsilon, \pi_3^\varepsilon\}$, where $\pi_i^\varepsilon \in \{\pi_2^\varepsilon, \pi_3^\varepsilon\}$ yields the good posterior μ_i^* with probability $\tau_i^\varepsilon := (1-\varepsilon)/2\mu_i^*$ and the bad posterior $\varepsilon\mu_i^*/(2\mu_i^* - (1-\varepsilon))$ with probability $1 - \tau_i^\varepsilon$. The latter is true irrespective of whether all voters receive common or independent effective messages.

If the politician chooses π_2^ε then she targets *approval by majority*, whereas if she chooses π_3^ε then she targets *approval by unanimity*. Thus, we say that the politician prefers to target majority (resp., unanimity) when π_2^ε (resp., π_3^ε) is the optimal signal. It is not difficult to verify that the optimal signal depends on various parameters, namely on the voter's preferences, on the politician's preferences and on the error probability.

Proposition 3. *Suppose that all voters receive a common effective message. Then, for every preference profile, the politician prefers to target approval by majority (resp., by unanimity) without noise if and only if she prefers to target approval by majority (resp., by unanimity) with noise, i.e., formally,*

$$\hat{v}(\pi_2^0) \geq \hat{v}(\pi_3^0) \textit{ if and only if } \hat{v}(\pi_2^\varepsilon) \geq \hat{v}(\pi_3^\varepsilon) \textit{ for every } \varepsilon > 0.$$

The previous result follows from $\hat{v}(\pi_2^\varepsilon) = (1-\varepsilon)v(2)/2\mu_2^*$ and $\hat{v}(\pi_3^\varepsilon) = (1-\varepsilon)v(3)/2\mu_3^*$ for every $\varepsilon \geq 0$. A direct consequence of the previous result is that, in equilibrium the politician always targets the same good posterior irrespective of the amount of noise. Nevertheless, the probability of achieving this good posterior decreases in ε, i.e., τ_i^ε is strictly decreasing in ε. This is because the bad posterior moves closer to the prior as ε increases. Hence, *when all voters receive the same effective signal, the noisier the channel, the less informative the optimal signal.*

Proposition 4. *Suppose that each voter receives an independent effective message. Then, there exist preference profiles and error probabilities such that, the politician prefers to target approval by majority without noise and approval by unanimity with noise, i.e., formally,*

$$\hat{v}(\pi_2^0) > \hat{v}(\pi_3^0) \textit{ and } \hat{v}(\pi_2^\varepsilon) < \hat{v}(\pi_3^\varepsilon) \textit{ for some } \varepsilon > 0.$$

The proof of the previous result is relegated to Appendix D, and is constructive. In particular, the specifications that we use in our (existence) example are: $\mu_1^* = 0.5$, $\mu_2^* = 0.6$ and $\mu_3^* = 0.7$, together with $v(2) = v(3) = 1$ and $\varepsilon = 0.1$. Notably, in the presence of small noise, the politician will target approval with unanimity although she is indifferent between winning the election with simple majority and winning with unanimous approval. Furthremore, unlike the previous case (with a common effective message), noise does not lead to a less noisy signal. In particular, the optimal signal with noise (viz., π_3^ε) is not less informative than the optimal signal without noise (viz., π_2^0). In fact the two are not comparable in Blackwell's sense, as both posteriors induced by π_3^ε are larger than the respective posteriors induced by π_2^0.

The general conclusion from combining the previous two results is that, *the politician's optimal strategy (over which voters to target) often depends on the source of the distortions.* For instance under the assumptions of Proposition 4, when $\varepsilon > 0$, the politician targets approval by majority (viz., the signal π_2^ε) if noise models the voters' own misunderstanding, and she targets approval by unanimity (viz., the signal π_3^ε) if noise models campaign mistakes. In other words, the politician adopts a more aggressive campaign when the voters misunderstand the political experiment. Notably, this preference reversal holds although the channel that describes the noise is the same in the two cases.

7.2. Persuading biased receivers

It is well-documented that people often appear to be biased in their judgments (e.g., Tversky and Kahneman, 1973, 1983). While this phenomenon has been traditionally attributed to the use of heuristics (for a review, see Gigerenzer and Gaissmaier, 2011), within the psychology literature there is an alternative explanation based on noise (e.g., Costello and Watts, 2014; Hilbert, 2012). Accordingly, agents store information in their memory and they retrieve it when they want to use it, e.g., when they are asked to report their probabilistic assessment or when they need to choose an action. However, information storage is noisy, in the sense that the information may have been deleted from their memory (with some small probability) by the time they try to retrieve it. In what follows, we show that the same idea can be used in the context of noisy persuasion, thus explaining cognitive biases by means of different forms of distortions, beyond memory loss. We illustrate our argument for a standard type of bias (viz., conservatism bias), but we confidently conjecture that similar explanations can be provided for other standard biases.

According to *conservatism bias*, people anchor their beliefs to their prior and update only partially (using Bayes rule). Formally, given an experiment, an agent (viz., the receiver in our case) updates his beliefs with probability $1-\delta$ and sticks to the prior with probability $\delta \in (0,1)$. In other words, the agent completely disregards the experiment with probability δ. Let us illustrate how this can occur in the context of our motivating example (Sections 2 and 6), while keeping in mind that we can directly extend our analysis to a general setting of noisy Bayesian persuasion.

Consider the binary state space $\Omega = \{G, I\}$ with a common prior $\mu_0 \in (0,1)$, and a noisy channel (p) with different input and output message spaces:

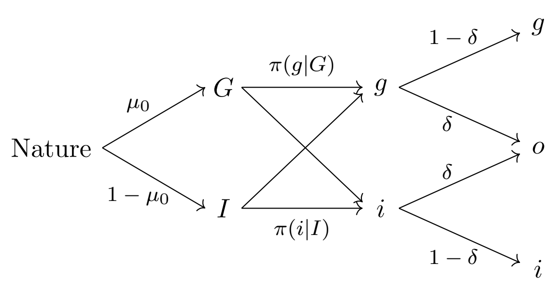


Figure 3: Noisy channel yielding conservatism bias in the prosecutor-judge example.

As we have already discussed, such a channel can be obtained by combining two partitional channels with a common input and output message space (Proposition E1). In Section E.2.2 we present one example of two such partitional channels.[9]

[9] In a nutshell the underlying idea is that messages can be decomposed into a verbal and a non-verbal part. All the relevant information is contained in the verbal part. However, the verbal part is only taken into account when coupled

The channel p has a very interesting property: for every signal, the posterior μ_o coincides with the prior μ_0. The probability of receiving o, and a fortiori forming the posterior μ_o, is equal to δ irrespective of the signal. This is the probability of the receiver disregarding the experiment, and in this sense it can be seen as a measure of the *conservatism bias*. In this case we say that p is a δ-conservative channel.

Proposition 5. *For a δ-conservative channel p, the value of the optimal signal is equal to*

$$\hat{v}_p^* = \delta \hat{v}_0(\mu_0) + (1-\delta)\hat{v}^*.$$

Interestingly, the optimal signal does not depend on δ, i.e., irrespective of how large δ is, the sender will always choose the optimal signal of the benchmark noiseless game. Thus, despite being more difficult to persuade a conservative receiver, the sender does not choose a more informative signal. This is because, there is nothing she can do if o is drawn. Thus, she designs her experiment by taking only the effective messages s and i into account. This is in contrast to models of persuasion where the sender chooses a more informative signal in response to the receiver having strategically made persuasion more difficult (Tsakas et al., 2017).

7.3. Other applications

In this section we briefly present a (non-exhaustive) list of additional applications of our noisy persuasion game.

Example 3. A doctor (the sender) chooses which medical tests to run. The results are truthfully communicated to the patient (the receiver) who decides which treatment to follow. Communication is often noisy due to language barriers or lack of understanding of medical terms on the receiver's side. Such sources of noise have been empirically documented and discussed in highly-influential medical journals (e.g., see Flores, 2006). Then it follows from our previous analysis that, if the preferences of the two agents are not aligned (e.g., if the doctor prefers the patient to choose an expensive treatment while the patient prefers the most effective treatment), combining more than one distortions could be beneficial for the doctor as opposed to removing one of the two. For instance, if the patient does not understand medical terms and at the same time is a non-native speaker, the doctor may prefer not to hire an interpreter in order to remove the second source of noise. ◁

Example 4. In a variant of our motivating example, the prosecutor conducts the investigation and communicates the outcomes to the judge via a mediator, e.g., her deputy is the one scheduled to appear in front of the judge to testify. The use of mediator adds an additional source of noise in

with specific non-verbal parts, e.g., when the prosecutor appears trustworthy or when she is not very aggressive.

the persuasion game. Notice that unlike the literature on cheap-talk games (e.g., see Ivanov, 2010; Ambrus et al., 2013), in this case the mediator is non-strategic and is only relevant to the extent that his presence introduces additional distortions. Once again following our monotonicity analysis, if there are other sources of data distortions, the prosecutor may prefer to use intermediation over direct communication. ◁

Example 5. A marketing campaign is designed by a firm in order to promote a new product. Similarly to the application on voting, the firm attempts to persuade multiple consumers to become regular customers. Persuasion is noisy due to various factors, e.g., the consumers disregard or do not understand the message of the campaign with some probability. Similarly to Section 7.1, depending on the source of noise, different consumers may receive a common effective signal or independent ones. Then, similarly to our earlier analysis, such distinction may have implications for the optimal signal, and in particular for the coalition of consumers that the firm will choose to target. ◁

8. Discussion

8.1. Modelling assumptions about the channel

Throughout the paper our basic underlying assumptions are that the sender chooses a *single signal* over a *given message space*. Are these assumptions natural for a model of noisy persuasion?

The fact that we exogenously *fix the message space* is in contrast to most of the literature on Bayesian persuasion. Indeed, with the exception of Le Treust and Tomala (2018), who also assume a fixed message space, every other paper allows for any finite message space. Let us explain the difference. Define the set Θ containing all possible experimental outcomes of all possible experiments that can be run. The agents cannot necessarily distinguish between any two different experimental outcomes, due to physical and technological restrictions. Such indistinguishable outcomes are then bundled together into the same equivalence class. We assume that these equivalence classes form a finite partition $\mathcal{P}$ of Θ. This is the difference to the rest of the literature, viz., they let $\mathcal{P}$ be countable. Back to our case, the partition $\mathcal{P}$ is identified by our set of messages S, i.e., each $P \in \mathcal{P}$ is given a name in the set S. This name does not have a meaning, viz., it is merely an enumeration of P. It is important to stress that S (and respectively $\mathcal{P}$) does not semantically represent a natural syntactic language. Instead, each $s \in S$ will only obtain meaning once some signal has been chosen. Note that, although finite, S is typically large. In fact we assume it to be at least large enough so that the most informative signal can be constructed, i.e., formally, recall that we have assumed $|S| \geq |\Omega|$. Overall, we find our assumption (of $\mathcal{P}$ being finite) to have a natural interpretation in terms of technological restrictions (e.g., experimental measurements) and/or physical constraints (e.g., bounded perception of the agents). In either case, this assumption does not have important implications for the applications that we study.

Let us now elaborate on our assumption that a *single message* is chosen by the sender. Actually our model is more permissive than it initially seems. The underlying idea is that we can in principle allow the sender to conduct a sequence of experiments, but the number of these experiments is bounded from above by some $n \in \mathbb{N}$, e.g., due to time and/or capacity restrictions. Then, the different signals can be bundled into a single experiment, with the finite message space now containing sequences of messages (of length n). Since the space of such sequences is fixed (by the fact that both S and n are exogenously given), we are back to our model.

8.2. Preferences over basic channels

We ask the following question: If the sender can choose the noisy channel from some set of basic channels, can we predict her choice, without knowing the preference profile? The answer is *sometimes*.

First, as we have already discussed (Section 5), the Blackwell order is complete in the set of *partitional channels*, i.e., for two partitional channels (p and q), it is the case that $p \succeq q$ if and only if $|\mathcal{T}_p| \geq |\mathcal{T}_q|$. Thus, the sender inherits a complete preference relation over the partitional channels, and therefore we can predict that the sender will choose the most informative partitional channel that is available, irrespective of the two agents' preference profile.

As it turns out, the same is not true for *canonical channels*. For instance, consider two channels p and q with transition matrices

$$P = \frac{1}{12}\begin{pmatrix} 7 & 3 & 2 \\ 3 & 8 & 1 \\ 2 & 1 & 9 \end{pmatrix} \text{ and } Q = \frac{1}{768}\begin{pmatrix} 413 & 195 & 160 \\ 195 & 541 & 32 \\ 160 & 32 & 576 \end{pmatrix}$$

respectively. The two channels are canonical and it is straightforward to confirm that neither of them is a garbling of the other. Thus, the preference relation over canonical channels that the sender inherits from the Blackwell order, is not complete, i.e., there are canonical channels that are Blackwell incomparable, and therefore one channel can be better under some preference profile, and the other better under some other profile. Hence, we cannot predict which one will be chosen, unless we know the agents' preference profile.

8.3. Persuasion with heterogeneous priors

A straightforward consequence of noise is that the sender and the receiver will often end up with different posterior beliefs. This is for instance the case when the noisy channel models communication errors from the sender (who is also assumed to be the data collector) and the receiver. Ours is not the only paper on Bayesian persuasion with this feature. In fact, the same is true in persuasion games with heterogeneous priors (e.g., see Alonso and Câmara, 2016; Laclau and Renou, 2017). However, in this last case disagreement arises in different forms compared to our model with noise. Furthermore, neither of the two models is a special case of the other, i.e., the set of feasible distributions of posteriors in one case is not a subset of those that can be achieved in the other.

A. Proofs of Section 3

A.1. Intermediate results

Lemma A1. *For an arbitrary channel p, there is a finite set $B \subseteq \Delta(S)$ such that $\Sigma_p = (\Delta(B))^\Omega$.*

PROOF. Define $B := \{b_1, \ldots, b_K\}$ such that $b_k(s) := p(s|s_k)$ for every $s \in S$ and every $k \in \{1, \ldots, K\}$. Clearly, observe that $b_k \in \Delta(S)$ for every k, while also noticing that different b_k's may coincide. Take arbitrary $\omega \in \Omega$ and $\pi \in \Pi$. Then, for every $s \in S$,

$$\sigma_\pi(s|\omega) = \sum_{k=1}^{K} p(s|s_k)\pi(s_k|\omega) = \sum_{k=1}^{K} b_k(s)\pi(s_k|\omega),$$

implying that $\sigma_\pi(\cdot|\omega) \in \Delta(B)$, and therefore $\Sigma_p \subseteq (\Delta(B))^\Omega$. Now, take an arbitrary $\sigma \in (\Delta(B))^\Omega$, implying that for each $\omega \in \Omega$ there exists some $(a_1^\omega, \ldots, a_K^\omega) \in \mathbb{R}_+^K$ with $\sum_{k=1}^K a_k^\omega = 1$, such that $\sum_{k=1}^K a_k^\omega b_k = \sigma(\cdot|\omega)$. Then, define $\pi \in \Pi$ by $\pi(s_k|\omega) := \alpha_k^\omega$, and observe that by construction $\sigma = \sigma_\pi$. Hence, $\Sigma_p \supseteq (\Delta(B))^\Omega$, thus completing the proof. $\square$

Definition A1. For an arbitrary $s \in S$, define the convex set

$$\hat{\Sigma}_s = \left\{ \sigma \in \Sigma : \sum_{\omega \in \Omega} \sigma(s|\omega) > 0 \right\}. \tag{A.1}$$

Then, we define the function $\hat{\mu}_s : \hat{\Sigma}_s \to M_s$ by

$$\hat{\mu}_s(\sigma)(\omega) = \frac{\mu_0(\omega)\sigma(t|\omega)}{\mathbb{E}_0[\sigma(t|\cdot)]}, \tag{A.2}$$

for each $\omega \in \Omega$.

Lemma A2. *M_s is convex and compact.*

Proof. Convexity. Notice that $\hat{\mu}_s$ is a fractional-linear function. Hence, since $\hat{\Sigma}_s$ is convex, the image $\hat{\mu}_s(\hat{\Sigma}_s)$ is also convex (Boyd and Vandenberghe, 2004, p.42). But then, observe that $\hat{\mu}_s$ is surjective, viz., $\hat{\mu}_s(\hat{\Sigma}_s) = M_s$, and therefore M_s is convex.

Compactness. Take an arbitrary sequence $(\tilde{\mu}_s^k)_{k=1}^{\infty}$ in M_s converging to some $\mu_s \in \mathbb{R}^{\Omega}$. Notice that $\mu_s \in \Delta(\Omega)$ since $\tilde{\mu}_s^k$ is a sequence in the closed set $\Delta(\Omega)$. We are going to prove that $\mu_s \in M_s$, which will then imply that M_s is closed, and hence compact.

Step 1. Since μ_s is the limit of a sequence in M_s, it belongs to $\text{clos}(M_s)$. Moreover, since M_s is convex – by the previous part of our result – its closure will also be convex. Now, for every $\beta \in (0,1]$ take $\beta\mu_0 + (1-\beta)\mu_s$, which belongs to M_s. Moreover, if we take any strictly decreasing sequence $(\beta_k)_{k=1}^{\infty}$ in $(0,1]$ that converges to 0, we obtain a sequence $(\mu_s^k)_{k=1}^{\infty}$ in M_s, defined by $\mu_s^k := \beta_k\mu_0 + (1-\beta_k)\mu_s$, which also converges to μ_s. Hence, it suffices to prove that μ_s^k (instead of $\tilde{\mu}_s^k$) converges to a point in M_s.

Step 2. For each $k \geq 1$, since $\mu_s^k \in M_s$, there exists some $\sigma^k \in \hat{\Sigma}_s$ such that $\hat{\mu}_s(\sigma^k) = \mu_s^k$. This implies that the effective signal σ^k satisfies

$$\beta_k\mu_0(\omega) + (1-\beta_k)\mu_s(\omega) = \frac{\mu_0(\omega)\sigma^k(s|\omega)}{\mathbb{E}_0[\sigma^k(s|\cdot)]}$$

for every $\omega \in \Omega$. In other words, $\sigma^k(s|\cdot) \in [0,1]^{\Omega}$ is a non-trivial solution to the system of linear equations

$$\big(\beta_k\mu_0(\omega) + (1-\beta_k)\mu_s(\omega)\big)\sum_{\omega' \in \Omega}\mu_0(\omega')x_{\omega'}^k - \mu_0(\omega)x_{\omega}^k = 0. \tag{A.3}$$

This system is homogeneous. Hence, since it has one non-trivial solution, it will have infinitely many.

Step 3. It follows from the previous step that, for every $k \geq 1$, there exists at least one $\omega \in \Omega$ such that x_{ω}^k is a free variable. Hence, it follows from Ω being finite that there exists some $\omega \in \Omega$ such that x_{ω}^k is a free variable for infinitely many $k \in \mathbb{N}$. This obviously defines a subsequence of μ_s^k. Now for each μ_s^k in this subsequence, we can pick a solution $\hat{\sigma}^k(s|\cdot)$ to the system (A.3) such that $\hat{\sigma}^k(s|\omega) = A_{\omega} > 0$, where A_{ω} is a constant that does not depend on k. Hence, we obtain a sequence $(\hat{\sigma}^k)_{k=1}^{\infty}$ in M_s with $\hat{\sigma}^k(s|\omega)$ being constant. Now, since Σ is a polytope, it is bounded, and therefore, by the Bolzano-Weierstrass theorem, there exists a subsequence of $\hat{\sigma}^k$ that converges to some $\sigma \in \Sigma$. However, notice that $\sigma(s|\omega) = \lim_{k\to\infty}\hat{\sigma}^k(s|\omega) = A_{\omega} > 0$, implying that $\sigma \in \hat{\Sigma}_s$, and therefore $\hat{\mu}_s(\sigma) \in M_s$. But then again, $\hat{\mu}_s$ is continuous in $\hat{\Sigma}_s$, implying that

$$\hat{\mu}_s(\sigma) = \lim_{k\to\infty}\hat{\mu}_s(\hat{\sigma}^k) = \lim_{k\to\infty}\mu_s^k = \mu_s,$$

which completes the proof. □

A.2. Proof of Proposition 1: Equilibrium existence

It follows from Lemma A2 that $\Delta(M)$ is compact, endowed with the topology of weak convergence. Define the mapping $\hat{\tau} : \Sigma \to \Delta(M)$, by $\hat{\tau}(\sigma_\pi) := \tau_\pi$.

Step 1. We are going to prove that $\hat{\tau}$ is continuous. Consider an arbitrary sequence $(\sigma^k)_{k=1}^\infty$ in Σ converging to some $\sigma \in \Sigma$. Then, it suffices to prove that $\hat{\tau}(\sigma^k) \overset{w^*}{\to} \hat{\tau}(\sigma)$. Thus, by the Portmanteau Theorem (see Aliprantis and Border, 1994, Thm. 15.3), it suffices to prove that for every continuous $f : M \to \mathbb{R}$, it is the case that $\int_M f(\mu) d\tau_k(\mu) \to \int_M f(\mu) d\tau(\mu)$, where $\tau_k := \hat{\tau}(\sigma^k)$ and $\tau := \hat{\tau}(\sigma)$. Since both σ^k and σ belong to Σ, both τ_k and τ have finite support. Hence, it suffices to prove that

$$\sum_{\omega \in \Omega} \mu_0(\omega) \sum_{s \in S} \big(\sigma^k(s|\omega) f(\mu_s^k) - \sigma(s|\omega) f(\mu_s) \big) \longrightarrow 0 \tag{A.4}$$

for an arbitrary continuous $f \in \mathbb{R}^M$, where $\mu_s^k \in \hat{\Sigma}_s$ is the receiver's posterior if s is observed given some effective signal σ^k with $\sigma^k(s|\omega) > 0$ for some $\omega \in \Omega$. If $\sigma^k(s|\omega) = 0$ for all $\omega \in \Omega$, then μ_s^k is chosen arbitrarily. Likewise we define $\mu_s \in M_s$. Now, let us consider two cases.

(a) $\sigma \in \hat{\Sigma}_s$: By continuity of $\hat{\mu}_s$ it follows that $\mu_s^k \to \mu_s$, and by continuity of f it follows that $f(\mu_s^k) \to f(\mu_s)$, which directly implies (A.4).

(b) $\sigma \notin \hat{\Sigma}_s$: By Lemma A2, the function f is continuous on a compact domain and therefore $f(\mu_s)$ is bounded. Hence, $\sigma(s|\omega) f(\mu_s) = 0$. Moreover, by $\sigma^k \to \sigma$, it follows that $\sigma^k(s|\omega) f(\mu_s^k) \to 0$, which again implies (A.4).

Combining the previous two cases proves our claim that $\hat{\tau}$ is continuous.

Step 2. Since Σ is a polytope (and therefore compact), its continuous image $\mathcal{M} := \hat{\tau}(\Sigma)$ is compact in $\Delta(M)$. Moreover, recall that $\mathbb{E}_\bullet[\hat{v}_0] : \mathcal{M} \to \mathbb{R}$ is upper semi-continuous (Kamenica and Gentzkow, 2011). Hence, it achieves a maximum $\tau^* \in \mathcal{M}$. Finally, by construction, it follows that $\pi^* \in \arg\max_{\pi \in \Pi} \hat{v}(\pi)$ if $\tau_{\pi^*} = \tau^*$, implying that there exists an optimal signal.

A.3. Proof of Proposition 2: Value of isometric channels

Take an arbitrary permutation matrix R. Then it suffices to prove that both PR and RP yield the same value as P. Since R is a permutation matrix, there exists some bijection $\rho : \{1, \ldots, K\} \to \{1, \ldots, K\}$ that associates each row with the column where 1 appears, i.e., if $\rho(k) = \ell$ then $P_{k,\ell} = 1$.

Step 1. Let $Q = PR$, i.e., Q is obtained by permuting the columns of P. In particular, the k-th column of P is the $\rho(k)$-th column of Q. Then, for every $\pi \in \Pi$ and every $\mu \in \Delta(\Omega)$,

$$\begin{aligned}\tau^p_\pi(\mu) &= \sum_{k=1}^{K} p(\{t \in S : \mu_t = \mu\}|s_k)\mathbb{E}_0[\pi(s_k|\cdot)] \\ &= \sum_{k=1}^{K} q(\{t \in S : \mu_t = \mu\}|s_{\rho(k)})\mathbb{E}_0[\pi(s_{\rho(k)}|\cdot)] \\ &= \tau^q_\pi(\mu),\end{aligned}$$

where τ^p_π and τ^q_π are the respective distributions of posteriors for the two channels, when π is chosen by the sender. Hence, $\mathbb{E}_{\tau^p_\pi}[\hat{v}_0] = \mathbb{E}_{\tau^q_\pi}[\hat{v}_0]$, thus implying $\hat{v}^*_p = \hat{v}^*_q$.

Step 2. Let $Q = RP$, i.e., Q is obtained by permuting the rows of P. In particular, the k-th row of P is the $\rho^{-1}(k)$-th row of Q. Now, for an arbitrary $\pi \in \Pi$ define $\tilde{\pi} \in \Pi$ by $\tilde{\pi}(s_k|\omega) := \pi(s_{\rho^{-1}(k)}|\omega)$, and observe that for every $\pi \in \Pi$ and every $\mu \in \Delta(\Omega)$,

$$\begin{aligned}\tau^p_\pi(\mu) &= \sum_{s \in S} p(\{t \in S : \mu_t = \mu\}|s)\mathbb{E}_0[\pi(s|\cdot)] \\ &= \sum_{s \in S} q(\{t \in S : \mu_t = \mu\}|s)\mathbb{E}_0[\tilde{\pi}(s|\cdot)] \\ &= \tau^q_{\tilde{\pi}}(\mu).\end{aligned}$$

Hence, $\mathbb{E}_{\tau^p_\pi}[\hat{v}_0] = \mathbb{E}_{\tilde{\tau}^q_\pi}[\hat{v}_0]$, thus implying $\hat{v}^*_p = \hat{v}^*_q$.

B. Proof of Theorem 1: Optimal signal

Proof of (i): Recall that $\mathcal{M} := \{\tau_\pi \mid \pi \in \Pi\}$. Then, it is obviously the case that $\mathcal{M} \subseteq \Delta(M)$. Therefore, we obtain

$$\hat{v}^*_p = \max_{\pi \in \Pi} \hat{v}(\pi) = \max_{\tau \in \mathcal{M}} \mathbb{E}_\tau[\hat{v}_0] \leq \max_{\tau \in \Delta(M)} \mathbb{E}_\tau[\hat{v}_0] = V_M(\mu_0).$$

Proof of (ii): Obviously, $\hat{v}_0(\mu_0)$ is equal to the sender's expected utility if she chooses the completely uninformative signal $[\mu_0]$ that puts probability 1 to μ_0. Hence, by the previous step,

$$\hat{v}([\mu_0]) = \hat{v}_0(\mu_0) = V_M(\mu_0) \geq \hat{v}^*_p \geq \hat{v}(\pi),$$

for every $\pi \in \Pi$, thus completing the proof.

C. Proofs of Section 5: Monotonicity

C.1. Intermediate results

Lemma C3. *Let P and Q be two doubly stochastic matrices such that P is nonsingular. Furthermore, let R be some stochastic matrix such that $Q = PR$. Then, R is doubly stochastic.*

PROOF. Since P is nonsingular, there exists a square (inverse) matrix B such that $PB = BP = I$, where I is the identity matrix. By $BP = I$ it follows that $\sum_{k=1}^{K} B_{n,k}P_{k,n} = 1$ and also $\sum_{k=1}^{K} B_{n,k}P_{k,m} = 0$ for all $m \neq n$. Thus, using the fact that P is doubly stochastic, we obtain

$$1 = \sum_{m=1}^{K}\sum_{k=1}^{K} B_{n,k}P_{k,m} = \sum_{k=1}^{K} B_{n,k}. \tag{C.1}$$

Now, multiply both sides of $Q = PR$ with B (from the left) to obtain $R = BQ$, thus implying

$$R_{n,m} = \sum_{k=1}^{K} B_{n,k}Q_{k,m}. \tag{C.2}$$

Since R is by hypothesis (row) stochastic, it suffices to prove that the entries of each column sum up to 1. Indeed,

$$\sum_{n=1}^{K} R_{n,m} = \sum_{n=1}^{K}\sum_{k=1}^{K} B_{n,k}Q_{k,m} = \sum_{k=1}^{K} Q_{k,m} = 1, \tag{C.3}$$

with (C.3) following from (C.1) and the fact that Q is doubly stochastic. □

Lemma C4. *Let P be isometric to a symmetric matrix $\tilde{P}$. Then, there is a permutation matrix I_p such that $\hat{P} := PI_p$ is symmetric, i.e., we can retrieve a symmetric matrix (perhaps other than $\tilde{P}$) by only permuting rows.*

PROOF. By hypothesis there are permutation matrices I_1 and I_2 such that $P = I_1\tilde{P}I_2$. Then, define the matrix $\hat{P} := I_2'I_1'P = I_2'\tilde{P}I_2$, where the transposes I_1' and I_2' are also the inverses of I_1 and I_2 respectively, as these are permutation matrices. Importantly, $\hat{P}$ is obtained by permuting (only) columns of P. Finally, by $\tilde{P}$ being symmetric, so is $I_2'\tilde{P}I_2$, and therefore the same is true for $\hat{P}$, thus completing the proof. □

C.2. Proof of Theorem 2

PROOF OF (I): By symmetry both P and Q are doubly stochastic. Moreover, by diagonal dominance P is nonsingular (Levy-Desplanques Theorem). Hence, by Lemma C3, it follows that R is doubly stochastic too. Furthermore, by Lemma C4, there exist permutation matrices I_p and I_q such that $\hat{P} := PI_p$ and $\hat{Q} := QI_q$ are symmetric. Define $\hat{R} := I_p' R I_q$, to obtain

$$\hat{Q} = QI_q = (PI_p)(I_p' R I_q) = \hat{P}\hat{R}. \tag{C.4}$$

Since R is doubly stochastic, so is $\hat{R}$. Therefore, every column of $\hat{Q}$ can be written as a convex combination of columns of $\hat{P}$. Formally, there is some $(\alpha_1, \ldots, \alpha_K) \in \mathbb{R}_+^K$ with $\sum_{k=1}^K \alpha_k = 1$, such that for every $\ell \in \{1, \ldots, K\}$,

$$\hat{Q}_{\ell,k} = \sum_{k=1}^K \alpha_k \hat{P}_{\ell,k}. \tag{C.5}$$

Since $\hat{P}$ and $\hat{Q}$ are symmetric, every column vector is also a row vector in each of them. Hence, every row of $\hat{Q}$ can also be written as a convex combination of rows of $\hat{P}$. Formally,

$$\hat{Q}_{k,\ell} = \sum_{k=1}^K \alpha_k \hat{P}_{k,\ell}. \tag{C.6}$$

Therefore, every extreme point of $\Sigma_{\hat{q}}$ belongs to $\Sigma_{\hat{p}}$, and thus obviously $\Sigma_{\hat{q}} \subseteq \Sigma_{\hat{p}}$, implying $\hat{v}^*_{\hat{q}} \leq \hat{v}^*_{\hat{p}}$. Finally, since P and Q are isometric with $\hat{P}$ and $\hat{Q}$ respectively, we obtain $\hat{v}^*_q = \hat{v}^*_{\hat{q}} \leq \hat{v}^*_{\hat{p}} = \hat{v}^*_p$.

Proof of (ii): Let $\mathcal{T}_p$ and $\mathcal{T}_p$ be the partitions of S induced by p and q respectively.

We will first show that, $p \succeq q$ if and only if $|\mathcal{T}_p| \geq |\mathcal{T}_q|$. Sufficiency is trivial. Indeed, if $|\mathcal{T}_p| \geq |\mathcal{T}_q|$ then take a surjective mapping $\rho : \mathcal{T}_p \to \mathcal{T}_q$. For an arbitrary $s \in S$, let $T(s) := \{t \in S : \text{if } s \in T_p \in \mathcal{T}_p \text{ and } t \in T_q \in \mathcal{T}_q \text{ then } \rho(T_p) = T_q\}$. Then, define $r(t|s) := 1/|T(s)|$ for every $t \in T(s)$ and every $s \in S$. Obviously, r is a stochastic matrix thus proving that $|\mathcal{T}_p| \geq |\mathcal{T}_q|$ implies $p \succeq q$. So let us turn to necessity. By definition there is some channel r such $Q = PR$. Assume that there are $T_1, T_2 \in \mathcal{T}_q$ such that $r(T_1|T_0) > 0$ and $r(T_2|T_0) > 0$ for some $T_0 \in \mathcal{T}_p$. Then, by q being partitional, there is some $T \in \mathcal{T}_q$ such that $T_1 \cup T_2 \subseteq T$. Hence, each $T_0 \in \mathcal{T}_p$ can be associated (via r) with at most one $T \in \mathcal{T}_q$, implying that there is a surjection from $\mathcal{T}_p$ to $\mathcal{T}_q$. Therefore, $|\mathcal{T}_p| \geq |\mathcal{T}_q|$.

By the previous step, p allows for more effective messages than q, implying that $\Sigma_q \subseteq \Sigma_p$. Hence, $\hat{v}^*_q \leq \hat{v}^*_p$, which completes the proof.

D. Proofs of Section 7

D.1. Proof of Proposition 4

We prove the result constructively. First, take $\mu_1^* = 0.5$, $\mu_2^* = 0.6$ and $\mu_3^* = 0.7$ and assume that $v(2) = v(3) = 1$. Note that without noise, we are back to the previous case of every voter receiving the same message (viz., the realized actual message), and therefore the politician will target approval by majority (viz., $\hat{v}(\pi_2^0) > \hat{v}(\pi_3^0)$), as $v(2) = v(3)$. Now suppose that $\varepsilon = 0.1$. Then, for each signal $\pi \in \{\pi_2^\varepsilon, \pi_3^\varepsilon\}$, a common actual message $s \in S$ is drawn, and three (perhaps different) effective messages are independently drawn from $p(\cdot|s)$, one for each voter. For notation simplicity and without loss of generality, let s_1 be the effective message that yields the good posterior and s_2 be the effective message yielding the bad posterior. When π_2^ε is chosen, the total probability of the proposal receiving at least two votes (which is in fact equal the probability of receiving exactly two votes) becomes

$$P_2^\varepsilon = \pi_2^\varepsilon(s_1)(1-\varepsilon)^2 + \pi_2^\varepsilon(s_2)\varepsilon^2,$$

where $\pi_i^\varepsilon(s) = (\pi_i^\varepsilon(s|\omega_1) + \pi_i^\varepsilon(s|\omega_2))/2$ for each $s \in S$ and each $i \in I$. On the other hand, when π_3^ε is chosen, the total probability of the proposal receiving at least two votes is equal to

$$P_3^\varepsilon = \pi_3^\varepsilon(s_1)\big(3(1-\varepsilon)^2\varepsilon + (1-\varepsilon)^3\big) + \pi_3^\varepsilon(s_2)\big(3(1-\varepsilon)\varepsilon^2 + \varepsilon^3\big).$$

Intuitively, when the politician targets approval by majority, she must persuade 1 and 2. On the other hand, when she targets approval by unanimity, any coalition of at least two voters would work. Then simple algebra reveals that $P_2^\varepsilon \approx 0.66$ and $P_3^\varepsilon \approx 0.67$, implying $\hat{v}(\pi_2^\varepsilon) < \hat{v}(\pi_3^\varepsilon)$, which completes the proof.

D.2. Proof of Proposition 5

Let us first prove that for every signal π, the receiver does not update his prior upon observing the message o. Indeed, for every $\omega \in \Omega$,

$$\mu_o(\omega) = \frac{\mu_0(\omega)\big(\delta\pi(g|\omega) + \delta\pi(i|\omega)\big)}{\sum_{\omega' \in \Omega} \mu_0(\omega')\big(\delta\pi(g|\omega') + \delta\pi(i|\omega')\big)} = \mu_0(\omega).$$

Hence, upon observing o the judge will always acquit the defendant, regardless of the signal, and the prosecutor will receive $\hat{v}_0(\mu_0)$. This will happens with probability δ.

Then, we prove that for every signal π, the receiver's posterior given the effective message $s \in \{g, i\}$ is equal to the posterior that he would have formed under π in the noiseless game. Indeed, for every $\omega \in \Omega$ and every $s \in \{s, i\}$,

$$\mu_s(\omega) = \frac{\mu_0(\omega)(1-\delta)\pi(s|\omega)}{\sum_{\omega' \in \Omega} \mu_0(\omega')(1-\delta)\pi(s|\omega')} = \mu_s(\omega).$$

Note that some $s \in \{g, i\}$ will be observed by the receiver with probability $1 - \delta$, thus completing the proof.

E. Supplementary material

E.1. Additional results

Proposition E1. *For each channel $q : S_1 \to \Delta(S_2)$ with $q(s_2|s_1) \in \mathbb{Q}$ for all $s_1 \in S_1$ and $s_2 \in S_2$, there exist two partitional channels $p : S \to \Delta(S)$ and $r : S \to \Delta(S)$, such that q is obtained by combining p and r, i.e., fomally, q is a garbling of p via r.*

Proof. Define the partition $\mathcal{S}_2$ of S_2 by bundling together messages that yield the same posterior belief for every signal under q, and let $K_1 := \prod_{U_2 \in \mathcal{S}_2} |U_2|$. Then, define $K := |S_1|K_1$, and take $S := \{s_1, \ldots, s_K\}$. Now, define the partition $\mathcal{T}_1$ by splitting S into $|S_1|$ subsets with K_1 messages in each one of them. Each $T_1 \in \mathcal{T}_1$ corresponds to one $s_1 \in S_1$. Then, construct a second partition $\mathcal{T}_2$, containing $|\mathcal{S}_2|$ subsets of S. Each $T_2 \in \mathcal{T}_2$ corresponds to one $s_2 \in S_2$. Specifically, the subset $T_2 \in \mathcal{T}_2$ that corresponds to s_2 will contain $q(s_2|s_1)K_1$ elements of the $T_1 \in \mathcal{T}_1$ that corresponds to $s_1 \in S_1$. Then, let p and r be the partitional channels characterized by $\mathcal{T}_1$ and $\mathcal{T}_2$ respectively, thus completing the proof. $\square$

Proposition E2. *$p \succeq q$ if and only if $\hat{v}_p^* \geq \hat{v}_q^*$ for every pair of aligned utility functions $u = v$.*

Proof. The proof is a direct application of Blackwell's Theorem (e.g., see Blackwell, 1951, 1953; Perez-Richet, 2016). In particular, similarly to our setting, let Ω be a finite state space, $\mu_0 \in \Delta(\Omega)$ be a prior, S be a finite set of messages, A be a compact set of actions and $u : A \times \Omega \to \mathbb{R}$ be a continuous utility function. An experiment is a function $\sigma : \Omega \to \Delta(S)$. The value of an experiment is given by

$$V(\sigma, u) := \max \left\{ \phi \in \mathbb{R} : \text{ there is } c : S \to A \text{ such that } \phi = \sum_{s \in S} \mathbb{E}_0[\sigma(s|\cdot)u(c(s), \cdot)] \right\}.$$

Then, Blackwell's Theorem states that for two experiments σ_1, σ_2 it is the case that, $\sigma_1 \succeq \sigma_2$ if and only if $V(\sigma_1, u) \geq V(\sigma_2, u)$ for every utility function u. Now turning back to our case notice that whenever $u = v$, it is the case that

$$\hat{v}_p^* = \max_{\pi \in \Pi} V(\pi \circ p, u)$$

Therefore, by Blackwell's Theorem, $p \succeq q$ is equivalent to $V(\pi \circ p, u) \geq V(\pi \circ p, u)$ for all $\pi \in \Pi$, which in turn is equivalent to $\hat{v}_p^* \geq \hat{v}_q^*$. $\square$

Proposition E3. *For diagonally dominant channels p and q, if $p \sim q$ then $\hat{v}_p^* = \hat{v}_q^*$.*

Proof. By $p \succeq q$ and $q \succeq p$, there are channels r_1 and r_2 such that $PR_1 = Q$ and $QR_2 = P$. Right-multiply both sides of the first equation by R_2 and of the second one by R_1, to obtain $P(R_1R_2 - I) = 0$

and $Q(R_2R_1-I)=0$. Since P and Q are diagonally dominant, they are both invertible (by the Levy-Desplanques Theorem), implying that $R_1R_2 = R_2R_1 = I$, and therefore $R_2 = R_1^{-1}$. Since R_1 and R_2 are stochastic matrices, they are permutation matrices (Mailath and Samuelson, 2006). Therefore, p and q are isometric. Hence, by Proposition 2, the respective values are equal, i.e., $\hat{v}_p^* = \hat{v}_q^*$. $\square$

E.2. Examples and applications revisited

E.2.1. Monotonicity: Example 2

Recall the two partitional channels p and r, respectively characterized by the partitions

$$\begin{aligned}
\mathcal{T}_p &= \{ \underbrace{\{s_1,\dots,s_{10}\}}_{T_1}, \underbrace{\{s_{11},\dots,s_{20}\}}_{T_2} \},\\
\mathcal{T}_r &= \{ \underbrace{\{s_1,s_2,s_{11},s_{12},s_{13}\}}_{T_1'}, \underbrace{\{s_3,s_4,s_5,s_{14},\dots,s_{18}\}}_{T_2'}, \underbrace{\{s_6,\dots,s_{10},s_{19},s_{20}\}}_{T_3'} \}.
\end{aligned}$$

The channel p is effectively equivalent to a noiseless channel with two messages, T_1 and T_2. Intuitively, for each signal $\pi:\Omega\to\Delta(S)$ the distribution of posteriors τ_π depends solely on the probability $\pi(T|\omega)$, and not on how $\pi(\cdot|\omega)$ is distributed across messages in $T\in\mathcal{T}_p$, i.e., formally, if $\pi(T|\omega)=\pi'(T|\omega)$ for all $T\in\mathcal{T}_p$ and all $\omega\in\Omega$, then $\tau_\pi=\tau_{\pi'}$. Hence, the sender can only achieve binary distributions of posteriors, implying that without loss of generality we can simply write $\pi:\Omega\to\Delta(\mathcal{T}_p)$. On the other hand, when we mix p with r, the induced channel q is not partitional. In fact, q is essentially equivalent to a channel with different input and output message space, viz., $\mathcal{T}_p$ and $\mathcal{T}_r$ respectively. Intuitively, it is still the case that the distributions of posteriors are identified by $\pi:\Omega\to\Delta(\mathcal{T}_p)$. However, when $T\in\mathcal{T}_p$ is realized, the receiver observes each $T'\in\mathcal{T}_r$ with probability $r(T'|T)=|T\cap T'|/|T|$. Hence, the sender's distribution of posteriors can put positive probability to three posteriors, one for each $T'\in\mathcal{T}_r$. Indeed, there is a signal π such that the receiver forms a different posterior for each $T'\in\mathcal{T}_r$ under q (see Example 2). In particular, τ_π puts positive probability to $\mu_1^q=(\frac{4}{21},\frac{7}{21},\frac{10}{21})$, $\mu_2^q=(\frac{2}{8},\frac{3}{8},\frac{3}{8})$ and $\mu_3^q=(\frac{10}{21},\frac{7}{21},\frac{4}{21})$. Crucially, there is no $\lambda\in\mathbb{R}$ such that $\mu_2^q=\lambda\mu_1^q+(1-\lambda)\mu_3^q$, i.e., the three posteriors are not collinear, implying that μ_0 belongs to the interior of the convex hull of these posteriors. Therefore, under p, at most one of these posteriors will occur with positive probability. Hence, by the fact that $v(\mu)=1$ for each $\mu\in\{\mu_1^q,\mu_2^q\mu_3^q\}$ and $v(\mu)=0$ otherwise, it follows that her expected utility under any Bayes-plausible binary distribution will be lower than the one she achieves when these three posteriors receive probability 1. Thus, q yields a higher value than p.

E.2.2. Persuading biased receivers: Conservatism bias

Recall our motivating prosecutor-judge example. Consider the message space $S = S_1 \times S_2$, where $S_1 := \{g, i\}$ is the verbal part of the message and $S_2 := \{s_1, \ldots, s_M\}$ is the non-verbal part. The verbal part encodes the experimental outcomes, whereas the non-verbal part encodes cues that are unintentionally used by the prosecutor, e.g., how trustworthy the prosecutor appears to be in front of the judge. We now consider two partitional channels q and r which are essentially combined into p. The channel q bundles together all the messages with the same verbal part, i.e., the corresponding partition is $\mathcal{T}_1 = \{T_g, T_i\}$ with $T_g = \{g\} \times S_2$ and $T_i = \{i\} \times S_2$. The channel r is associated with the partition $\mathcal{T}_2 = \{T_0, T_1, T_2\}$ where $T_0 := \{g, i\} \times \{s_1, \ldots, s_m\}$, $T_1 := \{g\} \times \{s_{m+1}, \ldots, s_M\}$ and $T_2 := \{i\} \times \{s_{m+1}, \ldots, s_M\}$, i.e., when a non-verbal part in $\{s_1, \ldots, s_m\} \subseteq S_2$ is realized, the judge fails to distinguish between g and i. This is for instance the case when $\{s_1, \ldots, s_m\}$ corresponds to the non-verbal messages that make the prosecutor look non-trustworthy. In other words, we identify the message space $\{o, g, i\}$ of p with the partition $\mathcal{T}_2$. Then, $\delta = m/M$.

Recall that by Proposition 5, the value of the optimal signal is decreasing in δ, implying that the sender prefers the verbal parts to be taken into account. The interpretation is straightforward and quite intuitive, viz., messages that appear unreliable are harmful for the sender. Indeed, it is harder to persuade people when they do not trust you. The latter explains why lawyers and politicians typically try to appear trustworthy in front of judges and their electorate respectively.

References

Aliprantis, C. & Border, K. (1994). *Infinite dimensional analysis.* Springer Verlag, Berlin.

Alonso, R. & Câmara, O. (2016). Bayesian persuasion with heterogeneous priors. *Journal of Economic Theory* 165, 672–706.

——— (2017). Persuading voters. *American Economic Review* 106, 35-90–3605.

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

Ambrus, A., Azevedo, E. & Kamada, Y. (2013). Hierarchical cheap talk. *Theoretical Economics* 8, 233–261.

Arieli, I. & Babichenko, Y. (2016). Private Bayesian persuasion. *Working Paper.*

Aumann, R.J. (1976). Agreeing to disagree. *Annals of Statistics* 4, 1236-1239.

Aumann, R. & Maschler, M. (1995). *Repeated games with incomplete information.* MIT Press.

Bergemann, D. & Morris, S. (2016a). Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics* 11, 487–522.

——— (2016b). Information design, Bayesian persuasion and Bayes correlated equilibrium. *American Economic Review, Papers and Proceedings* 106, 586–591.

Bizzotto, J., Rudiger, J. & Vigier, A. (2017). How to persuade a long-run decision maker. *Working Paper.*

Blackwell, D. (1951). The comparison of experiments. *Proceedings of the 2nd Berkeley Symposium on Mathematical Statistics and Probability* 93–102.

——— (1953). Equivalent comparisons of experiments. *Annals of Mathematical Statistics* 24, 265–272.

Blume, A. (2012). A class of strategy-correlated equilibria in sender-receiver games. *Games and Economic Behavior* 75, 510–517.

Blume, A. & Board, O.J. (2013). Language barriers. *Econometrica* 81, 781–812.

——— (2014). Intentional vagueness. *Erkenntnis* 79, 855–899.

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

Boyd, S. & Vandenberghe, L. (2004). *Convex optimization.* Cambridge University Press.

Brocas, I. & Carillo, J.D. (2007). Influence through ignorance. *RAND Journal of Economics* 38, 931–947.

Costello, F. & Watts, P. (2014). Surprisingly rational: probability theory plus noise explains biases in judgment. *Psychological Review* 121, 1939–1471.

Cover, T.M. & Thomas, J.A. (2006). *Elements of information theory.* John Wiley & Sons.

Crawford, V. & Sobel, J. (1982). Strategic information transmission. *Econometrica* 50, 1431–1451.

Dessein, W. (2002). Authority and communication in organizations. *Review of Economic Sudies* 69, 811–838.

Flores, G. (2006). Aligned delegation. *American Economic Review* 104, 66–83.

Frankel, A. (2014). Language barriers to health care in the United States. *New England Journal of Medicine* 355, 229–331.

Gennaioli, N. & Ponzetto, G. (2017). Optimally vague contracts and the law. *Working Paper.*

Gentzkow, M. & Kamenica, E. (2016). A Rothschild-Stiglitz approach to Bayesian persuasion. *American Economic Review, Papers and Proceedings* 106, 597–601.

——— (2017). Competition in persuasion. *Review of Economic Studies* 84, 300–322.

Gigerenzer, G. & Gaissmaier, W. (2011). Heuristic decision making. *Annual Review of Psychology* 62, 451–482.

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

Goltsman, M., Hörner, J., Pavlov, G. & Squintani, F. (2009). Mediation, arbitration and negotiation. *Journal of Economic Theory* 144, 1397–1420.

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

Hedlund, J. (2017). Bayesian persuasion by a privately informed sender. *Journal of Economic Theory* 167, 229–268.

Hernández, P. & von Stengel, B. (2014). Nash codes for noisy channels. *Operations Research* 62, 1221–1235.

Hilbert, M. (2012). Towards a synthesis of cognitive biases: how noisy information processing can bias human decision making. *Psychological Bulletin* 138, 211–237.

Ivanov, M. (2010). Communication via a strategic mediator. *Journal of Economic Theory* 145, 869–884.

Kamenica, E. & Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review* 101, 2590–2615.

Koessler, F. (2001). Common knowledge and consensus with noisy communication. *Mathematical Social Sciences* 42, 139–159.

Kolotilin, A., Mylovanov, T., Zapechelnyuk, A. & Li, M. (2018). Persuasion of a privately informed receiver. *Econometrica* (forthcoming).

LACLAU, M. & RENOU, L. (2017). Public persuasion. *Working paper.*

LANDERAS, P. & PÉREZ DE VILLAREAL, J.M. (2005). A noisy screening model of education. *Labour* 19, 35–54.

LE TREUST, M. & TOMALA, T. (2018). Persuasion with limited communication capacity. *Working paper.*

MAKUR, A. & POLYANSKYI, Y. (2016). Comparison of channels: Criteria for domination by a symmetric channel. *Working Paper.*

MAILATH, G. & SAMUELSON, L. (2006). *Repeated games and reputations: Long-run relationships.* Oxford University Press.

MARSHALL, A.W., OLKIN, I. & ARNOLD, B.C. (2011). *Inequalities: Theory of majorization and its applications.* Springer Series in Statistics, 2nd Edition.

MILGROM, P. & ROBERTS, J. (1986). Relying on the information of interested parties. *RAND Journal of Economics* 17, 18–32.

MYERSON, R. (1991). *Game theory: Analysis of conflict.* Harvard University Press.

PEREZ-RICHET, E. (2014). Interim Bayesian persuasion: First steps. *American Economic Review, Papers and Proceedings* 104, 469–474.

——— (2016). A proof of Blackwell's theorem. *Working Paper.*

PIERMONT, E. (2016). Distributional uncertainty and persuasion. *Working Paper.*

RAYO, L. & SEGAL, I. (2010). Optimal information disclosure. *Journal of Political Economy* 118, 949–987.

SCHOTTER, A. & TREVINO, I. (2014). Belief elicitation in the laboratory. *Annual Review of Economics* 6, 103–128.

SOBEL, J. (2013). Ten possible experiments on communication and deception. *Journal of Economic Behavior and Organization* 93, 408–413.

TANEVA, I. (2014). Information design. *Working paper.*

TSAKAS, E., TSAKAS, N. & XEFTERIS, D. (2017). Resisting persuasion. *Working paper.*

Tversky, A. & Kahneman, D. (1973). Availability: a heuristic for judging frequency andprobability. *Cognitive Psychology* 5, 207–232.

——— (1983). Extensional versus intuitive reasoning: the conjunction fallacy in probability judgment. *Psychological Review* 90, 293–315.

Wang, Y. (2013). Bayesian persuasion with multiple receivers. *Working paper*.