Resisting Persuasion

Elias Tsakas [*1], Nikolas Tsakas [†2], and Dimitrios Xefteris [‡2]

[1]Department of Economics, Maastricht University

[2]Department of Economics, University of Cyprus

December 29, 2020

Abstract

In the context of Bayesian Persuasion (Kamenica and Gentzkow, 2011), typically, a biased Sender designs a signal to influence the binary decision of an unbiased Receiver. Can the Receiver improve her payoffs by adopting a *resistance strategy*, i.e., by committing into incurring (deterministic or stochastic) costs if she picks the Sender-preferred action? We argue that deterministic resistance strategies cannot improve the Receiver's payoffs, whereas *stochastic resistance strategies* can increase both the informativeness of the signal and the Receiver's payoffs. We fully characterize the optimal resistance strategy and show that it always induces a substantial increase in the Receiver's welfare, as well as a perfectly informative signal.

Keywords: Bayesian persuasion; resistance; uncertainty; money burning.
JEL classification: D72, D82, D83, K40, M38

1. Introduction

Persuasion describes the process in which an agent intends to alter the behavior of another agent in his favor, and it is commonly observed in the context of economic and political decisions, such as for instance in product advertisement (Bertrand et al., 2010) and in elections (DellaVigna and Kaplan, 2007). For this reason, it has attracted interest in both economics (DellaVigna and Gentzkow, 2010;

[*]Department of Economics (AE1), Maastricht University, P.O. Box 616, 6200 MD, Maastricht, The Netherlands; E-mail: e.tsakas@maastrichtuniversity.nl

[†]Department of Economics, University of Cyprus, P.O. Box 20537, 1678, Nicosia, Cyprus; E-mail: tsakas.nikolaos@ucy.ac.cy; ORCID: 0000-0002-5397-5261 (Corresponding Author).

[‡]Department of Economics, University of Cyprus, P.O. Box 20537, 1678, Nicosia, Cyprus; E-mail: xefteris.dimitrios@ucy.ac.cy; ORCID:0000-0001-7397-5288

Glazer and Rubinstein, 2006; Kamenica and Gentzkow, 2011) and psychology (Petty and Cacioppo, 1986; Perloff, 2017).

Within microeconomic theory, starting with Kamenica and Gentzkow (2011), Bayesian persuasion models have recently surged. The main feature of this simple reduced-form model is that a (male) Sender publicly commits to a Bayesian signal, the outcome of which is observed by a (female) Receiver, who then takes an action that affects both of them.[1] Notably, in the (standard) analysis of most Bayesian persuasion games, the Receiver is essentially a dummy player who simply reacts to the information she receives from the Sender. As a result, the informativeness of the equilibrium signal is often relatively low, and thus so is the expected payoff of the Receiver, who, typically, prefers more informative signals. In this paper, we ask the following question: *is there a way for the Receiver to induce a more informative signal and increase her payoffs?*

As a potential answer, we propose a mechanism –henceforth, called a *resistance strategy*– that allows the Receiver to voluntarily commit ex ante to, *either deterministically or stochastically*, incur some cost (burn money) if she picks a Sender-preferred action. Intuitively, such a resistance strategy benefits the Receiver by making it harder for the Sender to persuade her to choose his preferred action, in the sense that he needs to provide more accurate information to her. Yet, at the same time, this requires the Receiver to commit to paying a cost.

We study this mechanism in a simple setup with binary state and action spaces, in which the Sender has state-independent preferences, whereas the Receiver wants to choose the action that is optimal in the given state of the world. Within this setup, we find that deterministic strategies cannot improve the Receiver's expected payoff, because the benefits from the increased informativeness of the designed signal are counterbalanced by the cost the Receiver needs to incur whenever she chooses the Sender's preferred action. However, we show that *there exists always a stochastic resistance strategy that makes the Receiver strictly better off*. In fact, in equilibrium, this strategy induces the Sender to design of the fully informative signal.

Consider the following example: a regulator (Receiver) is about to enter a series of meetings with a lobbyist of an industry (Sender) regarding whether the regulator maintains a low-tax regime for said industry or not. The lobbyist prefers the low-tax regime to be maintained, whereas the regulator prefers to make the right choice for the local economy, which depends on the state of the world economy. In an attempt to persuade the regulator to maintain the regime, the lobbyist announces the implementation of an economic experiment (signal) that will provide information on the state of the world economy.

Within this context, the regulator could use two types of resistance strategies: First, before the meetings with the lobbyist, she could publicly announce that she will abolish the low-tax regime. Therefore, if she eventually chooses to maintain it, she will appear to be inconsistent, which would necessarily incur a political cost. Thus, this would be a *deterministic resistance strategy*. Alterna-

[1] Commitment before the Sender observing his own type is what essentially distinguishes Bayesian persuasion from the earlier literature on cheap-talk (Crawford and Sobel, 1982).

tively, again prior to the meetings with the lobbyist, she could announce the hiring of a law firm to provide advice on whether the low-tax regime is legal or not. If the advice is positive (i.e. there are no legal issues with maintaining the low-tax regime), then either decision would not incur any cost, but ignoring a negative advice will lead to a cost. This would be an example of a *stochastic resistance strategy*.

Defining resistance in terms of burning money seems like a reasonable modelling choice for two reasons. First, it is always feasible for the Receiver to commit (contractually or otherwise) to incur costs, which is why such strategies are frequently used in the literature (Amador and Bagwell, 2020; Ambrus and Egorov, 2017; Austen–Smith and Banks, 2000; Kartik, 2007). Second, the types of applications that we have in mind typically involve some sort of public commitment, which leads to (reputation or other types of) costs if the commitment is broken. This is the case for instance in the literature on international negotiations (Leventoglu and Tarar, 2005; Tarar and Leventoglu, 2009) and auditing (Fearon, 1994, 1997; Tomz, 2007), but also in psychology in work on the susceptibility to normative influence (Batra et al., 2001), preference consistency (Wells and Iyengar, 2005), source credibility (Tormala and Petty, 2004) and attitude certainty (Tormala and Petty, 2002).

Now, consider in our earlier example the same common prior and utilities as in the standard example in Kamenica and Gentzkow (2011), viz., the prior probability that the low-tax regime is optimal given the state of the world is equal to 0.3, and each agent gets a unit of utility for preferred outcomes and zero units of utility otherwise. Then, in the benchmark case without resistance, the regulator would be persuaded to maintain the low-tax regime whenever her posterior belief about the low-tax regime being optimal would be at least 0.5, and an optimally designed experiment would achieve this with probability 0.6. Would this change if the regulator were to announce ex ante a plan to abolish the low-tax regime, i.e. to use a deterministic resistance strategy, reneging on which would incur a cost of $\kappa = 0.2$?[2] In this case, the minimum posterior for which she would be persuaded would be 0.6 and the optimal experiment could achieve this with probability only 0.5, as it would have to be more informative (in Blackwell's sense). Obviously, this would be detrimental for the lobbyist. But, would it benefit the regulator? It turns out that it would not, because the benefits from the increased informativeness of the experiment would be counterbalanced by the cost κ she would pay should she decides to maintain the low-tax regime.

This is not coincidental. Actually, within the current framework, any deterministic resistance strategy leaves the Receiver unaffected in equilibrium (Lemma 1). The only effect resistance could have in this case would be to lead to a more informative signal designed by the Sender, with the Sender being the one to internalize all costs. Hence, in case only deterministic resistance strategies are feasible, the Receiver will not have strict incentives to resist persuasion (Theorem 1(i)).

However, things are different if the regulator can employ stochastic resistance strategies. For instance, if she would not incur the cost of $\kappa = 0.2$ for sure, but only with probability 50% (in our

[2] That means that the regulator's expected payoff when choosing to maintain the low-tax regime would be $1 - \kappa$ if a low-tax regime is optimal given the state of the world, and $-\kappa$ otherwise.

working example, this would be the likelihood that the law firm would suggest that maintaining low taxes is illegal). In this case, if the lobbyist wants to persuade the regulator irrespective of the realized cost, he would have to design the same experiment as before, which would yield a posterior of 0.6 with probability 0.5. But, in this case the regulator would be strictly better off than with deterministic resistance, as the experiment is equally informative and her likelihood of having to incur a cost is halved. Alternatively, the lobbyist could design a less informative experiment that would persuade the regulator only when the realized cost is zero. Yet, interestingly, this is never optimal for him.

This is again not coincidental. In fact, there are stochastic resistance strategies that are even more beneficial for the regulator than the one described above. In our main theorem, we fully characterize the optimal stochastic resistance strategy, which we show that it always makes the Receiver strictly better off compared to the benchmark case (Theorem 1(ii)). This optimal strategy takes the following, arguably, elegant form: the Receiver assigns positive probability to incurring a zero cost, and spreads the remaining probability uniformly to all costs up to the maximum cost that would still allow her to pick the Sender-preferred action if she knew with certainty that it matched the state of world. Importantly, the Sender's optimal signal as a response to this resistance strategy is fully informative, and the Sender will be the one to internalize all costs.

Our work is part of the literature on Bayesian persuasion (Kamenica and Gentzkow, 2011), which has developed in several different directions. Alonso and Câmara (2017) study Bayesian persuasion with multiple Receivers and discuss commitment as a potential welfare enhancing strategy for them, which in that context takes a different form than here and, more importantly, is not incentive compatible for the Receivers. Kolotilin et al. (2017) study persuasion with a privately informed Receiver, which can reduce the persuading ability of the Sender. This is in contrast to our setup in which the two agents possess the same amount of information at any point of the process. Yet, similar to our results, under certain conditions, the Sender may design a signal that reveals the state of the world. Increased signal informativeness may also be a result of competition between Senders (Gentzkow and Kamenica, 2017a,b) or of the Receiver's ability to choose to wait longer before choosing (Bizzotto and Vigier, 2020). Furthermore, Alonso and Câmara (2016) have considered an extended model with heterogeneous priors, Hu and Weng (2020) study an environment where the Sender has limited knowledge about the Receiver's information source, Perez-Richet (2014) and Hedlund (2017) consider a similar game with a privately informed Sender, and Kolotilin (2015) one where not all information comes from the Sender's experiment. Finally, Hagmann and Loewenstein (2017) study biases in information processing that affect persuasion, based on whether the provided information supports the Receiver's prior beliefs. For a recent review of this literature, we refer to Kamenica (2019).

As already mentioned, our results relate to those of the wider money-burning literature and entail that persuasion approaches should control for potential resistance strategies that might be available to the subjects of the persuasion attempts. Indeed, it is known that in many instances destroying own utility is an effective means of convincing other players to behave according to one's interests (Ben-Porath and Dekel, 1992; van Damme, 1989). For instance, as far as communication frameworks are concerned, money burning is proved to expand the set of equilibrium outcomes (Austen–Smith and Banks, 2000; Kartik, 2007), bringing along possibilities for payoff enhancements.

Moreover, in the context of optimal delegation contract design (Amador and Bagwell, 2020) it has recently been shown that a principal can enhance her utility by delegating decisions to biased agents (Li, 2019) or by inducing action-contingent money burning to the agent (Ambrus and Egorov, 2017). That is, a contract designer may be better off by just assigning decision power to an agent who is biased against the Sender-preferred action or by punishing the agent for taking a Sender-preferred action, when she does not internalize the loss in utility experienced by the agent. Our work reinforces these intuitions by demonstrating that the Receiver can increase her payoffs by action-contingent burning of own utility.

Our work is also related more broadly to the literature on noisy strategic communication, either in the context of cheap talk (Blume et al., 2007) or in Bayesian persuasion (Tsakas and Tsakas, 2020). The common denominator of this earlier work and our paper is that the Receiver's action does not depend simply on the realized outcome of the experiment, but also on realizations of additional sources of uncertainty.

This is also the case in the literature on mediated cheap talk, where such additional uncertainty is due to the presence of the mediator (Ivanov, 2010; Goltsman et al., 2009).[3] In fact, there are clear parallels between the two approaches: First, in mediation the Sender's incentives to provide information arise from the uncertainty regarding the impact his message will have on the Receiver's action. The same is true here, since the Sender is uncertain about the extent to which the Receiver will be inclined against choosing the Sender's-preferred action. Nevertheless, in both cases, the behavior of the mediator and the Receiver, respectively, leads to higher expected payoffs for both agents, compared to the no-communication benchmark. Second, contrary to environments with delegation or arbitration, both mediation and resistance are characterized by lack of Receiver's commitment to particular actions.

Finally, our approach is linked to more traditional communication settings, related to price discrimination, in which a Receiver can make the Sender indifferent among all signals that are optimal for some cost realization (see for instance Bergemann et al., 2015; Roesler and Szentes, 2017). For another application to selling mechanisms see Ivanov (2020). This is rather reassuring, as it guarantees that the nature of the optimal resistance strategy in the context of Bayesian persuasion is well accepted in alternative frameworks of information transmission.

In what follows, we present the model (Section 2) and the formal results (Section 3) and then we conclude (Section 4).

[3]We thank an anonymous referee for pointing out these relationships to us.

2. The Model

The benchmark persuasion game (without resistance): Let $\Omega = \{G, B\}$ be a binary state space and $A = \{g, b\}$ be a binary action space. There are two agents, a male Sender and a female Receiver with utility functions $v : A \times \Omega \to \mathbb{R}$ and $u : A \times \Omega \to \mathbb{R}$, respectively. As usual, we assume that the Receiver wants to "match the true state", i.e., her utility is normalized as follows:

$$\begin{aligned} u_G = u(g, G) &> u(b, G) = 0, \\ 1 = u(b, B) &> u(g, B) = 0. \end{aligned}$$

This means that the Receiver prefers the action G if and only if she attaches probability at least

$$p_G := \frac{1}{1 + u_G}$$

to state G. The Sender is assumed to strictly prefer action g over action b, with his utilities being normalized as follows:

$$\begin{aligned} v_G = v(g, G) &> v(b, G) = 0, \\ 1 = v(g, B) &> v(b, B) = 0. \end{aligned}$$

It is not difficult to verify that the aforementioned normalizations are without loss of generality. Throughout the paper –to improve exposition– *we let the Sender have state-independent preferences*, i.e., $v_G = 1$.[4]

Both agents are Bayesian expected utility maximizers and share a common prior that assigns probability $p_0 \in (0, 1)$ to the state G. Before the Receiver chooses an action, the Sender chooses a signal/experiment $\pi : \Omega \to \Delta(S)$, which is represented by a pair of distributions $\pi(\cdot|G)$ and $\pi(\cdot|B)$ over a finite set of signal realizations, S. The choice of the signal is observed by both players, and so is the actual realization. Hence, information is symmetric throughout the game. Formally, given the signal π, upon observing a realization $s \in S$, both agents update their beliefs to a posterior identified by the probability

$$p_s := \frac{p_0 \pi(s|G)}{p_0 \pi(s|G) + (1 - p_0)\pi(s|B)}$$

attached to the state G. Then, using her updated belief, the Receiver chooses an action $a \in A$ that maximizes her interim utility,[5]

$$u_s(a) := p_s u(a, G) + (1 - p_s) u(a, B).$$

[4] In the Appendix we discuss the robustness of the result to the Sender having state-dependent preferences.

[5] To avoid confusion, we use the term *interim utility* to describe expected utilities conditional on the signal realization, and the term *ex-ante utility* to describe unconditional expected utilities.

Whenever the Receiver is indifferent between the two actions she chooses the action most preferred by the Sender, i.e., she picks g. We assume that $p_0 \in (0, p_G)$ to ensure that the problem is not trivial, viz., if $p_0 \geq p_G$, the Sender has no reason to attempt persuading the Receiver, who will anyway choose g given her prior.

Let us denote the Receiver's action for an arbitrary posterior $p \in [0,1]$ by $\hat{a}(p)$, i.e., $\hat{a}(p) = g$ if and only if $p \geq p_G$. Then, by state independence, the Sender's utility depends solely on p, i.e., it is the case that $v(p) = 1$ if $p \geq p_G$ and $v(p) = 0$ if $p < p_G$. Hence, the Sender's problem reduces to choosing a signal π that maximizes his ex-ante utility

$$\begin{aligned} V(\pi) &:= \sum_{s \in S} \left[p_0 \pi(s|G) + (1-p_0)\pi(s|B)\right] v(p_s) \\ &= p_0 \pi(\{s \in S : p_s \geq p_G\}|G) + (1-p_0)\pi(\{s \in S : p_s \geq p_G\}|B). \end{aligned}$$

Note that the Sender's optimal signal strategy exists always and is characterized by means of the standard concavification technique (Kamenica and Gentzkow, 2011).

Resistance strategies: Prior to the design of the signal, the Receiver sets up a *resistance strategy* against persuasion, which relies (deterministically or stochastically) on incurring costs $\kappa \geq 0$. For each realized cost κ, the Receiver's utility becomes

$$u^{\kappa}(a, \omega) := u(a, \omega) - \begin{cases} \kappa \text{ if } a = g \\ 0 \text{ if } a = b \end{cases}$$

while the Sender's utility remains unaffected, i.e., $v^{\kappa}(a, \omega) := v(a, \omega)$. Throughout the paper, we focus on costs $\kappa \leq u_G$, so that persuasion is possible and the problem non trivial (see Equation (1) below).

Formally, a resistance strategy r is a distribution over the interval of possible costs $\kappa \in [0, u_G]$ and can be fully identified by a cumulative distribution function $F_r : [0, u_G] \to [0,1]$. Resistance strategies that put probability one to a single $\kappa \in [0, u_G]$ are called *deterministic* and the set of these strategies is denoted by $\mathcal{D}$. The set $\mathcal{D}$ is a subset of the set $\mathcal{S}$ of *stochastic* resistance strategies, which includes all strategies that have *at least* one point in their support.

Persuasion game with resistance: The timing of our game is as follows (see Figure 1). First, the Receiver chooses a resistance strategy r from a choice set $\mathcal{R} \in \{\mathcal{D}, \mathcal{S}\}$, i.e., depending on the specific application that we have in mind, the Receiver may or may not have access to stochastic resistance strategies. Then, the Sender chooses a signal. They both observe the signal realization, $s \in S$, and update their beliefs. Similar to the benchmark case, both the signal and the realization are common knowledge. Subsequently, a cost $\kappa \in \text{supp}(r)$ is drawn and observed by both agents. Finally, the

Receiver chooses an action that maximizes her interim utility,

$$u_s^\kappa(a) := u_s(a) - \begin{cases} \kappa \text{ if } a = g \\ 0 \text{ if } a = b \end{cases}$$

The optimal action of the Receiver depends both on her posterior belief p_s and on the realized cost κ and is denoted by $\hat{a}(p_s, \kappa)$. Once again, indifference is resolved in favor of the Sender, i.e., $\hat{a}(p_s, \kappa) = g$ if and only if $u_s^\kappa(g) \geq u_s^\kappa(b)$, or equivalently whenever

$$p_s \geq p_G^\kappa := \frac{1+\kappa}{1+u_G}. \tag{1}$$

Obviously, p_G^κ is strictly increasing in κ, i.e., the threshold probability for choosing g increases in the realized cost.

Thus, the Sender's utility depends on the posterior $p \in [0,1]$ and the realized cost κ, i.e., $v(p,\kappa) = 1$ if $p \geq p_G^\kappa$ and $v(p,\kappa) = 0$ if $p < p_G^\kappa$. Hence, the Sender's ex-ante utility from choosing a signal π when the Receiver has chosen a resistance strategy r becomes

$$\begin{aligned} V_r(\pi) &:= \sum_{s\in S} \big(p_0\pi(s|G) + (1-p_0)\pi(s|B)\big) \int_0^{u_G} v(p_s,\kappa)dr(\kappa) \\ &= \int_0^{u_G} \Big[p_0\pi(\{s \in S : p_s \geq p_G^\kappa\}|G) + (1-p_0)\pi(\{s \in S : p_s \geq p_G^\kappa\}|B)\Big] dr(\kappa). \end{aligned}$$

The optimal signal for the Sender (given a resistance strategy r) is denoted by $\hat{\pi}_r$. Finally, the ex-ante utility of the Receiver from choosing a resistance strategy $r \in \mathcal{R}$ becomes:

$$U(r) = p_0 \int_0^{u_G} \sum_{s\in S} \hat{\pi}_r(s|G) u_s^\kappa(\hat{a}(\mu_s,\kappa), G) dr(\kappa) + (1-p_0) \int_0^{u_G} \sum_{s\in S} \hat{\pi}_r(s|B) u_s^\kappa(\hat{a}(\mu_s,\kappa), B) dr(\kappa)$$

Graphically, the timing of the persuasion game with a set of resistance strategies $\mathcal{R}$ is presented in Figure 1. Events above the line are observed by both players, whereas events below the line are not observed by anyone. The order of steps that correspond to $t = 3$ and $t = 4$ can be reversed, provided

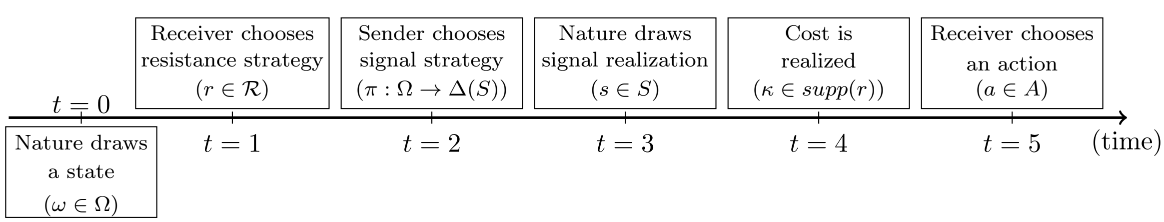


Figure 1: Persuasion game with resistance.

that the realization of the cost takes place after the choice of the signal by the Sender and before the choice of the action by the Receiver. Moreover, step 4 is trivial for deterministic strategies, in which case it is omitted. Finally, observe that *all players have access to exactly the same information at any instance of the game*. Yet, it is crucial that the Receiver can observe the realized cost before choosing an action, whereas the Sender can observe only the resistance strategy (i.e. the distribution from which the cost will be drawn) when choosing the signal.

3. Results

Let us start by pointing out that the Sender can design an optimal signal that puts positive probability to two signal realizations.[6] Hence, we can restrict the set of signal realizations to some $S = \{s_G, s_B\}$. Thus, a signal π is represented by a pair of probabilities $q := \pi(s_G|G)$ and $z := \pi(s_G|B)$ and will sometimes be mentioned as *signal* (q, z). For a signal (q, z), the Receiver may form two posteriors regarding the probability that the state is G, one for each signal realization,

$$p_{s_G} = \frac{p_0 q}{p_0 q + (1 - p_0) z} \qquad \text{or} \qquad p_{s_B} = \frac{p_0(1 - q)}{p_0(1 - q) + (1 - p_0)(1 - z)}. \tag{2}$$

Remark 1. The assumption $p_0 < \frac{1}{1+u_G}$ guarantees that the Sender cannot persuade the Receiver in both signal realizations irrespective of the cost. Indeed, $p_{s_G} \geq p_G^\kappa$ (resp., $p_{s_B} \geq p_G^\kappa$) implies $p_{s_B} < p_G^\kappa$ (resp., $p_{s_G} < p_G^\kappa$). Hence, the Sender focuses on persuading the Receiver to take action g in one realization, say realization s_G. ◁

Thus, the Sender chooses a signal that maximizes his ex-ante utility, subject to the constraint $p_{s_G} \geq p_G^\kappa$. For such a signal, the Receiver chooses action g upon observing signal realization s_G and action b upon observing s_B.

The next lemma characterizes the optimal signal and the ex-ante utilities of the two agents for an arbitrary deterministic resistance strategy, $r \in \mathcal{D}$.

Lemma 1. *Let the Receiver choose a deterministic resistance strategy, $r \in \mathcal{D}$, that assigns probability one to some $\kappa \in [0, u_G]$. Then, the optimal signal for the Sender is*

$$\hat{\pi}_r(s_G|G) = 1 \qquad \text{and} \qquad \hat{\pi}_r(s_G|B) = \frac{p_0(1 - p_G^\kappa)}{(1 - p_0) p_G^\kappa}.$$

Moreover, if the Receiver responds optimally to $\hat{\pi}_r$, the ex-ante utilities of the Receiver and the Sender

[6] The argument is similar to the one of Kamenica and Gentzkow (2011). In particular, the only complications of restricting to only two realizations appear when a stochastic resistance strategy has been used by the Receiver. For instance, consider a stochastic strategy in which the Receiver puts positive probability only to cost zero and to a positive cost. In this case, the Receiver is never persuaded for low posterior beliefs, she is always persuaded for high posterior beliefs, and is sometimes persuaded for intermediate posterior beliefs (depending on the cost that will be drawn). So, from the Sender's point of view, we can consider the Receiver's response at these intermediate beliefs as a third artificial action of the Receiver, and finally apply the standard concavification technique in a game with two states and three actions. Thus, using the usual Caratheodory-based argument, two signal realizations suffice for the Sender to achieve his maximum expected utility.

are

$$U(r) = 1 - p_0 \qquad \textit{and} \qquad V_r(\hat{\pi}_r) = \frac{p_0}{p_G^\kappa}$$

respectively.

All proofs can be found in the Appendix.

Some immediate observations can be made from the previous result. First, the Sender still designs the signal in a way that makes the Receiver indifferent between choosing each of the actions when the signal realization is s_G, as she did in the case without resistance. Obviously, the informativeness of the signal should be higher in order to compensate for the cost induced to the Receiver when choosing g. Geometrically, this is because the Receiver's cutoff probability has now shifted to the right –viz., formally, p_G^κ increases in κ (see Figure 2b). Nevertheless, the ex-ante utility of the Receiver is independent of the cost κ, since the indirect compensation she receives from the increased informativeness is exactly equal to the direct utility loss induced by the cost associated with her resistance strategy. Actually, this result holds irrespective of whether the Sender has state-independent preferences or not.

The inability of the Receiver to increase her ex-ante utility by committing to any deterministic resistance strategy puts under question the effectiveness of commitment to burn money as a successful means to resist persuasion. However, this is rushing to a false conclusion.

In fact, resistance can be beneficial if it makes the Sender uncertain of the exact cost that the Receiver will eventually bear (in case she chooses g). This can happen through stochastic resistance strategies. The important feature of stochastic resistance strategies is that the Receiver's cost is realized after the Sender has designed the signal, but before the Receiver chooses her action. As we will see, this is enough to increase the ex-ante utility of the Receiver, compared to the no-resistance case, despite the absence of any explicit benefit.

We proceed directly to the statement of the main result of our analysis.

Theorem 1 (Optimal Resistance Strategy)**.** *The Receiver benefits strictly from uncertainty, i.e.,* $\max_{r\in\mathcal{S}} U(r) > \max_{r\in\mathcal{D}} U(r)$. *Moreover, the following hold:*

(i) DETERMINISTIC RESISTANCE: *The Receiver has no incentive to resist persuasion. Formally,* $\arg\max_{r\in\mathcal{D}} U(r) = \mathcal{D}$.

(ii) STOCHASTIC RESISTANCE: *There is a unique optimal resistance strategy, which leads to the fully informative signal. Formally,* $\arg\max_{r\in\mathcal{S}} U(r) = \{\hat{r}\}$, *where* $\hat{r}$ *is identified by the cumulative distribution function*

$$\hat{F}_{\hat{r}}(\kappa) = p_G^\kappa$$

for each $\kappa \in [0, u_G]$. *Then, the Sender will respond by choosing*

$$\hat{\pi}_{\hat{r}}(s_G|G) = \hat{\pi}_{\hat{r}}(s_B|B) = 1.$$

Theorem 1(i) follows immediately from Lemma 1. All resistance strategies provide the same ex-ante utility to the Receiver, while at the same time higher κ reduces the Sender's ex-ante utility, thus the Receiver –who is assumed to choose the Sender's most preferred strategy when indifferent– has no incentive to resist persuasion.

Theorem 1(ii) shows that indeed the Receiver can improve her ex-ante utility by using stochastic resistance. In fact, the optimal stochastic resistance strategy makes the Sender design a perfectly informative signal. The improvement in the Receiver's ex-ante utility is achieved due to the introduction of uncertainty to the Sender about the size of the Receiver's cost. The optimal distribution is continuous. It places a positive mass only at $\kappa = 0$ and the remaining probability is distributed uniformly among the rest of the admissible costs. Under optimal resistance, the ex-ante utilities of the Receiver and the Sender are:

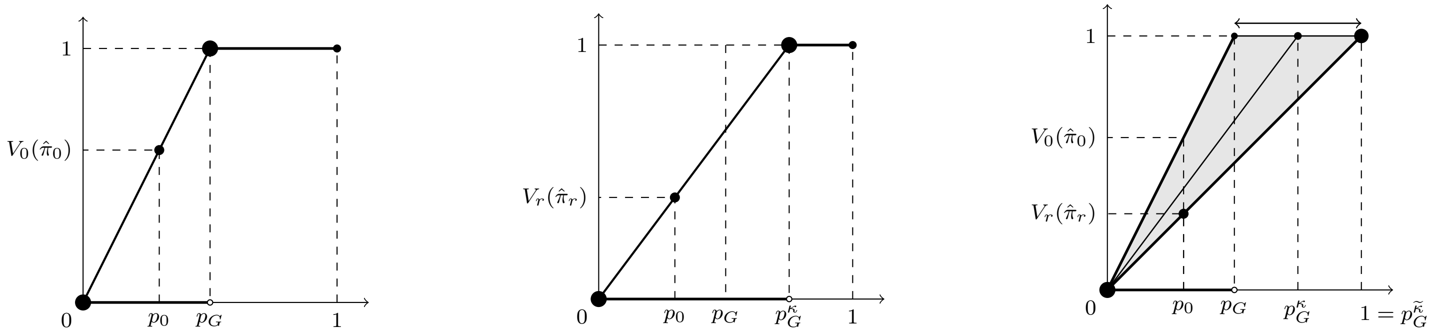
$$U(\hat{r}) = (1 - p_0) + p_0 u_G \left[1 + \frac{1}{1 + u_G}\right] \qquad \text{and} \qquad V_{\hat{r}}(\hat{\pi}_{\hat{r}}) = p_0$$

respectively.

The intuition behind the result is the following: When choosing a resistance strategy, the Receiver faces a trade-off. On the one hand, she wants to design a resistance strategy (CDF) that induces as small costs as possible, because this would lead her to higher payoffs when observing s_G (for which in equilibrium she gets persuaded to choose g). On the other hand, commiting to a sufficiently costly resistance strategy induces the design of a more informative signal by the Sender, which makes her more likely to choose the correct action, which in turn increases her payoff. Regarding the Sender's choice, each signal leads to a pair of possible posteriors for the Receiver, where the posterior formed upon observing s_G implicitly determines a cutoff cost below which the Receiver gets persuaded by s_G, while above it she does not. Thus, the Sender's problem essentially boils down to optimally selecting this cutoff cost. For a cutoff cost to be optimal, the standard concavification argument induces the chosen signal to never generate a posterior payoff/cutoff cost pair above a certain line (see Figure 2c). Thus, among all resistance strategies that lead to the same choice of cutoff cost by the Sender, the CDF that minimizes the cost for the Receiver is the one that generates pairs exactly along this line. This is exactly the CDF of the uniform distribution with an atom at zero cost. This leads the Sender to design the fully informative signal that persuades the Receiver when observing s_G irrespective of the realized cost, i.e. the optimal cutoff cost is u_G.

The proof of the result is constructive and clearly highlights the intuition described above. Namely, for each cost realization, the Receiver's choice is characterized by a threshold posterior, above which she chooses action g upon observing s_G. The signal chosen by the Sender determines the maximum cost –or equivalently the minimum posterior– for which the Receiver may be persuaded.

Essentially, the Sender has to decide how accurate the signal will be in state B (i.e., what z he will pick), while in state G he always prefers the signal to be fully accurate (i.e. he always picks $q = 1$). This generates a dilemma to the Sender, as a more accurate signal (lower z) increases the potential



(a) No resistance. (b) Deterministic resistance. (c) Stochastic resistance.

Figure 2: Persuasion with and without resistance: The three subfigures show the ex-ante utility of the Sender as a function of the posterior beliefs of the Receiver. Subfigure (a) shows the benchmark case without resistance. Subfigure (b) shows an example of deterministic resistance associated with cost κ, which shows the shift of the cutoff probability p_G^κ to the right. Subfigure (c) shows an example of stochastic resistance, where p_G^κ denotes the posterior for which the Receiver gets persuaded for costs below κ. In equilibrium, this coincides with the diagonal, which shows that for most of the realized costs the Receiver would be persuaded even with a less informative signal than the equilibrium one, which is why stochastic resistance increases the Receiver's ex-ante utility. In all subfigures, the thick dots indicate the points of the equilibrium posteriors, while $V_0(\hat{\pi}_0)$ and $V_r(\hat{\pi}_r)$ denote the ex-ante utility of the Sender in equilibrium without and with resistance respectively.

costs for which persuasion is possible, but at the same time it reduces the expected utility gained by a persuaded Receiver –as it reduces the expected frequency with which the Receiver chooses g in state B. Naturally, the Sender prefers to increase the informativeness of the signal whenever this leads to a sufficiently large increase of the likelihood that the Receiver will be persuaded. In other words, if a less informative signal leads to a threshold cost κ_1 –above (below) which persuasion fails (succeeds)– whereas a more informative signal would lead to a threshold cost $\kappa_2 > \kappa_1$, then the Sender prefers to provide the more informative signal if the cost distribution puts sufficiently large probability to costs in the interval $(\kappa_1, \kappa_2]$, for which the latter signal can persuade the Receiver, whereas the former could not.

The Receiver is able to anticipate this behavior and chooses her resistance strategy accordingly. Let us begin by describing the optimal distribution for the Receiver among all distributions that lead to the choice of a signal that induces the same maximum cost for which persuasion is possible, call it $\widetilde{\kappa}$. First, its support will be $[0, \widetilde{\kappa}]$, which means that it will allow persuasion to be possible for any cost realization. This is true because any signal that induces $\widetilde{\kappa}$ as a threshold satisfies –by construction– that a Receiver who observes a realized cost $\widetilde{\kappa}$ is exactly indifferent between the two available actions when obseving s_G. Thus, any distribution that puts some probability on costs above $\widetilde{\kappa}$ yields the same ex-ante utility to the Receiver as an otherwise identical distribution that assigns this probability on $\widetilde{\kappa}$. Yet, the latter is preferred for the Sender, as it increases the probability with which the Receiver gets persuaded by s_G.

Second, within $[0, \widetilde{\kappa}]$ the Receiver puts a positive mass only at $\kappa = 0$ and distributes the remaining probability uniformly among the rest of admissible costs. The uniform distribution balances a trade-off faced by the Receiver, who would like to put higher probability to lower costs, under the constraint that the Sender's optimal signal induces a threshold cost $\widetilde{\kappa}$. A non-uniform distribution would either no longer support $\widetilde{\kappa}$ as an optimal threshold for the Sender (if low costs were sufficiently likely), or it would lead to a higher expected cost for the Receiver without changing the choice of the Sender (if high costs were sufficiently likely). Essentially, the uniform distribution here makes the Sender exactly indifferent between all potentially optimal signals that induce any threshold cost between 0 and $\widetilde{\kappa}$. Furthermore, this indifference needs to be satisfied also for cost equal to zero. The latter happens whenever the distribution has a properly weighed mass exactly at zero, because the Sender's expected utility (as a function of $\widetilde{\kappa}$) is given by $V_r(\widetilde{\kappa}) = p_0 F_r(\widetilde{\kappa})\frac{1+u_G}{1+\widetilde{\kappa}}$.

The above analysis reduces the problem to one in which the Receiver needs to choose the maximum cost $\widetilde{\kappa}$ at which she wishes to be potentially persuaded. This decision entails again a trade-off for the Receiver. On the one hand, a lower $\widetilde{\kappa}$ would, obviously, lead to a lower expected cost to be paid when choosing g. Yet, on the other hand, a higher $\widetilde{\kappa}$ would lead to a more informative signal, thus also to fewer mistakes, i.e. fewer times in which the Receiver would choose g in state B. The relative intensity of the two terms is significantly affected by the positive mass that the optimal distribution puts at zero, which guarantees that the increase in the expected cost is going to be sufficiently small so as to always yield a smaller decrease in the Receiver's ex-ante utility than the increase induced by the higher accuracy of her choices. Hence, it is always optimal for the Receiver to choose the distribution that induces $\widetilde{\kappa} = u_G$, which in turn leads the Sender to design the fully informative signal.

4. Concluding remarks

We have made the case that, in a standard Bayesian persuasion context, committing to stochastic action-contingent money burning can improve the informativeness of the signal and increase the payoffs of the Receiver at the same time. While deterministic money burning also improves signal quality, without affecting the Receiver's payoffs –i.e. it is plausible that a Receiver employs such a strategy– it is (ex-ante) Pareto inefficient and, hence, with an arbitrarily small transfer the Sender can convince the Receiver not to engage in it. Small transfers are less effective when the Receiver has stochastic resistance strategies at her disposal, and, therefore, a substantial increase in signal quality is more probable in such instances.[7] This observation is particularly important when social welfare critically depends on the quality of the signal and less so on the payoffs of the two players (e.g. when the decision of the Receiver does not affect only her payoffs and those of the Sender, but also the welfare of the whole society, as with the regulator and the local economy in our initial example),

[7] In many contexts of empirical interest large monetary transfers from the Sender to the Receiver are illegal. For instance, while a lobbyist can contribute to the re-election campaign of an elected politician/regulator, there is a legal limit to such contributions.

and establishes that stochastic action-contingent money burning is resilient –to some extent– to the Sender's available countermeasures (e.g. bribes, threats, etc.).

To derive our result we have made the simplifying assumption that the Sender's payoffs do not depend on the state of the world. This helped us provide a more transparent formal analysis, and a concise and tight paper, but it is not key for our main results. The effect of resistance is still prevalent when the Sender has state-dependent preferences. The main results carry on to this case: Deterministic strategies still fail to increase the Receiver's ex-ante utility, whereas stochastic resistance do not. The only difference is that with stochastic resistance strategies it is no longer guaranteed –but it is still often the case– that the Receiver's optimal strategy would induce the design of a fully informative signal by the Sender.[8]

Our second restrictive assumption is that we only consider two states. Allowing for an arbitrary finite state space would not change the economic intuition of most our results, especially when the Sender has state-independent preferences over the Receiver's two actions. This is because the Receiver would still choose on the basis of a probability threshold: the only difference would be that this threshold would be defined in terms of a hyperplane cutting through the space of beliefs (i.e., the simplex over the state space). This would only make the analysis more cumbersome, but we confidently conjecture that the key insights would still go through.

Undeniably, the potential of stochastic money burning as an effective resistance strategy needs to be further explored in additional frameworks. Environments with limited commitment, imperfect/bounded rationality, external and/or several sources of information seem particularly interesting. While the analysis of all these cases is clearly beyond the scope of the current paper, we believe that the current results can properly inform and, at least partially, facilitate the subsequent analyses of such relevant settings.

A. Appendix

Proof of Lemma 1: The condition $p_{s_G} \geq p_G^{\kappa}$ is equivalent to $z \leq \frac{p_0(1-p_G^{\kappa})}{(1-p_0)p_G^{\kappa}} q$. Moreover, the ex-ante utility of Sender from selecting a signal (q, z) is as follows:

$$V_r(q, z) = qp_0 + z(1 - p_0)$$

$V_r(q, z)$ increases in both q and z, as long as the abovementioned condition holds, which implies that the optimal signal should satisfy $\hat{q}_r = 1$ and $\hat{z}_r = \frac{p_0(1-p_G^{\kappa})}{(1-p_0)p_G^{\kappa}}$, for which the Receiver chooses action g when observing s_G and action b otherwise. Substituting this into V_r, we directly obtain:

$$V_r(\hat{q}_r, \hat{z}_r) = p_0 \left(1 + \frac{1 - p_G^{\kappa}}{p_G^{\kappa}}\right) = \frac{p_0}{p_G^{\kappa}}$$

[8] The optimal resistance strategy can be obtained by a direct extension of the proof of Theorem 1(ii), and the relevant formal arguments are presented in the Appendix.

Analogously, the Receiver's ex-ante utility, anticipating that the Sender will choose optimally, is as follows:

$$\begin{aligned} U(r) &= p_0\hat{q}_r(u_G-\kappa)+(1-p_0)\left[-\hat{z}_r\kappa+(1-\hat{z}_r)\right]= \\ &= p_0(u_G-\kappa)+(1-p_0)-\frac{p_0(1-p_G^\kappa)}{p_G^\kappa}(1+\kappa)= \\ &= p_0(u_G-\kappa)+(1-p_0)-p_0\left(\frac{1+u_G}{1+\kappa}-1\right)(1+\kappa)= \\ &= (1-p_0) \end{aligned}$$

□

Proof of Theorem 1(ii): Consider a strategy $r\in\mathcal{S}$ identified by a CDF F_r, where the cost κ can take values in $[0,u_G]$. By Equation (1), for any $\kappa\in[0,u_G]$ that is drawn the Receiver chooses action g if her posterior satisfies $p\geq p_G^\kappa=\frac{1+\kappa}{1+u_G}$. According to this and given that the prior is low (see Remark 1), any signal (q,z) that the Sender chooses (as in Equation (2)) can persuade the Receiver in at most one of the two signal realizations, say s_G. For each κ, the condition $p\geq p_G^\kappa$ is satisfied by some signal (q,z) if this signal satisfies the condition $z\leq\frac{p_0}{1-p_0}\frac{u_G-\kappa}{1+\kappa}q=\widetilde{z}(\kappa)$. Note that, $\widetilde{z}(\kappa)$ is strictly decreasing in κ, therefore if s_G persuades the Receiver for some cost κ then it does so for all $\kappa'<\kappa$ as well. Given the bounds on the values of κ, if $z>\widetilde{z}(0)$ the Receiver is never persuaded (even if the realized cost is zero), whereas if $z\leq\widetilde{z}(u_G)$ the Receiver is persuaded by s_G for any cost realization. Thus, the Sender's ex-ante utility from a signal (q,z), given resistance strategy r associated with CDF F_r is equal to:

$$V_r(q,z)=\begin{cases} qp_0+z(1-p_0) & \text{if } z\leq\widetilde{z}(u_G) \\ F_r(k)[qp_0+z(1-p_0)] & \text{if } z=\widetilde{z}(k) \text{ for } k\in[0,u_G) \\ 0 & \text{if } z>\widetilde{z}(0) \end{cases}$$

Observe that, similarly to the deterministic case, irrespective of the value of z, it is always optimal to choose $\hat{q}_r=1$. Moreover, the fact that $qp_0+z(1-p_0)>0$ means that persuasion is always beneficial for the Sender, therefore it is never optimal to choose $z>\widetilde{z}(0)$. Furthermore, it is never optimal either to choose $z<\widetilde{z}(u_G)$, as it always yields lower ex-ante utility compared to $z=\widetilde{z}(u_G)$.

Therefore, potential optimal signals are those that satisfy $\widetilde{q}=1$ and $\widetilde{z}(\kappa)=\frac{p_0}{1-p_0}\frac{u_G-\kappa}{1+\kappa}$ for some $\kappa\in[0,u_G]$, which corresponds to the maximum cost for which the Receiver is persuaded by signal realization s_G. Hence, the problem of the Sender is equivalent to choosing a threshold value $\widetilde{\kappa}$ above which persuasion does not take place. The threshold value that maximizes his ex-ante utility is denoted by $\widetilde{\kappa}^*$. For some $\widetilde{\kappa}\in[0,u_G]$ and given the distribution F_r associated to the strategy r of the

Receiver, the ex-ante utility of the Sender can be rewritten as a function of $\widetilde{\kappa}$ as follows:

$$V_r(\widetilde{\kappa}) = F_r(\widetilde{\kappa})\left[p_0 + (1-p_0)\widetilde{z}(\widetilde{\kappa})\right] = p_0 F_r(\widetilde{\kappa})\left(1 + \frac{u_G - \widetilde{\kappa}}{1+\widetilde{\kappa}}\right) = p_0 F_r(\widetilde{\kappa})\frac{1+u_G}{1+\widetilde{\kappa}} \tag{A.1}$$

Note that the function $V_r(\widetilde{\kappa})$ is upper semicontinuous –because F_r is a CDF– and it is defined over a closed interval. Therefore, it achieves a global maximum. Yet, it might have multiple maxima –possibly, infinitely many– depending on the shape of F_r. Among these, we assume that the Sender chooses the most preferred for the Receiver. This is actually the (unique) greatest element of the set of these maxima, which the following observation guarantees that it exists.

Observation: Let $h : [a, b] \to \mathbb{R}$ be an upper-semicontinuous function. Hence, it achieves a maximum. Then, let $X := \arg\max h \subseteq [a, b]$ be the set of global maxima of h and, also, let $\bar{x} := \sup X$, i.e. $\bar{x}$ is the supremum of the set X of global maxima of h. Now, assume that $\max X$ does not exist. Then, X is infinite and there exists some infinite sequence $\{x_t\}_{t=1}^{\infty}$ with $x_t \in X$ for all $t \in \{1, \dots, \infty\}$ such that $x_t \to \bar{x}$. Yet, by upper semicontinuity of h, we have that $h(\bar{x}) \geq \limsup h(x_t) = \max h$, where, obviously, the latter can only hold with equality. That is, $h(\bar{x}) = \max h$, and therefore $\bar{x} = \max X$, which contradicts the assumption above.

Hence, any choice of distribution F_r by the Receiver induces some $\widetilde{\kappa}$ to be chosen by the Sender. Thus, let $\mathcal{F}_{\widetilde{\kappa}}$ be the set of all available distributions that induce $\widetilde{\kappa}$.[9] The ex-ante utility of the Receiver when choosing a strategy $r \in \mathcal{S}$ associated to a distribution $F_r \in \mathcal{F}_{\widetilde{\kappa}}$ is as follows:

$$\begin{aligned} U(r) &= [1 - F_r(\widetilde{\kappa})](1-p_0) + F_r(\widetilde{\kappa})\left\{p_0 u_G + (1-p_0)[1-\widetilde{z}(\widetilde{\kappa})]\right\} - [p_0 + (1-p_0)\widetilde{z}(\widetilde{\kappa})]\int_{[0,\widetilde{\kappa}]} \kappa dF_r(\kappa) = \\ &= p_0 u_G F_r(\widetilde{\kappa}) + (1-p_0)\left[1 - F_r(\widetilde{\kappa})\widetilde{z}(\widetilde{\kappa})\right] - [p_0 + (1-p_0)\widetilde{z}(\widetilde{\kappa})]\int_{[0,\widetilde{\kappa}]} \kappa dF_r(\kappa) \end{aligned} \tag{A.2}$$

where the (Lebesque) integral essentially refers to the expected cost for the Receiver in the region where persuasion is possible.

Our next step is to find the optimal distribution within each set $\mathcal{F}_{\widetilde{\kappa}}$, recalling that also the Receiver, when indifferent, chooses the most preferred resistance strategy for the Sender. Recall also that a signal chosen by the Sender that induces $\widetilde{\kappa}$, will be structured such that the Receiver will be indifferent between "always choosing action b" and "choosing action g when observing s_G". Therefore, the Receiver is indifferent between two distributions that distribute probability identically up to $\widetilde{\kappa}$ and one of them assigns positive mass to values $\kappa \in (\widetilde{\kappa}, u_G]$ while the other one puts the same mass exactly on $\widetilde{\kappa}$. Yet, the latter distribution is preferred by the Sender, because it increases the probability with which the Receiver will get persuaded, without affecting his optimal choice.[10]

[9]This set is never empty because it always contains the trivial distribution in which the Receiver puts probability one to the cost $\widetilde{\kappa}$.

[10]On the one hand, the Sender would not choose a signal that would induce $\widetilde{\kappa}' > \widetilde{\kappa}$, because that would require the

Hence, all potentially optimal distributions that induce $\widetilde{\kappa}$ satisfy $F(\widetilde{\kappa}) = 1$. Moreover, by definition, each of these distributions should satisfy the following condition for all $\kappa \in [0, \widetilde{\kappa})$:

$$V_r(\widetilde{\kappa}) \geq V_r(\kappa) \;\Leftrightarrow\; \frac{1}{1+\widetilde{\kappa}} \geq F_r(\kappa)\frac{1}{1+\kappa} \tag{A.3}$$

This is because, we have considered the distributions for which it is optimal for the Sender to induce $\widetilde{\kappa}$ as the maximum cost for which the Receiver can be persuaded. In fact, the equivalence relation guarantees that this inequality characterizes the set of all distributions in $\mathcal{F}_{\widetilde{\kappa}}$.

Among all the distributions satisfying Expression (A.3), the Receiver prefers the one with the minimum expected value (as this enters negatively in her expected utility, in Expression (A.2)). It is straightforward to see that there is a unique such distribution, which is the one that satisfies Expression (A.3) with equality for all $\kappa \in [0, \widetilde{\kappa}]$ and is denoted by $\widetilde{F}_{\widetilde{\kappa}}$, i.e.

$$\widetilde{F}_{\widetilde{\kappa}}(\kappa) = \begin{cases} \frac{1+\kappa}{1+\widetilde{\kappa}} & \text{if } \kappa \in [0, \widetilde{\kappa}) \\ 1 & \text{if } \kappa \in [\widetilde{\kappa}, 1] \end{cases} \tag{A.4}$$

This distribution would make the Sender indifferent between signals that induce any $\kappa \in [0, \widetilde{\kappa}]$. Yet, given that when indifferent he chooses the most preferred to the Receiver, his choice will be $\widetilde{\kappa}$.

It is important to notice that this distribution is differentiable in $[0, u_G]$, it puts positive mass $\frac{1}{1+\widetilde{\kappa}}$ only at $\kappa = 0$ and it is uniform in $(0, \widetilde{\kappa})$. Thus, it also has an associated well–defined continuous probability distribution function, which is $\widetilde{f}_{\widetilde{\kappa}}(\kappa) = \frac{1}{1+\widetilde{\kappa}}$ for every $\kappa \in (0, \widetilde{\kappa})$.

Therefore, we have shown that the Receiver can induce any $\widetilde{\kappa} \in [0, u_G]$ and we have found the optimal distribution for achieving so. Hence, the problem is summarized in finding the value of $\widetilde{\kappa}$ that would maximize the ex-ante utility of the Receiver, if it is induced. Given our previous findings, we can rewrite the ex-ante utility of the Receiver as a function of $\widetilde{\kappa}$ as follows:

$$\begin{aligned} U(\widetilde{\kappa}) &= p_0 u_G + (1-p_0)[1 - \widetilde{z}(\widetilde{\kappa})] - [p_0 + (1-p_0)\widetilde{z}(\widetilde{\kappa})] \int_0^{\widetilde{\kappa}} \kappa \widetilde{f}_{\widetilde{\kappa}}(\kappa) d\kappa \\ &= p_0 u_G + (1-p_0) - p_0 \frac{u_G - \widetilde{\kappa}}{1+\widetilde{\kappa}} - p_0 \frac{1+u_G}{1+\widetilde{\kappa}} \int_0^{\widetilde{\kappa}} \kappa \widetilde{f}_{\widetilde{\kappa}}(\kappa) d\kappa \\ &= p_0 u_G + (1-p_0) - p_0 \left[\frac{u_G - \widetilde{\kappa}}{1+\widetilde{\kappa}} + \frac{1+u_G}{2(1+\widetilde{\kappa})^2}\widetilde{\kappa}^2 \right] \end{aligned} \tag{A.5}$$

By differentiating with respect to $\widetilde{\kappa}$ we get that $U'(\widetilde{\kappa}) = p_0 \frac{1+u_G}{(1+\widetilde{\kappa})^3} > 0$. Thus, the ex-ante utility is strictly increasing in $[0, u_G]$. Hence, the Receiver wants to induce $\widetilde{\kappa} = u_G$, which means that she chooses distribution $\widetilde{F}_{u_G}$. Therefore, the optimal strategy $\hat{r}$ among all $r \in \mathcal{S}$ is characterized by the

signal to be more informative, without increasing the probability of persuasion (as the Receiver is always persuaded by s_G). On the other hand, the Sender would not choose a signal that would induce some $\widetilde{\kappa}' < \widetilde{\kappa}$, because if $\widetilde{\kappa}'$ is optimal now, then it should have also been optimal for the initial distribution, which cannot happen since $\widetilde{\kappa}$ was by definition the induced value that maximizes the ex-ante utility of the Sender for the chosen distribution. Therefore, the new distribution does not alter the subsequent signal choice of the Sender.

cumulative distribution function

$$\hat{F}_{\hat{r}}(\kappa) = \frac{1+\kappa}{1+u_G} = p_G^{\kappa}, \quad \text{for } \kappa \in [0, u_G]$$

For this resistance strategy the Sender designs the fully informative signal, i.e. $\hat{q}_{\hat{r}} = 1$ and $\hat{z}_{\hat{r}} = 0$. $\square$

Sender with state-dependent preferences: The previous results carry on to more general setups in which the incentives of the two players are still misaligned –the Receiver wants to match the state, whereas the Sender prefers g over b irrespective of the state– but the Sender has state-dependent preferences, i.e. $v_G \neq 1$. Namely, deterministic strategies are still unable to increase the Receiver's ex-ante utility –with the proof being essentially identical to that of Lemma 1–, whereas stochastic resistance strategies can do so. The only difference is that with stochastic resistance strategies it is no longer guaranteed –but still often possible– that the Receiver's optimal strategy would induce the design of the fully informative signal by the Sender.

The optimal resistance strategy can be obtained by a direct extension of the proof of Theorem 1(ii). Namely, with similar arguments as before, we can get that the optimal distribution for the Sender among all distributions that induce $\widetilde{\kappa}$ is as follows:

$$\widetilde{F}_{\widetilde{\kappa}}(\kappa) = \begin{cases} \left(v_G + \frac{u_G - \widetilde{\kappa}}{1+\widetilde{\kappa}}\right) \cdot \frac{1}{v_G + \frac{u_G - \kappa}{1+\kappa}} & \text{if } \kappa \in [0, \widetilde{\kappa}) \\ 1 & \text{if } \kappa \in [\widetilde{\kappa}, 1] \end{cases}$$

and for each choice of $\widetilde{\kappa}$ the ex-ante utility of the Receiver is as follows:

$$U(\widetilde{\kappa}) = p_0 u_G + (1-p_0) - p_0 \frac{u_G - \widetilde{\kappa}}{1+\widetilde{\kappa}} - p_0 \frac{1+u_G}{1+\widetilde{\kappa}} \int_0^{\widetilde{\kappa}} \kappa \widetilde{f}_{\widetilde{\kappa}}(\kappa) d\kappa$$

Given this, it can be shown that it is still always optimal for the Receiver to resist persuasion, i.e. to induce $\widetilde{\kappa} > 0$, although inducing the fully informative signal might not be optimal for the Receiver, i.e. the Receiver might prefer to induce $\widetilde{\kappa} < u_G$. For this to be shown, let use observe that $U(\widetilde{\kappa})$ is a differentiable function of $\widetilde{\kappa} \in [0, u_G]$, thus it admits a unique global maximum in this interval. Moreover, given that when indifferent the Receiver chooses the most preferred choice for the Sender –who prefers a $\widetilde{\kappa}$ as small as possible– and that by differentiability of $U(\widetilde{\kappa})$ we can find the minimum of $\text{argmax}_{\widetilde{\kappa} \in [0, u_G]} U(\widetilde{\kappa})$, there exists a unique optimal $\widetilde{\kappa}^*$ for the Receiver. Moreover, if we calculate the derivative of the ex-ante utility and evaluate its limit at zero, we get that: $\lim_{\widetilde{\kappa} \to 0} U'(\widetilde{\kappa}) = p_0(1+u_G) > 0$. This implies that $\widetilde{\kappa}^* > 0$, which means that resisting persuasion always yields higher payoffs to the Receiver than no persuasion.

References

Alonso, R., Câmara, O.: Bayesian persuasion with heterogeneous priors. J. Econ. Theory 165, 672–706 (2016)

Alonso, R., Câmara, O.: Persuading voters. Am. Econ. Rev. 106, 3590–3605 (2017)

Amador, M., Bagwell, K.: Money Burning in the Theory of Delegation. Gam. Econ. Behav. 121, 382–412 (2020).

Ambrus, A., Egorov, G.: Delegation and nonmonetary incentives. J. Econ. Theory 171, 101–135 (2017)

Austen–Smith, D., Banks, J.S.: Cheap Talk and Burned Money. J. Econ. Theory 91(1), 1–16 (2000)

Batra, R., Homer, P.M., Kahle, L.R.: Values, Susceptibility to Normative Influence, and Attribute Importance Weights: A Nomological Analysis. J. Consum. Psychol. 11(2), 115–128 (2001)

Ben-Porath, E., Dekel, E.: Signaling future actions and potential for sacrifice. J. Econ. Theory 57, 36–51 (1992)

Bergemann, D., Brooks, B., Morris, S.: The Limits of Price Discrimination. Am. Econ. Rev. 105 (3), 921–57 (2015)

Bertrand, M., Karlan, D.S., Mullainathan, S., Shafir, E., Zinman, J.: What's advertising content worth? Evidence fromn a consumer credit marketing field experiment. Q. J. Econ. 125, 263–305 (2010)

Bizzotto, J., Vigier, A.: Can a better informed listener be easier to persuade?. Econ. Theory (2020). https://doi.org/10.1007/s00199-020-01321-w

Blume, A., Board, O.J., Kawamura, K.: Noisy talk. Theor. Econ. 2, 395–440 (2007)

Brehm, J.W.: A theory of psychological reactance. New York: Academic Press (1966)

Cialdini, R.B., Cacioppo, J.T., Bassett, R., Miller, J.A.: Low-ball procedure for producing compliance: Commitment then cost. J. Pers. Soc. Psychol., 36(5), 463–476 (1978).

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

DellaVigna, S., Gentzkow, M.: Persuasion: Empirical Evidence. Annu. Rev. Econ. 2, 643–669 (2010)

DellaVigna, S., Kaplan, E.: The Fox News effect: media bias and voting. Q. J. Econ. 122(3), 1187–1234 (2007)

Fearon, J.: Domestic Political Audiences and the Escalation of International Disputes. Am. Polit. Sci. Rev. 88(3), 577–592 (1994)

Fearon, J.: Signaling Foreign Policy Interests: Tying Hands versus Sinking Costs. J. Conflict Resol. 41(1), 68–90 (1997)

Fransen, M.L., Smit, E.G., Verlegh, P.W.J.: Strategies and motives for resistance to persuasion: an integrative framework. Front. Psychol. 6:1201 (2015)

Gentzkow, M., Kamenica, E.: Bayesian persuasion with multiple senders and rich signal spaces. Gam. Econ. Behav. 111, 411–429 (2017a)

Gentzkow, M., Kamenica, E.: Competition in persuasion. Rev. Econ. Stud. 84, 300-322 (2017b).

Glazer, J., Rubinstein, A: A study in the pragmatics of persuasion: a game theoretical approach. Theor. Econ. 1, 395–410 (2006)

Goltsman, M., Hörner, J., Pavlov, G., Squintani, F.: Mediation, arbitration and negotiation. J. Econ. Theory 144, 1397–1420 (2009)

Gopinath, M., Nyer, N.U.: The effect of public commitment on resistance to persuasion: The influence of attitude certainty, issue importance, susceptibility to normative influence, preference for consistency and source proximity. Int. J. Res. Mark. 26, 60–68 (2009).

Hagmann, D., Loewenstein, G.: Persuasion with Motivated Beliefs. working paper (2017)

Hedlund, J.: Bayesian persuasion by a privately informed sender. J. Econ. Theory 167, 229–268 (2017)

Hu, J., Weng, X.: Robust persuasion of a privately informed receiver. Econ. Theory (2020). https://doi.org/10.1007/s00199-020-01299-5

Ivanov, M.: Communication via a strategic mediator. J. Econ. Theory 145, 869–884 (2010)

Ivanov, M.: Optimal monotone signals in Bayesian persuasion mechanisms. Econ. Theory (2020). https://doi.org/10.1007/s00199-020-01277-x

Jacks, J.Z., Cameron, K.A.: Strategies for resisting persuasion. Basic Appl. Soc. Psych. 25(2), 145–161 (2003)

Jellison, M.J., Mills, J.: Effect of public commitment upon opinions. J. Exp. Soc. Psych. 5, 340–346 (1969)

Kamenica, E.: Bayesian persuasion and information design. Annu. Rev. Econ. 11, 249–272 (2019)

Kamenica, E., Gentzkow, M.: Bayesian persuasion. Am. Econ. Rev. 101, 2590–2615 (2011)

Kartik, N.: A note on cheap talk and burned money. J. Econ. Theory 136(1), 749–758 (2007)

Knowles, E.S., Linn, J.A.: Resistance and persuasion. Psychology Press. New Jersey (2004)

Kolotilin, A.: Experimental design to persuade. Gam. Econ. Behav. 90, 215-226 (2015)

Kolotilin, A., Mylovanov, T., Zapechelnyuk, A., Li, M.: Persuasion of a privately informed receiver. Econometrica 85, 1949-1964 (2017)

Lehmann, E., Simula, L., Trannoy, A.: Tax me if you can! Optimal nonlinear income tax between competing governments. Q. J. Econ. 129(4), 1995-2030 (2014)

Leventoglu, B., Tarar, A.: Prenegotiation Public Commitment in Domestic and International Bargaining. Am. Polit. Sci. Rev. 99(3), 419–433 (2005)

Li, C.: Bayesian Persuasion and Authority in Organizations. SSRN Working Paper, nr. 3309376 (2019)

Lokhorst, A.M., Van Dijk, E., Staats, H.: Public commitment making as a structural solution in social dilemmas. J. Environ. Psychol. 29, 400–406 (2009)

Myerson, R. B.: Incentives to cultivate favored minorities under alternative electoral systems. Am. Polit. Sci. Rev. 87(4), 856-869 (1993)

Perez-Richet, E.: Interim Bayesian persuasion: First steps. Am. Econ. Rev. 104, 469–474 (2014)

Perloff, R.M.: The dynamics of persuasion: communication and attitudes in the 21st Century, Sixth Edition. Routledge, New York (2017)

Petty, R.E., Cacioppo, J.T.: Communication and persuasion: central and peripheral routes to attitude change. Springer, New York (1986)

Ringold, D.J.: Boomerang Effects in Response to Public Health Interventions: Some Unintended Consequences in the Alcoholic Beverage Market. J. Consum. Policy 25(1), 27–63 (2002)

Roesler, A. K., Szentes, B.: Buyer-optimal learning and monopoly pricing. Am. Econ. Rev. 107(7), 2072-80 (2017)

Stigler, G.J.: The Economics of information. J. Polit. Econ. 69, 213–225 (1961)

Tarar, A., Leventoglu, B.: Public Commitment in crisis bargaining. Int. Stud. Quart. 53, 817–839 (2009)

Tomz, M.: Domestic Audience Costs in International Relations: An Experimental Approach. Int. Organ. 61, 821–840 (2007)

Tormala, Z.L., Petty, R.E.: What Doesn't Kill Me Makes Me Stronger: The Effects of Resisting Persuasion on Attitude Certainty. J. Pers. Soc. Psychol. 83(6), 1298–1313 (2002)

Tormala, Z.L., Petty, R.E.: Source Credibility and Attitude Certainty: A Metacognitive Analysis of Resistance to Persuasion. J. Consum. Psychol. 14(4), 427–442 (2004)

Tsakas, E., Tsakas: Noisy persuasion. Working paper (2020)

van Damme, E.: Stable equilibria and forward induction. J. Econ. Theory 48, 476–496 (1989)

Wells, R.E., Iyengar, S.S.: Positive illusions of preference consistency: When remaining eluded by one's preferences yields greater subjective well-being and decision outcomes. Organ. Behav. Hum. Dec. 98, 66–87 (2005)