

Ricardo Alonso and Odilon Câmara

Persuading skeptics and reaffirming believers

Working paper

Original citation:

Alonso, Ricardo and Câmara, Odilon (2014) *Persuading skeptics and reaffirming believers*. USC Marshall School of Business Research Paper Series, [University of Southern California Marshall School of Business](#), Los Angeles, CA, USA.

Originally available from: USC Marshall School of Business, [author's personal page](#).

This version available at: <http://eprints.lse.ac.uk/58680/>

Available in LSE Research Online: August 2014

© 2014 The Author, [USC Marshall School of Business](#)

LSE has developed LSE Research Online so that users may access research output of the School. Copyright © and Moral Rights for the papers on this site are retained by the individual authors and/or other copyright owners. Users may download and/or print one copy of any article(s) in LSE Research Online to facilitate their private study or for non-commercial research. You may not engage in further distribution of the material or use it for any profit-making activities or any commercial gain. You may freely distribute the URL (<http://eprints.lse.ac.uk>) of the LSE Research Online website.

Persuading Skeptics and Reaffirming Believers*

RICARDO ALONSO[†]

ODILON CÂMARA[†]

Marshall School of Business

University of Southern California

May 28, 2014

Abstract

In a world where rational individuals may hold different prior beliefs, a sender can influence the behavior of a receiver by controlling the informativeness of a signal. We characterize the set of distributions of posterior beliefs that can be induced by a signal, and provide necessary and sufficient conditions for a sender to benefit from information control. We examine a class of models with no value of information control under common priors, and show that a sender generically benefits from information control under heterogeneous priors. We extend our analysis to cases where the receiver's prior is unknown to the sender.

JEL classification: D72, D83, M31.

Keywords: Persuasion, information control, heterogeneous priors.

*We thank Dan Bernhardt, Emir Kamenica, and Anton Kolotilin for detailed comments on earlier drafts of this paper. We also thank Isabelle Brocas, Juan Carrillo, Maxim Ivanov, Navin Kartik, Jin Li, Tony Marino, Niko Matouschek, John Matsusaka, Tymofiy Mylovanov, Michael Powell, Luis Rayo, Joel Sobel, Eric Van den Steen, Tim Van Zandt and Yanhui Wu for their suggestions, as well as the following audiences: 2013 SWET, Carey Business School, Kellogg School of Management, London School of Economics, McMaster University, Queen's University, University of Bonn, University of British Columbia, University of Southern California, and University of Western Ontario.

[†]USC FBE Dept, 3670 Trousdale Parkway Ste. 308, BRI-308 MC-0804, Los Angeles, CA 90089-0804. vralonso@marshall.usc.edu and ocamara@marshall.usc.edu.

1 Introduction

A notable feature of organizations is that those with decision making power are lobbied. In many cases, individuals influence decision makers by changing the information available to them. For instance, individuals can acquire and communicate hard evidence, or signal soft information. Another way of influencing decision makers' learning is by directly specifying the informativeness of the signals that they observe, that is, by engaging in *information control* (as in e.g. Brocas and Carrillo 2007 and Kamenica and Gentzkow 2011).

Information control is pervasive in economics and politics. A pharmaceutical company chooses which initial animal tests to perform, and the results influence the Food and Drug Administration's decision to approve further human testing. A central bank shapes the informativeness of a market index observed by households (such as inflation) by determining which information is collected and how to compute the index. A news channel selects the questions asked by the host of an electoral debate, and the answers affect voters' opinions about the candidates. In all these cases, changing the signal (e.g., changing the test, the rules to generate the index, or the questions asked) changes what decision makers can learn.

One rationale for an individual to engage in information control is the presence of conflicting interests, as designing what decision makers learn can sway the latter's choices to decisions favored by the former. Another important rationale, on which we focus in this paper, arises when individuals and decision makers disagree in their views of the world.¹ We ask: how does open disagreement affect an individual's benefit from persuading others, and her choice of an optimal signal?

The next example, where a novel political issue must be addressed by a policy maker, illustrates our main insights. As Callander (2011) points out, a large part of the difficulty in policy making is that the policy maker may be uncertain about which policies produce which outcomes, and much political disagreement is over beliefs about this mapping. This was certainly true in the late 19th century, when a fast succession of technological breakthroughs

¹Many papers study the role of heterogeneous priors in economics and politics. Giat et al. (2010) use data on pharmaceutical projects to study R&D under heterogeneous priors; Patton and Timmermann (2010) find empirical evidence that heterogeneity in prior beliefs is an important factor explaining the cross-sectional dispersion in forecasts of GDP growth and inflation; Gentzkow and Shapiro (2006) study the effects of prior beliefs on media bias.

created the electric power industry. Politicians had to decide how to regulate safety in this nascent industry, at a time when there was an increasing number of fatal electrocutions and significant disagreement over the dangers of electricity, even among members of the scientific community (for instance, much of the safety concerns were over voltage, instead of the more important amperage).

For concreteness, consider a policy maker (mayor) who must choose which policy $a \in [0, 1]$ to implement, where a lower a represents a liberal rule for the transmission of electricity, and a higher a represents strict restrictions, such as establishing a maximum voltage.² Let the uncertainty regarding the optimal regulation be captured by an unknown state of the world $\theta \in \{0, 0.5, 1\}$, so that the mayor’s payoff is $u_R(a, \theta) = -(a - \theta)^2$. A politically biased media outlet has a payoff increasing in the regulation level, $u_S(a, \theta) = a$. The media outlet (sender) has no private information, but can influence the mayor’s (receiver) decision through an investigative report.³ After observing the report’s finding, the mayor updates his expectation over the state and chooses policy $a^* = E_R[\theta]$. Therefore, the media chooses a signal that maximizes its ex ante expectation of the mayor’s ex post expectation of θ .

If the media and the mayor share a common prior belief, then the media doesn’t benefit from information control as the policy a is linear in the expected θ . This would not be the case if there is belief disagreement. Suppose that the priors over $\theta \in \{0, 0.5, 1\}$ are $p^R = (0.4, 0.5, 0.1)$ for the mayor and $p^S = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ for the media, so that $E_R[\theta] = 0.35 < E_S[\theta] = 0.5$. That is, from the media’s perspective the mayor is “skeptical” about the need for regulation. Clearly, information control is valuable in this case: by designing a “perfect” signal that fully reveals the state the media can increase the mayor’s expectation (and consequently his policy choice) from 0.35 to, on average, its own expectation 0.5. Nevertheless, a fully revealing signal is not optimal. The media’s optimal signal only determines whether $\theta = 1$ or not. If

²Two standards, alternating current (AC) and direct current (DC), were competing to dominate the market, in what became known as “the war of the currents.” An important comparative advantage of AC was its capacity to be transmitted over greater distances using higher voltage. Hence, stricter transmission rules were often championed by DC supporters — for example, the Edison Electric Light Company tried to influence the New York Board of Electrical Control to impose strict voltage limits in the city.

³The media can generate a report (signal) that is correlated with θ . The media can change the informativeness of its signal by changing, for example, its editorial board and the reporters assigned to cover the story. In Duggan and Martinelli (2011), a biased media outlet chooses the “slant” of its report.

the report reveals $\theta = 1$, then players share a common posterior. However, if the report shows that $\theta \neq 1$, then the mayor’s expectation becomes $\frac{0.4 \times 0 + 0.5 \times 0.5}{0.4 + 0.5} = \frac{5}{18}$, *strictly* higher than the media’s expectation $\frac{(1/3) \times 0 + (1/3) \times 0.5}{1/3 + 1/3} = 0.25$. With this signal the media converts the “skeptical” mayor into a “believer”, and expects the average policy to increase to $\frac{2}{3} \times \frac{5}{18} + \frac{1}{3} \times 1 = \frac{14}{27}$.

While it seems natural that the media benefits from providing information to a skeptic, it is less clear whether the same is true when the mayor is a believer. Suppose now that the mayor’s prior over states is $p^R = (0.1, 0.5, 0.4)$, while the media has the same prior as before, so that $E_S[\theta] = 0.5 < E_R[\theta] = 0.65$. Clearly, a signal that fully reveals the state does not benefit the media, as it expects the mayor’s expectation of θ to decrease on average. Perhaps surprisingly, the media can still benefit from designing the signal. The optimal signal only determines whether $\theta = 0.5$ or not. The mayor’s expectation decreases to 0.5 when the report reveals $\theta = 0.5$, and increases to $\frac{0.1 \times 0 + 0.4 \times 1}{0.1 + 0.4} = 0.8$ when the report shows that $\theta \neq 0.5$. With this signal the media expects the average policy to increase to $\frac{2}{3} \times 0.8 + \frac{1}{3} \times 0.5 = 0.7$. This is possible because, in spite of the mayor being a believer, the media assigns more probability (2/3) than the mayor (1/2) to the “beneficial” signal $\{\theta \neq 0.5\}$.

The previous example highlights two important points. First, while the common prior assumption may be appropriate for established policy issues with a long historical record of policy experimentation, technological breakthroughs and rapid social changes may create novel policy issues, with a potentially substantial initial belief disagreement. Second, open disagreement provides a separate rationale for information control — in the example, there is no value of information control when players share a common prior. In fact, Section 4 shows that in a more general class of models: (i) prior belief disagreement *generically*⁴ leads the sender to benefit from information control, and (ii) full information disclosure is often suboptimal, independently of whether the receiver is a skeptic or a believer.

Motivated by this example, we consider a general model in which a sender can influence a receiver’s behavior by designing his informational environment. After observing the realization of a signal, the receiver applies Bayes’ rule to update his belief, and chooses an action accordingly. The sender has no private information and can influence this action by designing what the receiver can learn from the signal, i.e. by specifying the statistical relation of the

⁴Genericity is interpreted over the space of pairs of prior beliefs.

signal to the underlying state. We make three assumptions regarding how Bayesian players process information. First, it is common knowledge that players hold different prior beliefs about the state, i.e. they “agree to disagree”. Second, this disagreement is non-dogmatic: each player initially assigns a positive probability to each possible state of the world. Third, the signal chosen by the sender is “commonly understood,” in the sense that if players knew the actual realization of the state, then they would agree on the likelihood of observing each possible signal realization.

We start our analysis by asking: from the sender’s perspective, what is the set of distributions of posterior beliefs that can be induced by a signal? When players share a common prior, Kamenica and Gentzkow (2011) (KG henceforth) establish that this set is defined by two properties: (i) posteriors must be homogeneous and (ii) the expected posterior must equal the prior. Now consider heterogeneous priors. Clearly, posteriors do not need to be homogeneous and, from the point of view of the sender, the receiver’s expected posterior does not need to equal either prior (as in the previous example). Our first contribution is to show that, given priors p^S and p^R , posteriors q^S and q^R form a bijection — q^R is derived from q^S through a perspective transformation. Moreover, this transformation is *independent* of the actual signal. Consequently, given prior beliefs, the probability distribution of posterior beliefs of *only one* player suffices to derive the *joint* probability distribution of posteriors generated by an arbitrary signal. This result allows us to characterize the set of distributions of posteriors that can be induced by a signal (Proposition 1). Importantly, our results imply that belief disagreement does not expand this set, that is, it does not allow the sender to generate “more” distributions of posterior beliefs.

We solve for the sender’s optimal signal (Proposition 2) and provide a simple geometric condition that is both necessary and sufficient for a sender to benefit from designing the signal (Corollary 1). We also obtain a necessary and sufficient condition for a sender to benefit from garbling a fully informative signal (Corollary 2).

In Section 4 we study pure-persuasion, i.e., models where the sender’s utility is not a function of the state. KG show that when players share a common prior, information control is valuable when the sender can exploit the non-concavity of the receiver’s action in his beliefs, or the convexity of the sender’s utility function in the receiver’s actions. We show that even in the absence of these features, the sender can still benefit from information

control by exploiting differences in players' prior beliefs. In fact, if the receiver's action is the expectation of a random variable and the state space has three or more distinct states, then a sender *generically* benefits from information control (Proposition 5), regardless of the curvature of the sender's utility. While the sender cannot induce "more" distributions over posterior beliefs, she can nevertheless benefit from a signal for which she puts more probability than the receiver on signal realizations that increase the receiver's expectation (and thus his action). Such signals exist for a generic pair of players' prior beliefs.

Our paper is primarily related to two strands in the literature.

Information Control: Some recent papers study the gains to players from controlling the information that reaches decision makers. In Brocas and Carrillo (2007), a leader without private information sways the decision of a follower in her favor by deciding the timing at which a decision must be made. As information arrives sequentially, choosing the timing of the decision is equivalent to shaping (in a particular way) the information available to the follower. Duggan and Martinelli (2011) consider one media outlet that can affect electoral outcomes by choosing the "slant" of its news reports. Gill and SgROI (2008, 2012) consider a privately-informed principal who can subject herself to a test designed to provide public information about her type, and can optimally choose the test's difficulty. Rayo and Segal (2010) study optimal advertising when a company can design how to reveal the attributes of its product, but it cannot distort this information. In a somewhat different setting, Ivanov (2010) studies the benefit to a principal of limiting the information available to a privately informed agent when they both engage in strategic communication (i.e. cheap talk). The paper most closely related to ours is KG. They analyze the problem of a sender who wants to persuade a receiver to change his action for an arbitrary state space and action space, and arbitrary, *but common*, prior beliefs, and arbitrary state-dependent preferences for both the sender and the receiver. We contribute to this literature by introducing and analyzing a new motive for information control: belief disagreement over an unknown state of the world.

Heterogeneous Priors and Persuasion: Several papers in economics, finance and politics have explored the implications of heterogeneous priors on equilibrium behavior and the performance of different economic institutions. In particular, Van den Steen (2004, 2009, 2010a, 2011) and Che and Kartik (2009) show that heterogeneous priors increase the incentives of agents to acquire information, as each agent believes that new evidence will back

their “point of view” and thus “persuade” others. Our work complements this view by showing that persuasion may be valuable even when others hold “beneficial” beliefs from the sender’s perspective. We also differ from this work in that we consider situations in which the sender has more leeway in shaping the signals that reach decision makers.

We present the model’s general setup in Section 2. Section 3 characterizes the value of information control. In Section 4 we examine pure persuasion models. Section 5 extends the model to the case of private priors. Section 6 concludes. All proofs are in the Appendices.

2 The Model

Our model features a game between a sender (she) and a receiver (he). The sender has no authority over the receiver’s actions, yet she can influence them through the design of a signal observed by the receiver. This setup can be regarded as a model of influence, a model of persuasion, or a model of managed learning where a sender “sways” a receiver into changing his action by carefully designing what he can learn. Our main departure from the previous literature on information control, particularly Brocas and Carrillo (2007) and Kamenica and Gentzkow (2011), is that we allow players to openly disagree about the uncertainty they face.

Preferences and Prior Beliefs: All players are expected utility maximizers. The receiver selects an action a from a compact set A . While in some applications it may be natural for the sender to also affect the outcome of the game directly by choosing an action, we abstract from this possibility in this paper. The sender and the receiver have preferences over actions $a \in A$ characterized by continuous von Neumann-Morgenstern utility functions $u_S(a, \theta)$ and $u_R(a, \theta)$, with $\theta \in \Theta$ and Θ a finite state space, common to both players.

Both players are initially uncertain about the realization of the state θ . A key aspect of our model is that players openly disagree about the likelihood of θ . Following Aumann (1976), this implies that rational players must then hold different prior beliefs.⁵ Thus let the receiver’s prior be $p^R = \{p_\theta^R\}_{\theta \in \Theta}$ and the sender’s prior be $p^S = \{p_\theta^S\}_{\theta \in \Theta}$. We assume that p^R and p^S belong to the interior of the simplex $\Delta(\Theta)$, that is, players have prior beliefs that

⁵See Morris (1994, 1995) and Van den Steen (2010b, 2011) for an analysis of the sources of heterogeneous priors and extended discussions of its role in economic theory.

are “totally mixed” as they have full support.⁶ This assumption will avoid known issues of non-convergence of posterior beliefs when belief distributions fail to be absolutely continuous with respect to each other (see Blackwell and Dubins 1962, and Kalai and Lehrer 1994).

In our base model these prior beliefs are common knowledge, i.e. players “agree to disagree” on their views of θ . This implies that differences in beliefs stem from differences in prior beliefs rather than differences in information. We extend the base model in Section 5 to consider cases where players have heterogenous prior beliefs drawn from some distribution $H(p^R, p^S)$. In that case, it will not be commonly known by players that they disagree on the likelihood of θ .

It is natural to inquire as to the sources of heterogenous prior beliefs and ponder whether these same sources may affect the way in which players process new information. For instance, mistakes in information processing will eventually lead players to different posterior beliefs, but will also call into question Bayesian updating. We take the view that players are Bayes rational, but may initially openly disagree on the likelihood of the state. Typically, this disagreement can come from lack of experimental evidence or historical records that would allow players to otherwise reach a consensus on their prior views. This was the case in our example in the Introduction where a poor understanding of electrical laws lead to widely varying views on the dangers of electricity. In fact, as argued in Van den Steen (2011), the Bayesian model specifies how new information is to be processed but is largely silent on how priors should be (or are actually) formed. Lacking a rational basis for selecting a prior, the assumption that, nevertheless, individuals should all agree on one may seem unfounded. In any case, open disagreement does not necessarily hinder players’ ability to process new information if heterogenous priors stem from insufficient data.

Signals and Information Control: All players process information according to Bayes rule. The receiver observes the realization of a signal π , updates his belief, and chooses an action. The sender can affect the receiver’s actions through the design of π . To be specific, a signal π consists of a finite realization space Z and a family of likelihood functions over Z , $\{\pi(\cdot|\theta)\}_{\theta \in \Theta}$, with $\pi(\cdot|\theta) \in \Delta(Z)$. Note that whether or not the signal realization is observed

⁶Actually, our results only require that players’ prior beliefs have a common support, which may be a strict subset of Θ . Assuming a full support eases the exposition without any loss of generality.

by the sender does not affect the receiver’s actions.

Key to our analysis is that π is a “commonly understood signal”: the sender’s choice of π is observed by the receiver and all players agree on the likelihood functions $\pi(\cdot|\theta)$, $\theta \in \Theta$.⁷ Common agreement over π generates substantial congruence in our model: all players agree on how a signal realization is generated given the state.⁸ To wit, if all players knew the actual realization of the state, then they would all agree on the likelihood of observing each $z \in Z$ for any signal π .

Our setup is closely related to models that study the incentives of agents to affect others’ learning, e.g. through “signal jamming” as in Holmström’s model of career concerns (Holmström 1999) or through obfuscation as in Ellison and Ellison (2009). In contrast to this literature, the sender in our model directly shapes the learning of the receiver by designing an “experiment” whose result is correlated with the underlying state. This interpretation of our model corresponds to several practical situations. For instance, rating systems and product certification fit this framework where consumers observe the result of an aggregate measure of the underlying quality of firms/products. Quality tests provide another example, as a firm may not know the quality of each single product, but can control the likelihood that a test detects a defective product. Finally, one can influence the information generated by a survey or a focus group, by specifying the questionnaire and the sampling methodology.

We make two important assumptions regarding the set of signals available to the sender. First, the sender can choose *any signal* that is correlated with the state. Thus our setup provides an upper bound on the sender’s benefit from information control in a setting with a more restricted space of signals. In particular, if the sender faces additional constraints, she will not engage in designing a signal if there is no value of information control in our unrestricted setup. Second, signals are costless to the sender. This is not a serious limi-

⁷Our assumption of a commonly understood signal is similar to the notion of “concordant beliefs” in Morris (1994). As Morris (1994) indicates, “beliefs are concordant if they agree about everything except the prior probability of payoff-relevant states”. Technically, his definition requires both agreement over the conditional distribution of signals given the state and that each player assigns positive probability to each signal realization. Our assumptions of a commonly understood signal and totally mixed priors imply that players’ beliefs are concordant in our setup.

⁸See Van den Steen (2011) and Acemoglu et al. (2006) for models where players also disagree on the informativeness of signals.

tation if all signals impose the same cost, and would not affect the choice of signal if the sender decides to influence the receiver. However, the optimal signal may change if different signals impose different costs. Gentzkow and Kamenica (2013) offer an initial exploration of persuasion with costly signals, where the cost of a signal is given by the expected Shannon entropy of the beliefs that it induces.

Our focus is on understanding when and how the sender benefits from designing the signal observed by the receiver. Given a signal π , for a signal realization z that induces the profile of posterior beliefs $(q^S(z), q^R(z))$, the receiver's choice in any Perfect Bayesian equilibrium must satisfy

$$a(q^R(z)) \in \arg \max_{a \in A} \sum_{\theta \in \Theta} q_{\theta}^R(z) u_R(a, \theta),$$

while the corresponding (subjective) expected utility of the sender after z is realized is

$$\sum_{\theta \in \Theta} q_{\theta}^S(z) u_S(a(q^R(z)), \theta).$$

We restrict attention to equilibria in which the receiver's choice only depends on the posterior belief induced by the observed signal realization. To this end we define a *language-invariant Perfect Bayesian equilibrium* as a Perfect Bayesian equilibrium where for every signals π and π' , and signal realizations z and z' for which $q^R(z) = q^R(z')$, the receiver selects the same action (or the same probability distribution over actions). Our focus on language-invariant equilibria allows us to abstract from the particular signal realization. Given an equilibrium $a(\cdot)$, we define the sender's expected payoff v when players hold beliefs (q^S, q^R) as

$$v(q^S, q^R) \equiv \sum_{\theta \in \Theta} q_{\theta}^S u_S(a(q^R), \theta), \text{ with } a(q^R) \in \arg \max_{a \in A} \sum_{\theta \in \Theta} q_{\theta}^R u_R(a, \theta). \quad (1)$$

We concentrate on equilibria for which the function v is upper-semicontinuous. This class of equilibria is non-empty: an equilibrium in which whenever the receiver is indifferent between actions he selects an action that maximizes the sender's expected utility, as a function of posterior beliefs only, is a (sender-preferred) language-invariant equilibrium for which v is upper semicontinuous.⁹ Given a language-invariant equilibrium that induces v , the sender's

⁹As noted in KG, this follows from Berge's maximum theorem. Upper-semicontinuity will prove convenient when establishing the existence of an optimal signal.

equilibrium expected utility is simply

$$\max_{\pi} \mathbb{E}_S^{\pi} [v(q^S(z), q^R(z))],$$

where the maximum is computed over all possible signals π .

Our primary interest in this paper are situations in which if the sender does not influence the receiver, then the receiver learns nothing about the state. In this case the sender's expected utility is simply $v(p^S, p^R)$. We thus define the *value of information control* as the maximum expected gain that can be attained by the sender in a Perfect Bayesian equilibrium, when in the absence of the sender's influence the receiver would remain uninformed. Note that the sender's maximum expected utility is attained in any sender-preferred equilibrium. Therefore, defining $V(p^S, p^R)$ as the expected utility of the sender in a sender-preferred equilibrium, the value of information control is

$$V(p^S, p^R) - v(p^S, p^R). \tag{2}$$

Trivially, a sender does not benefit from information control if and only if

$$V(p^S, p^R) = v(p^S, p^R).$$

Our framework also allows the study of the gains from obfuscating, or otherwise impeding, the receiver's learning. To accommodate this case, we simply posit that if the sender does not engage in information control, then the receiver observes a perfect signal of the state. Information control then takes the form of garbling — the sender can add noise to the receiver's signal in an arbitrary way. This effectively means that the sender can specify the statistical relation of every signal realization to the underlying state. We define the *value of garbling* as the maximum expected gain that can be attained by a sender in a Perfect Bayesian equilibrium when, absent her influence, the receiver learns the state.

Timing: The sender selects a signal $\pi (= (Z, \{\pi(\cdot|\theta)\}_{\theta \in \Theta}))$ after which the receiver observes a signal realization $z \in Z$, updates his beliefs according to Bayes' rule, selects an action, payoffs are realized and the game ends. As argued before, we concentrate on language-invariant perfect equilibria for which v is upper semicontinuous.

We have been silent regarding the true distribution governing the realization of θ . As our analysis is primarily positive and only considers the behavior of a sender when influencing a receiver, we remain agnostic as to the true distribution of the state.

Notational Conventions: For vectors $v, w \in \mathbb{R}^N$, we denote by $\langle v, w \rangle$ the standard inner product in \mathbb{R}^N , i.e. $\langle v, w \rangle = \sum_{i=1}^N v_i w_i$. As ours is a setup with heterogenous priors, this notation proves convenient when computing expectations where we need to specify both the information set and the individual whose perspective we are adopting. We also use a component-wise product of vectors, and denote it by vw , to refer to the vector whose components are the products of the components of each vector, i.e. $(vw)_i = v_i w_i$. Also, let $\cos(v, w)$ be the cosine of the angle between v and w , i.e. $\cos(v, w) = \frac{\langle v, w \rangle}{\|v\| \|w\|}$, and let $v_{\parallel W}$ be the orthogonal projection of v onto the linear subspace W . Finally, we will often use the subspace W of “marginal beliefs” defined as

$$W = \{\varepsilon \in \mathbb{R}^N : \langle 1, \varepsilon \rangle = 0\}. \quad (3)$$

This terminology follows from the fact that the difference between any two beliefs must lie in W .

3 The Value of Information Control under Open Disagreement

When is information control valuable to the sender? Our first contribution is to show that when players are subjected to a commonly understood signal the posterior belief of one player can be obtained from the posterior of another player, without explicit knowledge of the signal choice. This allows us to characterize the (subjective) distributions of posterior beliefs that can be induced by any signal (Proposition 1). Furthermore, it enables us to translate the search for an optimal signal to an auxiliary problem where the belief of each player is expressed in terms of the belief of a reference player, and then apply the techniques developed in KG to solve this auxiliary problem (Proposition 2). We then provide a simple necessary and sufficient condition for a sender to benefit from supplying information to an otherwise uninformed receiver (Corollary 1), and a necessary and sufficient condition for a sender to benefit from garbling a fully informative signal (Corollary 2). Finally, we contrast the gains from information control under open disagreement to the case where players share a common prior (Proposition 3).

3.1 Induced Distributions of Posterior Beliefs

From the sender's perspective, each signal π induces a (subjective) distribution over profiles of posterior beliefs. In any language-invariant equilibrium, the receiver's posterior belief uniquely determines his action. Therefore, two signals that, conditional on the state, induce the same distribution over profiles of beliefs generate the same value to the sender. That is, knowledge of the distribution of posterior beliefs suffices to compute the sender's expected utility from π .

If players share a common prior p , then following any realization of π players will also share a common posterior q . In this case, KG show that the martingale property of posterior beliefs $E^\pi[q] = p$ is both necessary and sufficient to characterize the set of distributions of beliefs that can be induced on Bayesian rational players by some signal. Consequently, KG are able to simplify the sender's problem by directly looking at this set of distributions, without having to specify the actual signal π that generates each distribution.

This leads us to ask: when players hold heterogeneous priors, what is the set of joint distributions of posterior beliefs that are consistent with Bayesian rationality? While it is still true that from the perspective of each player his expected posterior belief equals his prior, it is not true that the sender's expectation over the receiver's posterior belief always equals the receiver's prior. For instance, given any $p^S \neq p^R$, a signal π that is fully informative of the state implies that from the sender's perspective $E_S^\pi[q^R] = p^S \neq p^R$. Moreover, if there exist signal realizations z and z' such that both induce the same posterior q^S on the sender, but different posteriors q^R and $q^{R'}$ on the receiver (or vice versa), then knowledge of the specific signal would be necessary to compute the joint distribution of posteriors.

We next show that, given priors p^S and p^R , posteriors q^S and q^R form a bijection — q^R is derived from q^S through a perspective transformation. Moreover, this transformation is *independent* of the signal π and signal realization z . Proposition 1 establishes that the martingale property of the sender's beliefs and the perspective transformation (4) together characterize the set of distributions of posterior beliefs that are consistent with Bayesian rationality.

Proposition 1 *Let the totally mixed beliefs p^S and p^R be the prior beliefs of the sender and the receiver, and let r_θ^R be the state- θ likelihood ratio, $r_\theta^R = \frac{p_\theta^R}{p_\theta^S}$ with $r^R = \{r_\theta^R\}_{\theta \in \Theta}$. From the sender's perspective, a distribution over profiles of posterior beliefs $\tau \in \Delta(\Delta(\Theta) \times \Delta(\Theta))$*

is induced by some signal if and only if

(i) if $(q^S, q^R) \in \text{Supp}(\tau)$, then

$$q_\theta^R = q_\theta^S \frac{r_\theta^R}{\sum_{\theta' \in \Theta} q_{\theta'}^S r_{\theta'}^R} = \frac{q_\theta^S r_\theta^R}{\langle q^S, r^R \rangle}. \quad (4)$$

(ii) $E_\tau[q^S] = p^S$.

Proposition 1 shows that, in spite of the degrees of freedom afforded by heterogeneous priors, not all distributions over posterior beliefs are consistent with Bayesian rationality. Indeed, (4) implies that two signals that induce the same marginal distribution over the posterior beliefs of the sender must also induce the same marginal distribution over the posterior of the receiver. In fact, the set of joint distributions of players posterior beliefs under common priors and heterogeneous priors form a bijection. That is, belief disagreement does not allow the sender to generate “more” distributions of posterior beliefs. Equation (4) relies on both the assumptions of common support of priors and a commonly understood signal. One implication of a common support of priors is that any signal realization that leads the receiver to revise his belief must also induce a belief update by the sender — a signal realization is uninformative to the receiver if and only if it is uninformative to the sender.¹⁰ When players disagree on the likelihood functions that describe π (as is the case in Acemoglu et al, 2006 and Van den Steen 2011), then, even for Bayesian players, knowledge of the marginal distribution of posterior beliefs of one player may not be enough to infer the entire joint distribution, and thus it may not be enough to compute the sender’s expected utility from π .

Expression (4) affords a simple interpretation. Heterogeneous priors over θ imply that, for a given signal π , with signal space Z , players also disagree on how likely it is to observe each $z \in Z$. Just as the prior disagreement between the receiver and the sender is encoded in the likelihood ratio $r_\theta^R = p_\theta^R/p_\theta^S$,¹¹ we can encode the disagreement over z in the likelihood ratio

$$\lambda_z^R = \frac{\text{Pr}_R(z)}{\text{Pr}_S(z)}.$$

¹⁰If player j does not update his belief after observing z , then $q_\theta^j(z) = p_\theta^j$, implying that, for player i , $\langle q^j(z), r^i \rangle = 1$ and $q_\theta^j(z)r_\theta^i = p_\theta^j r_\theta^i = p_\theta^i$. Therefore, from (4) we must have $q_\theta^i(z) = p_\theta^i$.

¹¹For instance, a large class of measures of divergence between two probability distributions μ and ν take the form $\sum_{\theta \in \Theta} \mu_\theta f\left(\frac{\nu_\theta}{\mu_\theta}\right)$, which is the expectation of a (convex) function f of the likelihood ratio (Csiszar 1967).

The proof of Proposition 1 shows that this likelihood ratio can be obtained from r^R by

$$\frac{\Pr_R(z)}{\Pr_S(z)} = \langle q^S(z), r^R \rangle. \quad (5)$$

From (4) and (5) we can relate the updated likelihood ratio $q_\theta^R(z)/q_\theta^S(z)$ to r^R and λ_z^R ,

$$\frac{q_\theta^R(z)}{q_\theta^S(z)} = \frac{r_\theta^R}{\lambda_z^R}. \quad (6)$$

In words, the new state- θ likelihood ratio after updating based on z is obtained as the ratio of the likelihood ratio over states to the likelihood ratio over signal realizations. This implies that observing a signal realization z that comes more as a “surprise” to the receiver than the sender (so $\lambda_z^R < 1$) would lead to a larger revision of the receiver’s beliefs and thus a component-wise increase in the updated likelihood ratio. Moreover, both likelihood ratios (r_θ^R and λ_z^R) are positively related, in the sense that signals that come more as a surprise to the receiver than the sender are associated with states that the receiver believes to be less likely to occur.¹²

As a final remark, note that the likelihood ratio r^R is the Radon-Nikodym derivative of p^R with respect to p^S . Therefore (4) states that Bayesian updating under a commonly understood signal simply induces a linear scaling of the Radon-Nikodym derivative. Importantly, given the sender’s posterior belief, the proportionality factor does not depend on the signal π .

¹²Formally, given signal π , consider the probability distribution $\zeta^j(\theta, z)$ in $\Theta \times Z$ defined by $\zeta^j(\theta, z) = \pi(z|\theta)p_\theta^j$. Define the random variables $r^i(\theta, z) = r_\theta^i$ and $\lambda^i(\theta, z) = \lambda_z^i$. Then r^i and λ^i are positively (linear) correlated under $\zeta^j(\theta, z)$. To see this note that

$$\begin{aligned} \mathbb{E}_{\zeta^i} [\lambda^i r^i] &= \sum_{z \in Z} \sum_{\theta \in \Theta} \frac{\langle \pi(z), p^i \rangle}{\langle \pi(z), p^j \rangle} \frac{p_\theta^i}{p_\theta^j} \pi(z|\theta) p_\theta^j = \sum_{z \in Z} \left(\frac{\langle \pi(z), p^i \rangle}{\langle \pi(z), p^j \rangle} \right)^2 \langle \pi(z), p^j \rangle \\ &\geq \left(\sum_{z \in Z} \frac{\langle \pi(z), p^i \rangle}{\langle \pi(z), p^j \rangle} \langle \pi(z), p^j \rangle \right)^2 = 1 \\ \mathbb{E}_{\zeta^i} [r^i] &= \sum_{z \in Z} \sum_{\theta \in \Theta} \frac{p_\theta^i}{p_\theta^j} \pi(z|\theta) p_\theta^j = 1 \\ \mathbb{E}_{\zeta^i} [\lambda^i] &= \sum_{z \in Z} \sum_{\theta \in \Theta} \frac{\langle \pi(z), p^i \rangle}{\langle \pi(z), p^j \rangle} \pi(z|\theta) p_\theta^j = \sum_{z \in Z} \langle \pi(z), p^i \rangle = 1 \end{aligned}$$

3.2 Value of Information Control

The sender in our setup can neither use monetary incentives, nor restrictions on the receiver's choice set, to affect the latter's decisions. The only alternative available to change the receiver's decision is to (literally) change his beliefs over θ by providing signal π . Therefore, the sender's expected utility from π is uniquely determined by the sender's subjective distribution of posterior beliefs induced by π . In other words, if $\tau \in \Delta(\Delta(\Theta) \times \Delta(\Theta))$ represents a distribution over (q^S, q^R) , then the sender's problem can be written as

$$V(p^S, p^R) = \sup_{\pi} \mathbb{E}_{\tau} [v(q^S(z), q^R(z))] \quad (7)$$

s.t. τ is induced by π ,

where τ obtains from π and the sender's prior p^S , and the receiver's posterior q^R follows from applying Bayes' rule to the prior p^R .

We use Proposition 1 to translate the optimization problem (7), where the choice set are joint distributions of (q^S, q^R) , to the following equivalent, but lower dimensional, optimization problem, where the choice set are distributions over q^S .

$$V(p^S, p^R) = \sup_{\sigma} \mathbb{E}_{\sigma} [v(q^S, q^R)] \quad (8)$$

s.t. $\sigma \in \Delta(\Delta(\Theta)), \mathbb{E}_{\sigma} [q^S] = p^S, \left\{ q_{\theta}^R = \frac{q_{\theta}^S r_{\theta}^R}{\langle q^S, r^R \rangle} \right\}_{\theta \in \Theta}$,

where the receiver's posterior beliefs q^R are expressed through the perspective transformation (4) as a function of q^S .

The next Proposition establishes that an optimal signal exists, that it can use a limited number of distinct signal realizations, and computes the sender's expected utility under an optimal signal. For this purpose, and following KG, for an arbitrary real-valued function f define \tilde{f} as the concave closure of f ,

$$\tilde{f}(q) = \sup \{w | (q, w) \in co(f)\},$$

where $co(f)$ is the convex hull of the graph of f . In other words, \tilde{f} is the smallest upper semicontinuous and concave function that (weakly) majorizes the function f .

Proposition 2 (i) *An optimal signal exists. Furthermore, there exists an optimal signal with signal space Z such that $card(Z) \leq \min\{card(A), card(\Theta)\}$.*

(ii) Define the function V_S by

$$V_S(q^S) = v\left(q^S, \frac{q^S r^R}{\langle q^S, r^R \rangle}\right). \quad (9)$$

The sender's expected utility under an optimal signal is

$$V(p^S, p^R) = \tilde{V}_S(p^S). \quad (10)$$

The existence of an optimal signal in Proposition 2(i) follows from our assumption of a finite state space and our focus on equilibria for which v is upper semicontinuous. The characterization in Proposition 2(ii) follows from combining Proposition 1 and the insights provided by KG. Consider program (8). For any distribution σ over posterior beliefs of the sender induced by some signal, we have that the sender's expected utility is $E_\sigma\left[v\left(q^S, q^S \frac{r^R}{\langle q^S, r^R \rangle}\right)\right]$ with $E_\sigma[q^S] = p^S$. In other words, $\left(p^S, E_\sigma\left[v\left(q^S, q^S \frac{r^R}{\langle q^S, r^R \rangle}\right)\right]\right)$ belongs to the convex hull of the graph of the function V_S given by (9). Moreover, for any point (p^S, w) in the convex hull of the graph of V_S there exists a signal that induces $\sigma(p^S, w)$ over posteriors of the sender, and such that $w = E_{\sigma(p^S, w)}\left[v\left(q^S, q^S \frac{r^R}{\langle q^S, r^R \rangle}\right)\right]$ with $E_{\sigma(p^S, w)}[q^S] = p^S$. Therefore, the maximum expected utility of the sender is $\sup\{w \mid (p^S, w) \in co(V_S)\} = \tilde{V}_S(p^S)$.

Our model is essentially static as, once signal π is selected, all learning is performed after observing its realization z . Could the sender strictly benefit from further releasing information contingent on z ? Releasing further information is tantamount to inducing a different distribution over posteriors that still has to satisfy Bayesian rationality. In particular, Proposition 1 still holds for the composition of multiple signals. It follows that the sender cannot increase her expected utility by sequentially releasing information, since the posterior beliefs under sequential updating can be replicated with a single signal that induces the same distribution over beliefs given the state.¹³ Thus, in contrast to Brocas and Carrillo (2007), sequential disclosure has no value as the set of signals available to the sender is sufficiently rich.

Proposition 2 shows that the value of information control is $\tilde{V}_S(p^S) - V_S(p^S)$. Direct application of Proposition 2 to establish whether a sender benefits from information control would require the derivation of the concave closure of an upper semicontinuous function, a

¹³This is similar to the observation made by KG that a sender with full commitment cannot strictly benefit from sequential disclosure. In our case this remains true given our assumptions of a commonly understood signal and totally mixed priors.

task typically not amenable to standard algorithms. Nevertheless, the following Corollary provides conditions that make it easier to verify if information control is valuable.

Corollary 1 *There is no value of information control if and only if there exists a vector $\gamma \in \mathbb{R}^{\text{card}(\Theta)}$ such that*

$$\langle \gamma, q^S - p^S \rangle \geq V_S(q^S) - V_S(p^S), \quad q^S \in \Delta(\Theta). \quad (11)$$

In particular, if V_S is differentiable at p^S , then there is no value of information control if and only if

$$\langle \nabla V_S(p^S), q^S - p^S \rangle \geq V_S(q^S) - V_S(p^S), \quad q^S \in \Delta(\Theta). \quad (12)$$

This Corollary provides a geometric condition for a sender not to benefit from information control: a sender optimally releases no information if and only if V_S admits a supporting hyperplane at p^S . It is immediate to see that (11) is sufficient: consistent beliefs require that, for any signal that induces distribution σ over q^S , we must have $E_\sigma[q^S] = p^S$, implying $E_\sigma[\langle \gamma, q^S - p^S \rangle] = 0$ and $0 \geq E_\sigma[V_S(q^S)] - V_S(p^S)$. Conversely, Proposition 2 establishes that if there is no value of information control, then $\tilde{V}_S(p^S) = V_S(p^S)$. As $-\tilde{V}_S$ is a proper, convex function, any element from the non-empty set of subdifferentials $\partial(-\tilde{V}_S(p^S))$ would provide a majorizing affine function to \tilde{V}_S and hence to V_S .

We conclude this section by pointing out that in some applications it will be convenient to rewrite the sender's problem as follows. Define a new utility function for the sender,

$$\check{u}_S(a, \theta) = \frac{u_S(a, \theta)}{r_\theta^R}. \quad (13)$$

For any signal $\pi = (Z, \{\pi(\cdot|\theta)\}_{\theta \in \Theta})$ and the receiver's decision rule $a(z)$, $z \in Z$, we have

$$E_S[u_S(a(z), \theta)] = \sum_{\theta \in \Theta} \sum_{z \in Z} \pi(z|\theta) p_\theta^S u_S(a(z), \theta) = \sum_{\theta \in \Theta} \sum_{z \in Z} \pi(z|\theta) p_\theta^R \frac{u_S(a(z), \theta)}{r_\theta^R} = E_R[\check{u}_S(a(z), \theta)].$$

That is, given the receiver's behavior, the expected utility of a sender with prior p^S and utility u_S is the same as the expected utility of a sender who shares the receiver's prior p^R , but has utility \check{u}_S . Therefore, under a commonly understood signal one can convert the sender's original problem to one with common priors as follows. Rewrite (1) as $\check{v}(q^S, q^R) \equiv \sum_{\theta \in \Theta} q_\theta^S \check{u}_S(a(q^R), \theta)$, and define

$$V_R(q^R) = \check{v}(q^R, q^R). \quad (14)$$

Then the claims of Proposition 2 remain valid if one substitutes $V_R(q^R)$ for $V_S(q^S)$. However, note that in many cases the transformed utility \tilde{u}_S is hard to interpret and defend on economic grounds. Moreover, by maintaining the original formulation one is able to gather a better economic understanding of the effects of a commonly understood signal on heterogeneous priors. For example, an important result in Section 4 is that on the space of priors the sender generically benefits from information control. Such result would be hard to postulate and interpret if one only examines the transformed problem.

3.3 The Value of Garbling

Our previous analysis is well suited to cases where, absent the sender’s signal, receivers would not be able to acquire further information on their own. In many situations, however, a sender’s influence takes the form of obfuscation or “signal jamming”, i.e. a sender attempts to “confound” receivers by garbling the information that would otherwise reach them. Corollary 2 provides a simple necessary and sufficient condition for a sender to benefit from introducing noise into a fully informative signal observed by the receiver. That is, under these conditions a fully revealing signal does not solve the sender’s problem defined by (8). For this purpose, let $\mathbf{1}_\theta$ be the posterior belief that puts probability 1 on state θ .

Corollary 2 *A sender does not benefit from garbling a perfectly informative signal if and only if*

$$\sum_{\theta \in \Theta} q_\theta^S u_S(a(\mathbf{1}_\theta), \theta) \geq V_S(q^S), \quad q^S \in \Delta(\Theta). \quad (15)$$

Condition (15) admits a simple interpretation. Suppose that players observe a signal realization that induces q^S in the sender. The right hand side of (15) is the sender’s expected utility if she discloses no more information, while the left hand side of (15) is the sender’s expected utility if she allows the receiver to perfectly learn the state. Then a sender does not benefit from garbling a perfectly informative signal if and only if after *every* possible signal and signal realization she is not worse off by fully revealing the state.

4 Pure Persuasion

In this section we apply our results to the case when the sender’s utility is independent of the state, i.e. the case of “pure persuasion”. We first show that there is always a prior belief disagreement that renders information control valuable. We then characterize when and why a sender values information control, as a function of the players preferences and the extent of prior belief disagreement. In particular, we show that if the receiver’s action is a linear function of his beliefs, then information control is generically valuable and the optimal signal is often not fully revealing of the state.

4.1 The Role of Heterogenous Priors

What are the possible reasons for a sender to benefit from designing a receiver’s access to information? The literature has explored two broad sources of value from information control under the assumption of a common prior. One source is based on the value of information: a sender who benefits from adapting decisions to the underlying state would certainly benefit from providing an informative signal to a decision maker that shares her preferences. The other source is based on conflicting interests. For instance, under pure persuasion, the sender draws no value from knowing the state if she could make decisions herself. However, KG and Brocas and Carrillo (2007) show that she can still benefit from information control if it is a receiver who instead makes decisions — when players share a common prior, information control is valuable when the sender can exploit the non-concavity of the receiver’s action in his beliefs, or the convexity of the sender’s utility function in the receiver’s actions.

We now argue that open disagreement provides a third, distinct rationale for a sender to benefit from information control. To make our point as clear as possible, Proposition 3 considers a pure persuasion setup where $u_S(a(q^R))$ is everywhere concave, so that both previous rationales are absent: under common priors the sender does not benefit from information control as the function V_S given by (9) is everywhere concave. Proposition 3 shows that belief disagreement can reverse this result.

Proposition 3 *(i) Suppose that $u_S(a(q^R))$ is twice-continuously differentiable and for each belief $q^R \in \Delta(\Theta)$ the Hessian Matrix of $u_S(a(q^R))$ is negative definite. Then for any totally*

mixed prior p^R there exists a neighborhood of p^R such that a sender with prior belief $p^S \in N(p^R)$ does not benefit from information control.

(ii) For every bounded u_S and totally mixed prior p^R for which

$$u_S(a(p^R)) < \max_{q^R \in \Delta(\Theta)} u_S(a(q^R)),$$

there exists a totally mixed p^S such that a sender with prior p^S benefits from information control.

Proposition 3(i) states that, as long as $u_S(a(q^R))$ is strictly concave for all directions in which beliefs may be updated, then small belief disagreements are not sufficient for a sender to provide some information to a receiver. Nevertheless, Proposition 3(ii) shows that if the receiver is not already choosing the sender's preferred decision, then there always exists a level of prior belief disagreement such that information control is valuable. The logic of the proof is simple: if the sender's utility increases when the receiver has a belief $\tilde{q}^R \neq p^R$, then one can construct a signal π and a belief p^S such that a sender with prior p^S expects signal π to induce \tilde{q}^R almost certainly. Interestingly, it is not a sender with prior belief $p^S = \tilde{q}^R$ the one who is most confident of inducing \tilde{q}^R in the receiver. Indeed, the proof of the Proposition constructs a signal π that induces two different posteriors, where one of them is \tilde{q}^R , and shows that a sender becomes more confident of inducing \tilde{q}^R through π as her prior belief puts more probability on the state θ' that maximizes $\tilde{q}_\theta^R/p_\theta^R$, i.e. on the state θ' such that $\tilde{q}_{\theta'}^R/p_{\theta'}^R \geq \tilde{q}_\theta^R/p_\theta^R$, $\theta \in \Theta$.

4.2 Value of Information Control under Pure Persuasion

When does a sender benefit from providing an informative signal to a receiver under pure persuasion? To answer this question, we rewrite function V_R defined in (14) as

$$\begin{aligned} V_R(q^R) &= E_R [\check{u}_S(a(q^R))] = u_S(a(q^R)) E_R \left[\frac{1}{r^R} \right] \\ &= u_S(a(q^R)) \langle q^R, r^S \rangle, \end{aligned} \tag{16}$$

with $r_\theta^S = p_\theta^S/p_\theta^R$, $r^S = \{r_\theta^S\}_{\theta \in \Theta}$. Representation (16) suggests that to understand the sender's gain from information control one should consider the sender's risk preferences over

decisions, the shape of the receiver's actions given his beliefs, and the extent of prior belief disagreement as captured by $\langle q^R, r^S \rangle = \Pr_S(q^R) / \Pr_R(q^R)$.

To simplify the exposition, we assume $A \subset \mathbb{R}$ and $u'_S > 0$ so that the sender's utility is increasing in the receiver's action. If the sender and the receiver share a common prior, then $\langle q^R, r^S \rangle = 1$ and the value of information control is obtained directly from the curvature of $u_S(a(q^R))$. Given $u'_S > 0$, a concave u_S and $a(q^R)$ imply that $u_S(a(q^R))$ is concave and the sender does not benefit from information control. However, if u_S is strictly convex and $a(q^R)$ is convex, then $u_S(a(q^R))$ is strictly convex and the sender benefits from information control. That is, as shown by KG, under a common prior belief the sender can benefit from the provision of a signal by exploiting non-concavities in the receiver's action, or her own positive attitude towards risk.

Proposition 4 emphasizes the role of the curvature of the receiver's action. We first establish that the sender can exploit non-concavities in the action of the receiver to her advantage, irrespective of her risk attitudes and of the extent of belief disagreement. We then characterize situations in which the sender benefits from information control even when $a(q^R)$ is concave.

Proposition 4 *Suppose $A \subset \mathbb{R}$, $u'_S > 0$ and $a(q^R)$ is twice continuously differentiable. Let $A^+ = \{q^R \in \Delta(\Theta) : a(q^R) > a(p^R)\}$ be the (open) upper contour set of the receiver's action at the prior belief p^R , $T = \{q^R \in \Delta(\Theta) : \langle \nabla a(p^R), q^R - p^R \rangle = 0\}$ be the tangent hyperplane of $a(q^R)$ at p^R , $H(a(p^R))$ be the Hessian matrix of $a(q^R)$ at $q^R = p^R$, and W defined in (3).*

(i) *Suppose that $\nabla a(p^R)_{\parallel W} \neq 0$. If $T \cap A^+ \neq \emptyset$, then the value of information control is positive for all $p^S \in \text{int}(\Delta(\Theta))$ and for all strictly increasing u_S . In particular, if the restriction of $H(a(p^R))$ to T is not negative semidefinite then the value of information control is positive.*

(ii) *Suppose that $a(q^R)$ is concave at $q^R = p^R$. Let λ_{\min} be the smallest eigenvalue of $H(a(p^R))$, and define*

$$\begin{aligned} m &= \nabla a(p^R), \\ n &= \frac{u''_S(a(p^R))}{u'_S(a(p^R))} \nabla a(p^R) + 2r^S. \end{aligned}$$

If the projections $m_{\parallel W}$ and $n_{\parallel W}$ of m and n on W satisfy

$$\frac{1}{2} \|m_{\parallel W}\| \|n_{\parallel W}\| \left[1 + \cos(m_{\parallel W}, n_{\parallel W}) \right] > |\lambda_{\min}|, \quad (17)$$

then the sender benefits from information control.

As the proof of the Proposition shows, under the conditions of Proposition 4(i) the sender can always find a signal such that every signal realization induces the receiver to choose a strictly higher action. Therefore, a sender with monotone preferences will increase her expected utility with this signal, irrespective of her risk attitudes and of the extent of belief disagreement. Suppose now that the action of the receiver is concave in his beliefs, so that the set of posterior beliefs that weakly raise the receiver's action is convex and the conditions of Proposition 4(i) do not hold. The martingale property of posterior beliefs implies that any signal that induces higher actions in the receiver must also have signal realizations that lead the receiver to choose a lower action. When this is the case, the sender's risk preferences and the extent of prior belief disagreement play a role in dictating whether the sender benefits from information control. Proposition 4(ii) provides conditions such that information control can be valuable, even when $u_S(a(q^R))$ is concave. Recall from (16) that V_R is the product of $u_S(a(q^R))$ and the concave function $\langle q^R, r^S \rangle$. Condition (17) guarantees that the product of these two concave functions is locally strictly convex in at least one direction of feasible posterior beliefs.

4.3 Persuading Skeptics and Believers

Proposition 4(ii) provides sufficient conditions for the sender to benefit from information control when she cannot depend on non-concavities in the receiver's action. In this section we maintain the assumption $u'_S > 0$ and restrict attention to the subcase of Proposition 4(ii) where the receiver's action exhibits linear increments in beliefs. This assumption is equivalent to the existence of a random variable x such that $a(q^R)$ satisfies

$$a(q^R) = \sum_{\theta \in \Theta} q_{\theta}^R x(\theta) = \langle q^R, x \rangle, \quad (18)$$

where, to avoid trivialities, we assume that x is non-constant.

This action choice is consistent with a receiver with preferences $u_R(a, \theta) = -(a - x(\theta))^2$. For instance, in many political economy models and in the example in the Introduction, action a can be interpreted as a policy choice in a left-right policy spectrum. Alternatively, (18) can be derived in a moral hazard setup in which $u_R(a, \theta) = x(\theta)a - \frac{a^2}{2}$, where $x(\theta)$ is the receiver's marginal benefit of effort, and $\frac{a^2}{2}$ is his personal cost of effort; or in a resource allocation problem with an infinitely divisible budget of 1 which the receiver needs to allocate between two projects, when his utility from allocating a to the first project is $u_R(a, \theta) = x(\theta) \ln a - (1 - x(\theta)) \ln(1 - a)$. In all these cases, the sender would like to induce the highest possible action by providing information that induces in the receiver the highest possible expectation of x .¹⁴

The specification (18) allows a simple categorization of the type of receiver that the sender may face. A sender views a receiver as holding adverse beliefs if she would be made better off by a receiver who shares her point of view, that is, if

$$\langle q^R, x \rangle < \langle q^S, x \rangle. \quad (19)$$

Conversely, a sender views a receiver as holding favorable beliefs if she would not be made better off by the receiver sharing her point of view, that is, if

$$\langle q^R, x \rangle \geq \langle q^S, x \rangle. \quad (20)$$

When (19) holds we refer to the receiver as a “skeptic,” and when (20) holds as a “believer”.

If the sender faces a skeptic, a fully revealing signal would raise her expectation over the receiver's actions. If instead the sender faces a believer, a fully revealing signal would (weakly) decrease her expectation over the receiver's actions. Whether such signal raises or decreases the sender's expected utility will depend on her risk preferences. Nevertheless,

¹⁴Our results in this Section translate readily to the more general setup where, for each q^R , the indirect utility of the sender can be written as an increasing function of the receiver's expectation of x , $u_S(a(q^R)) = F(E_R[x(\theta)])$, with $F'(\cdot) \geq 0$. For example, consider a receiver who takes a binary action $\{0, 1\}$, and chooses action 1 if and only if $E_R[\theta] \geq \eta$ (e.g., vote for candidate A or B, approve or not approve a project, vote to convict or to acquit a defendant). The random variable η follows some distribution F , is orthogonal to θ , and becomes public information after the sender chooses the signal, but before the receiver chooses his action. The sender receives payoff 1 if action 1 is taken, and zero otherwise. In this case, the sender's expected utility becomes $E_S[u_S(a^*(q^R))] = Pr(E_R[\theta] \geq \eta) = F(E_R[\theta])$.

Proposition 5 shows that when the sender has access to a richer set of signals and the state space includes at least three states, then the sender *generically* benefits from information control, regardless of whether she is facing a skeptic or a believer, and regardless of her risk attitudes.¹⁵ To present our results we recall the following definition.

Definition: Vectors v and w are *negatively collinear with respect to the subspace Q* if there exist $\lambda < 0$ such that the projections $v_{||Q}$ and $w_{||Q}$ satisfy

$$v_{||Q} = \lambda w_{||Q}. \quad (21)$$

In particular, by considering $Q = W$, with W defined by (3), condition (21) is equivalent to the existence of λ_0 and $\lambda_1 > 0$ such that

$$v = \lambda_0 \mathbf{1} - \lambda_1 w,$$

or, alternatively, to the existence of $\lambda_1 > 0$ such that

$$v_\theta - v_{\theta'} = -\lambda_1 (w_\theta - w_{\theta'}), \quad \theta, \theta' \in \Theta. \quad (22)$$

Our interest in this definition is given by the following Lemma.

Lemma 1 *Every signal realization z of a signal π that increases the receiver's action is perceived to be more likely by the receiver than by the sender if and only if x and r^S are negatively collinear with respect to W , defined by (3).*

We now state our main proposition in this Section.

Proposition 5 *Suppose that the receiver's action is given by (18), $\text{card}(\Theta) > 2$, u_S is twice continuously differentiable with $u'_S > 0$, and $p^R \neq p^S$.*

(i) *If x and r^S are not negatively collinear w.r.t. W , then the sender benefits from information control.*

(ii) *If u_S is concave, then the sender benefits from information control if and only if x and r^S are not negatively collinear w.r.t. W .*

¹⁵Genericity is interpreted over the space of pairs of prior beliefs.

When there are at least three states and regardless of the curvature of u_S , Proposition 5(i) implies that the sender benefits from information control whenever she can construct a signal π and a signal realization z to which she assigns more probability than the receiver, and z increases the receiver’s action (by Lemma 1). Moreover, if u_S is concave, so that information control has no value under a common prior, then information control has value under heterogenous priors if and only if x and r^S are not negatively collinear w.r.t. W . One important implication of Proposition 5, given Lemma 1, is that if the state space is rich enough and x is injective (i.e. takes different values for different states), then the sender *generically* benefits from information control under belief disagreement, regardless of the curvature of the utility function u_S .

To provide some intuition for Proposition 5, define $\Lambda = \{q^R : \langle q^R - p^R, x \rangle = 0, q^R \in \Delta(\Theta)\}$, which is a hyperplane of beliefs that includes the prior of the receiver, and such that the receiver’s action is constant in Λ . One can then find signals supported on Λ that leave the expected utility of the sender unchanged. Moreover, if x and r^S are not collinear, then the sender and the receiver generically disagree over the likelihood of any posterior q^R in Λ . The sender can then exploit this disagreement by switching to a signal that modifies the posterior of the receiver in the direction of a higher action only for those beliefs that the sender perceives as more likely than the receiver. If the state is binary, however, then Λ is a singleton and the previous argument cannot be applied.

4.4 Garbling Information to Skeptics and Believers

In many situations, the effect of lobbying is to reduce the amount of information that reaches decision makers. To examine these situations, suppose that the receiver perfectly learns the state if the sender does not engage in information control. When would the sender benefit from reducing the information that reaches the receiver? To answer this question we apply Corollary 2 to the function V_S in (9) when the receiver’s action satisfies (18), which here takes the simple form

$$V_S(q^S) = u_S \left(\frac{\langle q^S, r^R x \rangle}{\langle q^S, r^R \rangle} \right). \quad (23)$$

Expression (23) suggests that the sender’s gain from garbling depends both on her “risk attitudes” (i.e. on the curvature of u_S) and the type of receiver she is facing. The next

proposition formalizes this intuition.

Proposition 6 (i) Suppose that u_S is convex, x non-decreasing in θ , and¹⁶ $p^S \succeq_{LR} p^R$. Then the sender does not benefit from garbling. (ii) Suppose that u_S is absolutely continuous and there exist states θ and θ' such that

$$(x(\theta') - x(\theta)) \left((r_{\theta'}^S)^2 u'_S(x(\theta')) - (r_{\theta}^S)^2 u'_S(x(\theta)) \right) < 0. \quad (24)$$

Then the sender benefits from garbling.

It is immediate to see that any garbling reduces the variance of the receiver's posterior beliefs. Moreover, likelihood ratio orders are preserved under Bayesian updating. In particular, if $p^S \succeq_{LR} p^R$, then the receiver will remain a skeptic after any signal realization that does not fully reveal the state, meaning that by fully revealing the state the sender can increase, on average, the receiver's action. Proposition 6(i) establishes that when u_S is convex and the receiver remains a skeptic after every partially informative signal, then the sender cannot do better than letting the receiver fully learn the state. That is, garbling is not valuable. Nevertheless, Proposition 6(ii) argues that if at least one of these conditions is relaxed, then garbling is valuable as long as (24) is satisfied. For example, it follows from (24) that if u_S is linear and $p^S \not\prec_{LR} p^R$, then the sender benefits from garbling, even if the receiver is a skeptic. Proposition 6(ii) also implies that if $p^R \succeq_{LR} p^S$ and u_S is concave, then the sender would optimally restrict the information available to the receiver.

4.5 Optimal Signal

When information control is valuable, what is the optimal signal? To provide some intuition, we now restrict attention to the case in which the sender's utility is linear, $u_S = \beta a$, $\beta \neq 0$, and $a(q^R)$ satisfies (18) with $x = \theta$, so that the receiver's action is the expected state. The sender's optimal signal maximizes $E_S[E_R[\theta|\pi]]$ when $\beta > 0$, and minimizes $E_S[E_R[\theta|\pi]]$ when $\beta < 0$. With a common prior, this expectation is constant in the space of all signals, thus information control is not valuable. With heterogenous priors, however, the sender can always find a signal that increases or decreases, on average, the receiver's expectation over θ , as long as $\theta_{\parallel W}$ and $r_{\parallel W}^S$ are not collinear.

¹⁶The order \succeq_{LR} denotes the likelihood ratio order as in Shaked and Shanthikumar (2007, pg 42).

Corollary 3 *Suppose that the state θ and the likelihood ratio p_θ^S/p_θ^R are not collinear w.r.t. W . Then there exist signals π and π' such that $E_S [E_R [\theta|\pi]] < E_R [\theta] < E_S [E_R [\theta|\pi']]$.*

Proposition 7 characterizes an optimal signal when $\beta > 0$, and it is straightforward to restate the proposition in the case $\beta < 0$.

Proposition 7 *Suppose that $u_S = \beta a$, $\beta > 0$, and $a(q^R)$ satisfies (18) with $x = \theta$. Then*

- (i) after each realization of an optimal signal the receiver is a believer;*
- (ii) if every combination of three elements of θ and r^S are not negatively collinear, then after each realization of an optimal signal the receiver puts positive probability in at most two states;*
- (iii) a completely uninformative signal is optimal if and only if $\theta_{||W}$ and $r_{||W}^S$ are negatively collinear;*
- (iv) a fully revealing signal is optimal if and only if $p^S \succeq_{LR} p^R$.*

Proposition 7 obtains from Propositions 5 and 6 with the aid of two simple observations. First, optimization by the sender implies that there is no value in further releasing any information after any signal realization. In particular, the conditions in Proposition 5 must hold after each realization of an optimal signal. Second, negative collinearity of θ and r^S w.r.t. W and likelihood ratio order relations between beliefs are both preserved under Bayesian updating¹⁷. This follows trivially from (6) as Bayesian updating induces a rescaling of likelihood ratios.

Proposition 7(i) follows from the fact that if a signal realization leaves the receiver being a skeptic, then the sender would strictly benefit from fully disclosing the state, contradicting the premise of optimality. Proposition 7(ii) exploits the invariance of collinearity under Bayesian updating: if no three components of θ and r^S are negatively collinear, then an optimal signal must narrow down the receiver's uncertainty to at most two states. Proposition 7(iii) and (iv) follow immediately from Propositions 5 and 6, respectively.

We now apply our results to solve for the optimal signal in the example presented in the Introduction. There are three possible states, $\theta \in \{0, 0.5, 1\}$, and the sender's utility is linear in the receiver's expectation of θ . Consider priors $p^R = (0.4, 0.5, 0.1)$ for the receiver and

¹⁷To be precise, this is true when considering only the elements in the support of the posteriors.

$p^S = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ for the sender, so that the receiver is a skeptic. The likelihood ratio and the state are not negatively collinear, hence the sender benefits from persuasion (Proposition 5). Moreover, Proposition 2 implies that there is an optimal signal with at most three signal realizations. We now use Proposition 7 to solve the problem without having to explicitly derive the concave closure of V_S . First, after every signal realization the players must attach positive probability to at most two states — cf. Proposition 7(ii). Second, after each signal realization the players cannot attach positive probabilities to states 0 and 1 at the same time, nor to states 0.5 and 1 at the same time, otherwise the receiver would be a “skeptic” — cf. Proposition 7(i). Consequently, after each realization of an optimal signal, players must know with certainty whether state 1 occurred or not. Does the sender benefit from further disclosing information about states 0 and 0.5? Note that conditional on learning that state 1 has not occurred, the receiver becomes a “believer”, and there are only two possible states left. Further information disclosure is not beneficial since likelihood ratios and the state are negatively collinear in the partition $\{0, 0.5\}$. Thus, the optimal signal only reveals whether $\theta = 1$ or not.

Now consider the “believer” case in the second part of the example, $p^R = (0.1, 0.5, 0.4)$. After each signal realization, individuals cannot assign positive probability to states 0.5 and 1 at the same time, otherwise θ and r^S are not negatively collinear and further information disclosure is optimal. Moreover, conditional on learning that the state is not 1, further information disclosure is not optimal because θ and r^S are negatively collinear in the partition $\{0, 0.5\}$. Similarly, conditional on learning that the state is not 0.5, further information disclosure is not optimal. Therefore, we can focus on a binary signal $\{z_L, z_H\}$, where state 0.5 generates signal z_L with probability one, state 1 generates signal z_H with probability one, and state 0 generates signal z_L with probability α and z_H with probability $1 - \alpha$. In this example, the sender’s expected utility decreases in α ,¹⁸ hence her optimal choice is $\alpha = 0$.

¹⁸Her expected utility is $(\frac{1}{3} + \alpha\frac{1}{3}) \left(0.5\frac{0.5}{0.5+\alpha 0.1}\right) + (\frac{1}{3} + (1 - \alpha)\frac{1}{3}) \left(1\frac{0.4}{0.4+(1-\alpha)0.1}\right)$.

5 Private Priors

So far we have assumed that the sender knows the prior belief of the receiver. It is immediate to extend the analysis to a case in which the sender is uncertain about the receiver's prior beliefs when designing the signal π . Suppose for concreteness that prior beliefs are drawn from a distribution $H(p^R, p^S)$ with conditional distribution $h(p^R|p^S)$.¹⁹ Proposition 1 still applies for each (p^R, p^S) . Consequently, given p^S and $h(p^R|p^S)$, knowledge of the sender's posterior q^S suffices to compute the joint distribution of posterior beliefs. Moreover, the restriction to language-invariant equilibria implies that, given realization (p^R, p^S) , the receiver's choice only depends on his posterior belief q^R . Therefore, after a signal realization that induces posterior q^S , we can compute the sender's expected payoff V_S using the implied distribution of q^R . More specifically, (9) translates to

$$V_S(q^S) = E_S[v(q^S, q^R)|p^S] = \int v\left(q^S, \frac{q^S p^R}{p^S}\right) dh(p^R|p^S). \quad (25)$$

With this modification, the expected utility of a sender under an optimal signal is $\tilde{V}_S(p^S)$ and the sender would benefit from information control under the conditions of Corollary 1. Moreover, the expected value to the sender of a perfectly informative signal is independent of the receiver's prior belief. Therefore, the value of garbling is positive whenever (25) satisfies the conditions in Corollary 2.

As an application of (25), consider the pure persuasion model from Section 4. When the sender knows the receiver's prior, Proposition 5(i) provides conditions on the likelihood ratio of priors such that information control is valuable. Suppose these conditions are met and the sender strictly benefits from providing signal π to a particular receiver. By a continuity argument, the same signal π strictly benefits the sender when she faces another receiver whose beliefs are not too different. Consequently, even if the sender does not know the receiver's prior, information control remains beneficial when the receiver's possible priors are not too dispersed. Proposition 8 provides an upper bound on how dispersed these beliefs can be. To

¹⁹Note that the receiver's preferences are unaffected by his beliefs about the sender's prior. Therefore, the sender's choice of signal conveys no additional information to the receiver. This would not be true if the sender privately observes a signal about the state. See Sethi and Yildiz (2012) for a model of communication where players have private prior beliefs and also receive a private signal about the state.

this end, let R be the set of likelihood ratios induced by the priors in the support of $h(p^R|p^S)$,

$$R = \{r^R : \{r_\theta^R = p_\theta^R/p_\theta^S\}_{\theta \in \Theta}, p^R \in \text{Supp}(h(p^R|p^S))\}. \quad (26)$$

Proposition 8 *Suppose that r^R and $r^R x$ are not collinear w.r.t. W for all $r^R \in R$, and let $m = \frac{1}{2} \frac{\max |u_S''(a)|}{\min u_S'(a)} > 0$. If for all $r^R, r^{R'} \in R$*

$$\|r^R - r^{R'}\| \leq \beta, \quad (27)$$

with β given by (46), then the sender benefits from information control.

The condition on r^R and $r^R x$ implies that if the sender knew the receiver's prior, then she could find a signal with a positive value (cf. Proposition 5). The bound $\beta(m)$ is defined by (46) in the Appendix B, as a function of the curvature of u_S . From (27), $\beta(m)$ represents a lower bound on the cosine of the angle between any two likelihood ratios in the support of $h(p^R|p^S)$. Therefore, (27) describes how different the receiver's possible prior beliefs can be for the sender still to benefit from information control, by imposing an upper bound on the angle between any two likelihood ratios in R .

6 Conclusion

In this paper we study the gain to an individual (sender) from controlling the information available to a decision maker (receiver) when they openly disagree on their views of the world. Our first contribution is to characterize the set of distributions over posterior beliefs that can be induced through a signal, under our assumption of ‘‘commonly understood signal’’ (i.e., when players agree on the statistical relation of the signal to the payoff-relevant state). This allows us to compute the gain from information control, both when the receiver would otherwise remained uninformed and when the receiver would perfectly learn the state absent the sender's influence. One implication of our analysis is that differences in prior beliefs are a separate rationale for persuasion, and that, under mild conditions, there always exists a difference in prior beliefs that renders information control valuable.

In Section 4 we apply our results to a large class of pure persuasion models, where the sender's payoff is an increasing function of the receiver's expectation of a random variable.

One could think that a sender would be hurt by providing information to an overly optimistic receiver, as she expects information to corroborate her pessimistic point of view. However, we show that if the state space is rich enough, then the sender *generically* benefits from providing some information, even when facing an overly optimistic receiver. Moreover, this result does not depend on the sender’s risk-attitudes. We then analyze the gain from garbling an otherwise fully informative signal. We show that a sender may benefit from garbling the signal even in situations when the receiver holds overly pessimistic beliefs.

To focus on the impact of heterogeneous priors on information control, we have restricted our analysis in several ways. First, we have eschewed the possibility that the sender has private information. Second, we consider a single receiver. In many situations, however, the sender may want to affect the beliefs of a collective, where she is typically constrained to use a public signal. Third, we have considered a fixed decision making process. However, in some instances the sender can both offer a contract and provide some information to a receiver, i.e. the sender designs a grand-mechanism that specifies the information to be released and several contractible variables. Similarly, one can examine how the optimal signal varies across different mechanisms of preference aggregation (e.g., Alonso and Câmara 2014 examine information control in a voting model). We leave these promising extensions for future work.

A Proofs

Proof of Proposition 1: *Necessity*: Consider a signal $\pi = (Z, \{\pi(\cdot|\theta)\}_{\theta \in \Theta})$ that induces, from the sender’s perspective, the distribution τ and let $\pi(z) = \{\pi(z|\theta)\}_{\theta \in \Theta}$ and $q^R(z)$ and $q^S(z)$ be the posterior beliefs of the receiver and the sender if $z \in Z$ is realized. Clearly, the marginal distribution over the sender’s posterior beliefs satisfies the martingale property, i.e. $E_\tau[q^S] = p^S$. Furthermore, as priors are totally mixed, the receiver assigns positive probability to z if and only if the sender also assigns positive probability to z .²⁰ Suppose then that $\pi(z) \neq \mathbf{0}$. Bayesian updating implies that, after observing z , the sender’s posterior is

$$q_\theta^S(z) = \frac{\pi(z|\theta)p_\theta^S}{\langle \pi(z), p^S \rangle},$$

²⁰Indeed, we have $\Pr_R[z] = \langle \pi(z), p^R \rangle = 0 \Leftrightarrow \pi(z|\theta) = 0, \theta \in \Theta \Leftrightarrow \Pr_S[z] = \langle \pi(z), p^S \rangle = 0$.

so we can write

$$q_\theta^S(z) \langle \pi(z), p^S \rangle \frac{p_\theta^R}{p_\theta^S} = \pi(z|\theta) p_\theta^R,$$

and summing over $\theta \in \Theta$ we obtain

$$\langle \pi(z), p^S \rangle \langle q^S(z), r^R \rangle = \langle \pi(z), p^R \rangle.$$

Then we can relate the two posterior beliefs by

$$q_\theta^R(z) = \frac{\pi(z|\theta) p_\theta^R}{\langle \pi(z), p^R \rangle} = \frac{\pi(z|\theta) p_\theta^S}{\langle \pi(z), p^S \rangle} \frac{p_\theta^R}{p_\theta^S} = q_\theta^S(z) \frac{r_\theta^R}{\langle q^S(z), r^R \rangle}.$$

Sufficiency: Given a distribution τ satisfying (i) and (ii), let $\tau_S(q^S)$ be the marginal distribution of the sender's posterior beliefs and define the signal space $Z = \{q^S : q^S \in \text{Supp}(\tau_S)\}$ and the likelihood functions $\pi(q^S|\theta) = \frac{q_\theta^S \text{Pr}_{\tau_S} q^S}{p_\theta^S}$. Then simple calculations reveal that the signal $\pi = \left(Z, \{\pi(q^S|\theta)\}_{\theta \in \Theta} \right)$ induces τ . ■

Proof of Proposition 2: Part (i) See KG. Part (ii) In the text.

Proof of Corollary 1: The first part of the claim can be rephrased in terms of the subdifferential $\partial V(p)$ of a function V evaluated at p , which we take to be the set of linear functionals f such that

$$f(q - p) \leq V(q) - V(p), \quad q \in \mathbb{R}^N.$$

With this terminology, the first part of Corollary 1 states that the sender does not benefit from information control if and only if $\partial(-V_S(p^S)) \neq \emptyset$. The second part of Corollary 1 then follows immediately as, if V_S is differentiable at p^S , then $\partial(-V_S(p^S))$ can have at most one element.

Sufficiency: As the concave closure \tilde{V}_S is the lower envelope of all affine functions that majorize V_S and, by assumption, the majorizing affine function $f(q^S) = V_S(p^S) + \langle \gamma, q^S - p^S \rangle$ satisfies $V_S(p^S) = f(p^S)$, then

$$V_S(p^S) = f(p^S) \geq \tilde{V}_S(p^S) \geq V_S(p^S),$$

implying that $\tilde{V}_S(p^S) = V_S(p^S)$ and, by Proposition 2, there is no value of information control.

Necessity: Suppose that there is no value of information control. From Proposition 2 this implies that $\tilde{V}_S(p^S) = V_S(p^S)$. As \tilde{V}_S is the concave closure of an upper semicontinuous

function in a compact set, the differential of $-\tilde{V}_S(q^S)$ is non-empty for all $q^S \in \text{int}(\Delta(\Theta))$. Any element of $\partial(-\tilde{V}_S(p^S))$ would then satisfy (11). \blacksquare

Proof of Corollary 2: *Sufficiency:* Suppose that (15) is satisfied. Then any signal π that, from the sender's point of view, induces the distribution over posterior beliefs σ must satisfy $E_\sigma[q^S] = p^S$, implying that

$$\sum_{\theta \in \Theta} p_\theta^S u_S(a(\mathbf{1}_\theta), \theta) = E_\sigma \left[\sum_{\theta \in \Theta} q_\theta^S u_S(a(\mathbf{1}_\theta), \theta) \right] \geq E_\sigma [V_S(q^S)].$$

Thus, a fully informative signal weakly dominates any other signal π and is thus optimal.

Necessity: Fix any belief $q^S \in \Delta(\Theta)$ and let $\bar{\delta}$ be defined as

$$\bar{\delta} = \max \left\{ \delta : p_\theta^S - \frac{\delta}{1-\delta} (q_\theta^S - p_\theta^S) \geq 0, \delta \in [0, 1] \right\}$$

As the prior belief $p^S \in \text{int}(\Delta(\Theta))$ we have $1 > \bar{\delta} > 0$. Letting $\mathbf{1}_\theta$ be the belief that assigns probability 1 to state θ , consider now a signal that induces belief q^S with probability $\bar{\delta}$ and belief $\mathbf{1}_\theta$ with probability $(1 - \bar{\delta}) \left(p_\theta^S - \frac{\bar{\delta}}{1-\bar{\delta}} (q_\theta^S - p_\theta^S) \right) = p_\theta^S - \bar{\delta} q_\theta^S \geq 0$ for each $\theta \in \Theta$. The expected utility of the sender under this signal is

$$\bar{\delta} V_S(q^S) + \sum_{\theta \in \Theta} (p_\theta^S - \bar{\delta} q_\theta^S) u_S(a(\mathbf{1}_\theta), \theta) = \bar{\delta} \left(V_S(q^S) - \sum_{\theta \in \Theta} q_\theta^S u_S(a(\mathbf{1}_\theta), \theta) \right) + \sum_{\theta \in \Theta} p_\theta^S u_S(a(\mathbf{1}_\theta), \theta).$$

Full disclosure is optimal by assumption, therefore we must have

$$\bar{\delta} \left(V_S(q^S) - \sum_{\theta \in \Theta} q_\theta^S u_S(a(\mathbf{1}_\theta), \theta) \right) + \sum_{\theta \in \Theta} p_\theta^S u_S(a(\mathbf{1}_\theta), \theta) \leq \sum_{\theta \in \Theta} p_\theta^S u_S(a(\mathbf{1}_\theta), \theta),$$

which, given that $\bar{\delta} > 0$, we must then necessarily have (15). \blacksquare

Proof of Proposition 3: *Part (i).* Letting $\hat{u}_S(q^R) = u_S(a(q^R))$, (14) translates to

$$V_R(q^R) = \langle q^R, r^S \rangle \hat{u}_S(q^R),$$

with gradient

$$\nabla V_R(p^R) = \hat{u}_S(p^R) r^S + \nabla \hat{u}_S(p^R).$$

Let W be the subspace of ‘‘marginal beliefs’’, and for $q^R \in \Delta(\Theta)$ let $\delta \in \mathbb{R}$ and $\varepsilon \in W$ be such that $q^R = p^R + \delta \varepsilon$ with ε unitary (i.e. $\langle \varepsilon, \varepsilon \rangle = 1$). Then the condition (12) in Corollary 1 can be expressed as,

$$\langle \nabla V_R(p^R), \delta \varepsilon \rangle \geq V_R(q^R) - V_R(p^R),$$

which can be expanded to

$$\begin{aligned}\hat{u}_S(p^R) \langle r^S, \delta\varepsilon \rangle + \langle \nabla \hat{u}_S(p^R), \delta\varepsilon \rangle &\geq \langle q^R, r^S \rangle \hat{u}_S(q^R) - \hat{u}_S(p^R), \\ \hat{u}_S(p^R) + \langle \nabla \hat{u}_S(p^R), \delta\varepsilon \rangle - \hat{u}_S(q^R) &\geq (\hat{u}_S(q^R) - \hat{u}_S(p^R)) \langle r^S, \delta\varepsilon \rangle.\end{aligned}\quad (28)$$

The left hand side of (28) is the excess of the linear approximation $\hat{u}_S(p^R) + \langle \nabla \hat{u}_S(p^R), \delta\varepsilon \rangle$ over the function $\hat{u}_S(q^R)$, which is positive for concave \hat{u}_S . The mean value theorem in integral form implies

$$\begin{aligned}\hat{u}_S(p^R) + \langle \nabla \hat{u}_S(p^R), \delta\varepsilon \rangle - \hat{u}_S(q^R) &= - \int_0^1 \langle \nabla \hat{u}_S(p^R + t\delta\varepsilon), \delta\varepsilon \rangle dt + \langle \nabla \hat{u}_S(p^R), \delta\varepsilon \rangle \\ &= - \int_0^1 \langle \nabla \hat{u}_S(p^R + t\delta\varepsilon) - \nabla \hat{u}_S(p^R), \delta\varepsilon \rangle dt \\ &= -\delta^2 \varepsilon^T M(\delta\varepsilon) \varepsilon,\end{aligned}$$

where

$$(M(\delta\varepsilon))_{\theta_i, \theta_j} = \int_0^1 \int_0^t \frac{\partial^2 \hat{u}_S(p^R + \tau\delta\varepsilon)}{\partial q_{\theta_i}^R \partial q_{\theta_j}^R} d\tau dt.$$

Therefore (28) translates to

$$\begin{aligned}-\delta^2 \varepsilon^T M(\delta\varepsilon) \varepsilon &\geq \langle r^S, \delta\varepsilon \rangle \left\langle \int_0^1 \nabla \hat{u}_S(p^R + t\delta\varepsilon) dt, \delta\varepsilon \right\rangle, \\ -\varepsilon^T M(\delta\varepsilon) \varepsilon &\geq \langle r^S, \varepsilon \rangle \left\langle \int_0^1 \nabla \hat{u}_S(p^R + t\delta\varepsilon) dt, \varepsilon \right\rangle.\end{aligned}\quad (29)$$

We finish our proof by making three observations. First, a negative definite Hessian implies that the left hand side of (29) is bounded away from zero for $\delta \geq 0$. Let ξ be such a bound, i.e.

$$-\varepsilon^T M(\delta\varepsilon) \varepsilon \geq \xi > 0, \text{ for all } (\delta, \varepsilon) \text{ such that } \langle \varepsilon, \varepsilon \rangle = 1, \delta \in [0, 1].$$

Second, smoothness of \hat{u}_S implies that the term $\left| \left\langle \int_0^1 \nabla \hat{u}_S(p^R + t\delta\varepsilon) dt, \eta \right\rangle \right|$ is uniformly bounded in $\{(\delta, \varepsilon) : \langle \varepsilon, \varepsilon \rangle = 1, \delta \in [0, 1]\}$. Let M be such an upper bound. Third, the prior of the sender only enters (29) through the term $\langle r^S, \varepsilon \rangle$. Clearly, if the sender and the receiver share a common prior then $r^S = \mathbf{1}$ and $\langle r^S, \varepsilon \rangle = 0$. As $\langle r^S, \varepsilon \rangle$ is continuous in p^S and $\langle r^S, \varepsilon \rangle = 0$ when $p^S = p^R$, then there exists a neighborhood of p^R , $N(p^R)$, such that for $p^S \in N(p^R)$ we have $|\langle r^S, \varepsilon \rangle| < \xi/M$. That is, condition (29) is satisfied, implying that the sender does not benefit from information control for every $p^S \in N(p^R)$.

Part (ii). Suppose that for belief $q^R(+)$ we have $\hat{u}_S(p^R) < \hat{u}_S(q^R(+))$. Define the collection of signals $\{\pi(\delta), \delta \in \Xi \subset [0, 1]\}$, such that each signal induces only two posteriors, $q^R(+)$ and $q^R(-)$, with

$$q_\theta^R(-) = p_\theta^R - \frac{\delta}{1-\delta}(q_\theta^R(+)-p_\theta^R),$$

and where the receiver assigns probability δ to $q^R(+)$. Let $\bar{\delta}$ be the maximum admissible δ

$$\bar{\delta} = \max \{ \delta : q^R(-) \in \Delta(\Theta) \}.$$

The full support assumption on p^R implies $\bar{\delta} > 0$. Furthermore, from the definition of $q_\theta^R(-)$, we have

$$1 - \frac{\bar{\delta}}{1-\bar{\delta}} \left(\max_{\theta \in \Theta} \frac{q_\theta^R(+)}{p_\theta^R} - 1 \right) = 0,$$

yielding

$$\bar{\delta} = \frac{1}{\max_{\theta \in \Theta} \frac{q_\theta^R(+)}{p_\theta^R}}.$$

From (5) the probability that a sender with prior p^S assigns to the signal $\pi(\bar{\delta})$ inducing $q^R(+)$ in the receiver is

$$\Pr_S(q^R(+)) = \Pr_R(q^R(+)) \langle q^R(+), r^S \rangle = \bar{\delta} \langle q^R(+), r^S \rangle = \sum_{\theta \in \Theta} p_\theta^S \frac{q_\theta^R(+)/p_\theta^R}{\max_{\theta \in \Theta} q_\theta^R(+)/p_\theta^R}. \quad (30)$$

Let $\Delta(\bar{\delta}, p^S)$ be the sender's expected gain from signal $\pi(\bar{\delta})$, i.e.

$$\Delta(\bar{\delta}, p^S) = \Pr_S(q^R(+)) (\hat{u}_S(q^R(+)) - \hat{u}_S(p^R)) + (1 - \Pr_S(q^R(+))) (\hat{u}_S(q^R(-)) - \hat{u}_S(p^R)).$$

As $\hat{u}_S(q^R)$ is bounded in the simplex $\Delta(\Theta)$, let M be the maximum variation $M = \sup \hat{u}_S(q^R) - \inf \hat{u}_S(q^R)$. Then

$$\Delta(\bar{\delta}, p^S) \geq \Pr_S(q^R(+)) \left[\hat{u}_S(q^R(+)) - \hat{u}_S(p^R) + \frac{(1 - \Pr_S(q^R(+)))}{\Pr_S(q^R(+))} M \right].$$

Let $\varphi = \hat{u}_S(q^R(+)) - \hat{u}_S(p^R) > 0$. As $\Pr_S(q^R(+))$, given by (30), is continuous in p^S , and $\Pr_S(q^R(+))$ converges to 1 as the prior belief p^S tends to $\mathbf{1}_{\theta'}$, where θ' satisfies $q_{\theta'}^R(+)/p_{\theta'}^R = \max_{\theta \in \Theta} q_\theta^R(+)/p_\theta^R$, we can always find p'^S such that

$$\sum_{\theta \in \Theta} p_\theta'^S \frac{q_\theta^R(+)/p_\theta^R}{\max_{\theta \in \Theta} q_\theta^R(+)/p_\theta^R} > \frac{M}{M + \varphi},$$

which implies that $\Delta(\bar{\delta}, p^S) > 0$, i.e. a sender with prior p^S is so confident of inducing the favorable belief $q^R(+)$ with $\pi(\bar{\delta})$ that, regardless of u_S , she benefits from information control. ■

The proofs of Propositions 4 and 5 will make use of the following Lemma.

Lemma A.1 *Let $x, y \in \mathbb{R}^N$, and W defined by (3). Then,*

$$\frac{1}{2} (\|x_{\parallel W}\| \|y_{\parallel W}\| + \langle x_{\parallel W}, y_{\parallel W} \rangle) = \max_{v \in W, \|v\|=1} \langle x, v \rangle \langle y, v \rangle, \quad (31)$$

Proof: For notational convenience, let $\rho(x, y)$ be the angle formed by the vectors x and y , where trivially for any v we have $\rho(x, y) = \rho(x, v) + \rho(v, y)$. If $v \in W$, then $\langle v, x \rangle = \langle v, x_{\parallel W} \rangle$ and $\langle v, y \rangle = \langle v, y_{\parallel W} \rangle$. Therefore, for every $v \in W, \|v\| = 1$, we have

$$\begin{aligned} \langle x, v \rangle \langle y, v \rangle &= \langle v, x_{\parallel W} \rangle \langle v, y_{\parallel W} \rangle = \|x_{\parallel W}\| \|y_{\parallel W}\| \|v\|^2 \cos \rho(v, x_{\parallel W}) \cos \rho(v, y_{\parallel W}) \\ &= \|x_{\parallel W}\| \|y_{\parallel W}\| \frac{\cos(\rho(v, x_{\parallel W}) + \rho(v, y_{\parallel W})) + \cos(\rho(v, x_{\parallel W}) - \rho(v, y_{\parallel W}))}{2} \\ &= \|x_{\parallel W}\| \|y_{\parallel W}\| \frac{\cos(2\rho(v, x_{\parallel W}) + \rho(x_{\parallel W}, y_{\parallel W})) + \cos(\rho(x_{\parallel W}, y_{\parallel W}))}{2}, \end{aligned}$$

which implies

$$\begin{aligned} &\max_{v \in W, \|v\|=1} \langle x, v \rangle \langle y, v \rangle \\ &= \|x_{\parallel W}\| \|y_{\parallel W}\| \left[\frac{\cos(\rho(x_{\parallel W}, y_{\parallel W}))}{2} + \max_{v \in W, \|v\|=1} \frac{\cos(2\rho(v, x_{\parallel W}) + \rho(x_{\parallel W}, y_{\parallel W}))}{2} \right] \\ &= \|x_{\parallel W}\| \|y_{\parallel W}\| \left[\frac{\cos(\rho(x_{\parallel W}, y_{\parallel W}))}{2} + \frac{1}{2} \right], \end{aligned}$$

where the maximum is achieved by selecting a vector v such that $\rho(v, x_{\parallel W}) = -\frac{1}{2}\rho(x_{\parallel W}, y_{\parallel W})$. Rewriting this last expression one obtains (31). ■

Proof of Proposition 4: *Part (i).* We will first show that if $T \cap A^+ \neq \emptyset$, then there exists a signal with two signal realizations such that the sender chooses a strictly higher decision after either realization. This implies that information control is valuable to the sender, for any strictly increasing u_S and totally mixed p^S .

Let $q_0^R \in T \cap A^+$. Since $a(q^R)$ is continuous then A^+ is open and there is a neighborhood of q_0^R with all posterior beliefs leading to strictly higher decisions. In particular, there exists $\bar{\eta} > 0$ such that $q_0^R - \bar{\eta} \nabla a(p^R)_{\parallel W} \in A^+$. Next, define the vector $t(\bar{\eta}) = -(q_0^R - \bar{\eta} \nabla a(p^R)_{\parallel W} - p^R)$.

We now show that there is a belief of the receiver $q^R = p^R + \lambda t(\bar{\eta})$, $\lambda > 0$, that leads to a higher action, i.e. such that $a(p^R + \lambda t(\bar{\eta})) - a(p^R) > 0$.

Since $\nabla a(p^R)_{\|W} \neq 0$, then

$$\begin{aligned} \langle \nabla a(p^R), t(\bar{\eta}) \rangle &= -\langle \nabla a(p^R), q_0^R - p^R \rangle + \bar{\eta} \langle \nabla a(p^R), \nabla a(p^R)_{\|W} \rangle \\ &= \bar{\eta} \langle \nabla a(p^R)_{\|W}, \nabla a(p^R)_{\|W} \rangle > 0. \end{aligned}$$

This implies that the derivative of the function $a(\lambda) = a(p^R + \lambda t(\bar{\eta}))$ is strictly positive at $\lambda = 0$. Therefore, there exists $\lambda^* > 0$ such that

$$a(p^R + \lambda^* t(\bar{\eta})) - a(p^R) > 0.$$

Consider now a signal with two signal realizations that induce posterior beliefs in the receiver $q_0^R - \bar{\eta} \nabla a(p^R)_{\|W}$ and $p^R + \lambda^* t(-\bar{\eta})$. By construction, these two posterior beliefs lie on the same line that contains the prior belief p^R and thus can be induced by a signal. Importantly, both posteriors lead to a higher decision for the receiver, and thus information control is valuable for the sender.

To prove the second part, note that if the restriction of the Hessian matrix $H(a(p^R))$ to the tangent hyperplane has a positive eigenvalue then the intersection $T \cap A^+$ is non-empty.

Part (ii). Under pure persuasion, the representation (14) translates to

$$V_R(q^R) = u_S(a(q^R)) \langle q^R, r^S \rangle. \quad (32)$$

Our proof strategy is to show that, whenever (17) holds, one can find a direction in the space W of "marginal beliefs" along which V_R is locally strictly convex at $q^R = p^R$. This implies that $\tilde{V}_R(p^R) > V_R(p^R)$, and thus information control is valuable.

Consider a vector $v \in W$, and the function $V(\lambda; v) = V_R(p^R + \lambda v)$. Twice differentiating (32) and evaluating it at $\lambda = 0$ we can write

$$\begin{aligned} \left. \frac{\partial^2 V(\lambda; v)}{\partial \lambda^2} \right|_{\lambda=0} &= u_S''(a(p^R)) \langle \nabla a(p^R), v \rangle^2 + u_S'(a(p^R)) v^T H(a(p^R)) v \\ &\quad + 2u_S'(a(p^R)) \langle \nabla a(p^R), v \rangle \langle v, r^S \rangle. \end{aligned} \quad (33)$$

Next, let Ψ denote the function

$$\Psi(v) = \frac{u_S''(a(p^R))}{u_S'(a(p^R))} \langle \nabla a(p^R), v \rangle^2 + 2 \langle \nabla a(p^R), v \rangle \langle v, r^S \rangle.$$

It can be readily seen that Ψ is a quadratic form satisfying the functional form of Lemma A.1. Letting $m = \nabla a(p^R)$ and $n = \frac{u''_S(a(p^R))}{u'_S(a(p^R))} \nabla a(p^R) + 2r^S$, Lemma A.1 then implies that

$$\frac{1}{2} \|m_{\|W}\| \|n_{\|W}\| (1 + \cos(m_{\|W}, n_{\|W})) = \max \Psi(v), s.t., v \in W, \|v\| = 1.$$

Furthermore, if λ_{\min} is the smallest eigenvalue of $H(a(p^R))$, one has that

$$\lambda_{\min} \|v\|^2 \leq v^T H(a(p^R))v.$$

We can now establish the existence of a vector v such that $V(\lambda) = V_R(p^R + \lambda v)$ is locally stricly convex. Indeed, taking into account the definition of $\Psi(v)$, we can rewrite (33) as

$$\left. \frac{d^2 V}{d\lambda^2} \right|_{\lambda=0} = u'_S(a(p^R)) (\Psi(v) + v^T H(a(p^R))v) \geq u'_S(a(p^R)) (\Psi(v) + \lambda_{\min} \|v\|^2).$$

Condition (17) guarantees that $\max \Psi(v) > |\lambda_{\min}|$ and thus the existence of a vector $v^* \in W$, $\|v^*\| = 1$, such that $\Psi(v^*) = \max \Psi(v)$ and

$$\left. \frac{\partial^2 V(\lambda; v^*)}{\partial \lambda^2} \right|_{\lambda=0} \geq u'_S(a(p^R)) (\Psi(v^*) + \lambda_{\min} \|v^*\|^2) > 0,$$

thus $V_R(p^R)$ is locally strictly convex in the direction v^* . \blacksquare

Proof of Lemma 1: Let $\varepsilon = q^R - p^R \in W$ with $q^R \in \Delta(\Theta)$. Posterior belief q^R does not decrease the receiver's action if and only if $\langle \varepsilon, x \rangle \geq 0$, while (5) implies that for any signal, the sender does not assign more probability to the receiver having a belief q^R if and only if $\langle \varepsilon, r^S \rangle \leq 0$. Therefore, for any signal and every signal realization, the sender never assigns more probability to any belief that (weakly) increases the receiver's action if and only if

$$\langle \varepsilon, x \rangle \langle \varepsilon, r^S \rangle \leq 0, \varepsilon = q^R - p^R, q^R \in \Delta(\Theta).$$

Since the set $\{\varepsilon : \varepsilon = q^R - p^R, q^R \in \Delta(\Theta)\} \subset W$ contains a neighborhood of 0 in W , then the previous condition is satisfied if and only if the following global condition is true:

$$\langle \varepsilon, x \rangle \langle \varepsilon, r^S \rangle \leq 0 \text{ for } \varepsilon \in W,$$

or, in other words, iff the quadratic form $\langle \varepsilon, x \rangle \langle \varepsilon, r^S \rangle$ is negative semidefinite in W .

Consider the orthogonal decompositions $x = x_{\|W} + \alpha_x 1$ and $r^S = r^S_{\|W} + \alpha_r 1$. Whenever $\varepsilon \in W$ we have $\langle \varepsilon, x \rangle = \langle \varepsilon, x_{\|W} \rangle$ and $\langle \varepsilon, r^S \rangle = \langle \varepsilon, r^S_{\|W} \rangle$, implying that negative semidefiniteness of $\langle \varepsilon, x \rangle \langle \varepsilon, r^S \rangle$ in W is equivalent to negative semidefiniteness of $\langle \varepsilon, x_{\|W} \rangle \langle \varepsilon, r^S_{\|W} \rangle$

in W . From Lemma A.1 we have

$$0 = \max_{\varepsilon \in W, \|\varepsilon\|=1} \langle \varepsilon, x_{\parallel W} \rangle \langle \varepsilon, r_{\parallel W}^S \rangle \Leftrightarrow \langle x_{\parallel W}, r_{\parallel W}^S \rangle = -\|x_{\parallel W}\| \|r_{\parallel W}^S\|,$$

If $x_{\parallel W} \neq 0$ and $r_{\parallel W}^S \neq 0$, then $\langle x_{\parallel W}, r_{\parallel W}^S \rangle = -\|x_{\parallel W}\| \|r_{\parallel W}^S\|$ iff $\cos(x_{\parallel W}, r_{\parallel W}^S) = -1$ which is equivalent to the existence of $\alpha > 0$ such that $x_{\parallel W} = -\alpha r_{\parallel W}^S$. ■

Lemma A.2 *Suppose that $N = \text{card}(\Theta) \geq 3$ and consider the subspace $W = \{\varepsilon \in \mathbb{R}^N : \langle \varepsilon, 1 \rangle = 0\}$ with the derived topology. Then, for $v \notin W$, the rational function $\langle \varepsilon, w \rangle / \langle \varepsilon, v \rangle$, $\varepsilon \in W$, is bounded in a neighborhood of $\mathbf{0}$ if and only if $v_{\parallel W}$ and $w_{\parallel W}$ are collinear.*

Proof: Consider the linear subspace $W_{v,1} = \{\varepsilon \in \mathbb{R}^N : \langle \varepsilon, v \rangle = 0, \langle \varepsilon, 1 \rangle = 0\}$. As, by assumption $v \notin W$, then $W_{v,1}$ is a linear subspace of dimension $N - 2 \geq 1$. Consider now the subspace $W_w = \{\varepsilon \in \mathbb{R}^N : \langle \varepsilon, w \rangle = 0\}$. The ratio $\langle \varepsilon, w \rangle / \langle \varepsilon, v \rangle$ is locally unbounded in W iff $W_{v,1} \cap W_w^c \neq \emptyset$. First, if the projections $v_{\parallel W}$ and $w_{\parallel W}$ are not collinear then the orthogonal projection $w_{\parallel W_{v,1}}$ is non-zero, implying that $\langle w_{\parallel W_{v,1}}, v \rangle = 0$ but $\langle w_{\parallel W_{v,1}}, w \rangle > 0$. This establishes that $W_{v,1} \cap W_w^c \neq \emptyset$. Now suppose that $v_{\parallel W} = \lambda w_{\parallel W}$ for some $\lambda \neq 0$. Then $\langle \varepsilon, v_{\parallel W} \rangle = 0$ iff $\langle \varepsilon, w_{\parallel W} \rangle = 0$, implying $W_{v,1} \cap W_w^c = \emptyset$. ■

Proof of Proposition 5: The representation (14) applied to our setup yields

$$V_R(q^R) = u_S(\langle q^R, x \rangle) \langle q^R, r^S \rangle,$$

with gradient at the prior belief p^R

$$\nabla V_R(p^R) = u'_S(\langle p^R, x \rangle) \mathbf{x} + u_S(\langle p^R, x \rangle) \mathbf{r}^S.$$

Corollary 1 implies that the value of information control is zero if and only if

$$\langle \nabla V_R(p^R), q^R - p^R \rangle \geq V^R(q^R) - V^R(p^R), \quad q^R \in \Delta(\Theta),$$

which in our case leads to

$$u'_S(\langle p^R, x \rangle) \langle x, q^R - p^R \rangle - \langle q^R, r^S \rangle (u_S(\langle q^R, x \rangle) - u_S(\langle p^R, x \rangle)) \geq 0, \quad q^R \in \Delta(\Theta). \quad (34)$$

To ease notation, let $\varepsilon = q^R - p^R \in W$ and define Δ as the left hand side of (34),

$$\Delta = u'_S(\langle p^R, x \rangle) \langle x, \varepsilon \rangle - \langle q^R, r^S \rangle (u_S(\langle q^R, x \rangle) - u_S(\langle p^R, x \rangle)). \quad (35)$$

Part (i) - To show that for an arbitrary smooth and increasing u_S the value of information control is positive whenever $x_{\|W}$ and $r_{\|W}^S$ are not negatively collinear, it suffices to find a feasible q^R such that $\Delta < 0$. First, with the help of the identities

$$\begin{aligned}
u_S(\langle q^R, x \rangle) - u_S(\langle p^R, x \rangle) &= \int_{\langle p^R, x \rangle}^{\langle q^R, x \rangle} u'_S(t) dt, \\
u_S(\langle q^R, x \rangle) - u_S(\langle p^R, x \rangle) - \langle x, \varepsilon \rangle u'_S(\langle p^R, x \rangle) &= \int_{\langle p^R, x \rangle}^{\langle q^R, x \rangle} u'_S(t) dt - \int_{\langle p^R, x \rangle}^{\langle q^R, x \rangle} u'_S(\langle p^R, x \rangle) dt, \\
&= \int_{\langle p^R, x \rangle}^{\langle q^R, x \rangle} (u'_S(t) - u'_S(\langle p^R, x \rangle)) dt, \\
&= \int_{\langle p^R, x \rangle}^{\langle q^R, x \rangle} \int_{\langle p^R, x \rangle}^t u''_S(\tau) d\tau dt,
\end{aligned}$$

and $\langle p^R, r^S \rangle = 1$, we can rewrite Δ in (35) as

$$\Delta = - \int_{\langle p^R, x \rangle}^{\langle q^R, x \rangle} \int_{\langle p^R, x \rangle}^t u''_S(\tau) d\tau dt - \langle \varepsilon, r^S \rangle \int_{\langle p^R, x \rangle}^{\langle q^R, x \rangle} u'_S(t) dt.$$

Given q^R , The smoothness condition on $u'_S(a)$ implies that $u'_S(a)$ and $u''_S(a)$ are bounded in the compact set $A = \{a : a = \langle q^R, z \rangle, q^R \in \Delta(\Theta)\}$. Thus, let $m_S = \min \{u'_S(a) : a \in A\}$ and $M_S = \max |u''_S(a)| : a \in A$, with $m_S > 0$ since $u'_S(a) > 0$. Then

$$\begin{aligned}
\Delta &\leq M_S \int_{\langle p^R, x \rangle}^{\langle q^R, x \rangle} \int_{\langle p^R, x \rangle}^t d\tau dt - \langle \varepsilon, r^S \rangle m_S \int_{\langle p^R, x \rangle}^{\langle q^R, x \rangle} dt \\
&= \frac{1}{2} \langle \varepsilon, x \rangle^2 M_S - \langle \varepsilon, r^S \rangle \langle \varepsilon, x \rangle m_S \\
&= m_S \langle \varepsilon, x \rangle^2 \left(\frac{1}{2} \frac{M_S}{m_S} - \frac{\langle \varepsilon, r^S \rangle}{\langle \varepsilon, x \rangle} \right).
\end{aligned}$$

From Lemma A.1, if $x_{\|W}$ and $r_{\|W}^S$ are not negatively collinear then there exists a neighborhood $N(\mathbf{0})$ of $\mathbf{0}$ in W such that $\langle \varepsilon, r^S \rangle / \langle \varepsilon, x \rangle$ admits no upper bound in $N(\mathbf{0})$. This establishes the existence of $\varepsilon \in N(\mathbf{0})$, and thus a feasible $q^R = p^R + \varepsilon$, such that

$$\frac{1}{2} \frac{M_S}{m_S} - \frac{\langle \varepsilon, r^S \rangle}{\langle \varepsilon, x \rangle} < 0,$$

implying that $\Delta < 0$.

Part (ii)- We show that if, in addition, u_S is concave then the condition on x and r^S is also necessary for the sender to benefit from information control. Our proof strategy is

to establish the contrapositive: if $x_{\parallel W}$ and $r_{\parallel W}^S$ are negatively collinear then the value of information control is zero.

Concavity of u_S yields the following bound

$$u_S(\langle q^R, x \rangle) - u_S(\langle p^R, x \rangle) \leq u'_S(\langle p^R, x \rangle) \langle \varepsilon, x \rangle,$$

which, applied to (35), implies

$$\Delta \geq -u'_S(\langle p^R, x \rangle) \langle \varepsilon, x \rangle \langle \varepsilon, r^S \rangle. \quad (36)$$

As $x_{\parallel W}$ and $r_{\parallel W}^S$ are negatively collinear, Lemma 1 implies that

$$\langle \varepsilon, x \rangle \langle \varepsilon, r^S \rangle \leq 0 \text{ for } \varepsilon \in W,$$

which applied to (36) leads to

$$\Delta \geq -u'_S(\langle p^R, x \rangle) \langle \varepsilon, x \rangle \langle \varepsilon, r^S \rangle \geq 0 \text{ for } \varepsilon \in W, .$$

Since $\Delta \geq 0$ for all feasible beliefs, Corollary 1 implies that the value of information control is zero. ■

Proof of Proposition 6: *Part (i)* - First, likelihood ratio orders are preserved by bayesian updating with commonly understood signals (Whitt 1979, Milgrom 1981). Thus, induced posteriors satisfy $q^S(z) \succeq_{LR} q^R(z)$ if $p^S \succeq_{LR} p^R$ for any signal π and signal realization z , and, as x is increasing in θ , we must then have $\langle q^S(z), x \rangle \geq \langle q^R(z), x \rangle$. Therefore

$$q_\theta^S u_S(\langle \mathbf{1}_\theta, x \rangle) \geq u_S(\langle q^S, x \rangle) \geq u_S(\langle q^R, x \rangle) = V_S(q^S), \quad q^S \in \Delta(\Theta),$$

where the first inequality follows from convexity of u_S . Corollary 2 then implies that garbling is not valuable.

Part (ii) - Consider two states θ and θ' and the indexed family of receiver's posterior beliefs $q^R(\delta)$ and associated sender's beliefs $q^S(\delta)$ given by

$$\begin{aligned} q^R(\delta) &= \delta \mathbf{1}_{\theta'} + (1 - \delta) \mathbf{1}_\theta, \delta \in [0, 1], \\ q^S(\delta) &= \lambda(\delta) \mathbf{1}_{\theta'} + (1 - \lambda(\delta)) \mathbf{1}_\theta, \text{ with } \lambda(\delta) = \delta r_{\theta'}^S / (\delta r_{\theta'}^S + (1 - \delta) r_\theta^S). \end{aligned}$$

Define $W(\delta, \theta, \theta')$ as

$$\begin{aligned} W(\delta, \theta, \theta') &= \lambda(\delta) u_S(x(\theta')) + (1 - \lambda(\delta)) u_S(x(\theta')) \\ &\quad - u_S(\delta x(\theta') + (1 - \delta) x(\theta')). \end{aligned}$$

From Corollary 2, if for some $(\delta, \theta, \theta')$ we have $W(\delta, \theta, \theta') < 0$, then the value of garbling is positive. After some algebraic manipulations we can express $W(\delta, \theta, \theta')$ as

$$W(\delta, \theta, \theta') = \frac{\delta(1-\delta)}{(\delta r_{\theta'}^S + (1-\delta)r_{\theta}^S)} S(\delta, \theta, \theta'),$$

with

$$S(\delta, \theta, \theta') = r_{\theta'}^S \frac{1}{(1-\delta)} \int_{\delta x(\theta') + (1-\delta)x(\theta)}^{x(\theta')} u'_S(t) dt - r_{\theta}^S \frac{1}{\delta} \int_{x(\theta)}^{\delta x(\theta') + (1-\delta)x(\theta)} u'_S(t) dt,$$

where we have exploited the absolute continuity of u_S to express it as the integral of its derivative. Evaluating $S(\delta, \theta, \theta')$ at the extremes we obtain

$$S(0, \theta, \theta') = (x(\theta') - x(\theta)) (r_{\theta'}^S \bar{u}'_S - r_{\theta}^S u'_S(x(\theta))), \quad (37)$$

$$S(1, \theta, \theta') = (x(\theta') - x(\theta)) (r_{\theta'}^S u'_S(x(\theta')) - r_{\theta}^S \bar{u}'_S), \quad (38)$$

with

$$\bar{u}'_S = \frac{1}{(x(\theta') - x(\theta))} \int_{x(\theta)}^{x(\theta')} u'_S(t) dt.$$

By assumption, there exist θ' and θ , $\theta' > \theta$, such that $(r_{\theta'}^S)^2 u'_S(x(\theta')) < (r_{\theta}^S)^2 u'_S(x(\theta))$. This implies that $\frac{r_{\theta'}^S}{r_{\theta}^S} u'_S(x(\theta')) < \frac{r_{\theta}^S}{r_{\theta'}^S} u'_S(x(\theta))$, which implies that either $S(0, \theta, \theta')$ or $S(1, \theta, \theta')$ is strictly negative. To see this, suppose for example that $S(0, \theta, \theta') \geq 0$. Then

$$\frac{r_{\theta'}^S}{r_{\theta}^S} u'_S(x(\theta')) - \bar{u}'_S < \frac{r_{\theta}^S}{r_{\theta'}^S} u'_S(x(\theta)) - \bar{u}'_S = -\frac{S(0, \theta, \theta')}{(x(\theta') - x(\theta)) r_{\theta'}^S} \leq 0 \Rightarrow S(1, \theta, \theta') < 0. \blacksquare$$

Proof of Corollary 3: The claim follows immediately by applying Proposition 5(i) to the cases where $x = \theta$ and $x = -\theta$. \blacksquare

Proof of Proposition 7: In the text.

Proof of Proposition 8: See Appendix B.

References

- [1] ACEMOGLU, D., V. CHERNOZHUKOV, AND M. YILDIZ (2006): “Learning and Disagreement in an Uncertain World,” *NBER Working Paper No. 12648*.
- [2] ALONSO, R. AND O. CÂMARA. (2014): “Persuading Voters,” *mimeo*.
- [3] AUMANN, R. J. (1976): “Agreeing to Disagree,” *The Annals of Statistics*, 4(6), 1236-1239.

- [4] BLACKWELL, D., AND L. DUBINS. (1962): “Merging of Opinions with Increasing Information,” *The Annals of Mathematical Statistics*, 33(3), 882-886.
- [5] BROCAS, I., AND J. CARRILLO (2007): “Influence through Ignorance,” *Rand Journal of Economics*, 38(4), 931-947.
- [6] CALLANDER, S. (2011): “Searching for Good Policies,” *American Political Science Review*, 105(4), 643-662.
- [7] CHE, Y.-K., AND N. KARTIK (2009): “Opinions as Incentives,” *Journal of Political Economy*, 117(5), 815-860.
- [8] CSISZAR, I. (1967): “Information-type Measures of Probability Distributions and Indirect Observations,” *Studia Math. Hungarica*, 2, 299-318.
- [9] DUGGAN, J., AND C. MARTINELLI (2011): “A Spatial Theory of Media Slant and Voter Choice,” *Review of Economic Studies*, 78(2), 640-666.
- [10] ELLISON, G., AND S. F. ELLISON (2009): “Search, Obfuscation, and Price Elasticities on the Internet,” *Econometrica*, 77, 427-452.
- [11] GENTZKOW, M. AND E. KAMENICA (2013): “Costly Persuasion,” *mimeo*, University of Chicago.
- [12] GENTZKOW, M. AND J. M. SHAPIRO (2006): “Media Bias and Reputation,” *Journal of Political Economy*, 114(2), 280-316.
- [13] GIAT, Y., HACKMAN, S. AND A. SUBRAMANIAN (2010): “Investment under Uncertainty, Heterogeneous Beliefs, and Agency Conflicts,” *Review of Financial Studies*, 23(4), 1360-1404.
- [14] GILL, D., AND D. SGROI (2008): “Sequential Decisions with Tests,” *Games and Economic Behavior*, 63(2), 663-678.
- [15] GILL, D., AND D. SGROI (2012): “The Optimal Choice of Pre-Launch Reviewer,” *Journal of Economic Theory*, 147(3), 1247-1260.
- [16] HOLMSTRÖM, B. (1999): “Managerial Incentive Problems: A Dynamic Perspective,” *Review of Economic Studies*, 66, 169-182.
- [17] IVANOV, M. (2010): “Informational Control and Organizational Design,” *Journal of Economic Theory*, 145(2), 721-751.
- [18] KALAI, E., AND E. LEHRER (1994): “Weak and Strong Merging of Opinions.” *Journal of Mathematical Economics*, 23(1), 73-86.

- [19] KAMENICA, E., AND M. GENTZKOW (2011): “Bayesian Persuasion,” *American Economic Review*, 101, 2590-2615.
- [20] MILGROM, P. (1981): “Good News and Bad News: Representation Theorems and Applications,” *Bell Journal of Economics*, 12(2), 380-91.
- [21] MORRIS, S. (1994): “Trade with Heterogeneous Prior Beliefs and Asymmetric Information,” *Econometrica*, 62(6), 1327-1347.
- [22] MORRIS, S. (1995): “The Common Prior Assumption in Economic Theory,” *Economics and Philosophy*, 11, 227-227.
- [23] PATTON, A. J. AND A. TIMMERMANN (2010): “Why do Forecasters Disagree? Lessons from the Term Structure of Cross-sectional Dispersion,” *Journal of Monetary Economics*, 57(7), 803-820.
- [24] RAYO, L., AND I. SEGAL (2010): “Optimal Information Disclosure,” *Journal of Political Economy*, 118(5), 949-987.
- [25] SETHI, R., AND M. YILDIZ (2012): “Public Disagreement,” *American Economic Journal: Microeconomics*, 4(3), 57-95.
- [26] SHAKED, M., AND J. G. SHANTHIKUMAR (2007): *Stochastic Orders*, Springer.
- [27] VAN DEN STEEN , E. (2004): “Rational Overoptimism (and Other Biases),” *American Economic Review*, 94(4), 1141-1151.
- [28] VAN DEN STEEN , E. (2009): “Authority versus Persuasion,” *American Economic Review: P&P*, 99(2), 448-453.
- [29] VAN DEN STEEN , E. (2010a): “Interpersonal Authority in a Theory of the Firm,” *American Economic Review*, 100(1), 466-490.
- [30] VAN DEN STEEN , E. (2010b): “On the Origin of Shared Beliefs (and Corporate Culture),” *Rand Journal of Economics*, 41(4), 617-648.
- [31] VAN DEN STEEN , E. (2011): “Overconfidence by Bayesian-Rational Agents,” *Management Science*, 57(5), 884-896.
- [32] WHITT, W. (1979): “A Note on the Influence of the Sample on the Posterior Distribution,” *Journal of the America Statistical Association*, 74(366a), 424-426.

B On-line Supplemental Material

Lemma B.1 *Let R be defined by (26) and $m = \frac{1}{2} \frac{\max |u_S''(a)|}{\min u_S'(a)} > 0$, and for each $r^R \in R$ define $\Delta_S = \frac{\langle q^S, r^R x \rangle}{\langle q^S, r^R \rangle} - \langle p^R, x \rangle$, and define $l_{r^R}(\varepsilon)$ as*

$$l_{r^R}(\varepsilon) = \frac{\langle \varepsilon, r^R \rangle}{\Delta_S}. \quad (39)$$

For any ε and $r^R \in R$ such that

$$l_{r^R}(\varepsilon) < -m \text{ and } \Delta_S > 0, \text{ with } p^S + \varepsilon \in \Delta(\Theta), \quad (40)$$

there exists a signal π with the following properties: (i) Some realization of π induces in the sender the belief $p^S + \varepsilon$ and (ii) π increases the expected utility of the sender when the receiver's associated likelihood ratio is r^R .

Proof: The function $l_{r^R}(\varepsilon)$ has an immediate interpretation as a measure of disagreement: the numerator $\langle \varepsilon, r^R \rangle$ is the difference in the probability that the receiver and sender attach to a signal realization inducing a posterior $q_S = p_S + \varepsilon$ on the sender, divided by the probability that the sender ascribes to such signal realization, while the denominator is the change in the receiver's action when the sender changes her belief to q_S . We first show that if some ε satisfies (40), then the value of information control is positive. Consider V_S defined in (9), which in this case can be written as

$$V_S(q^S) = u_S \left(\frac{\langle q^S, r^R x \rangle}{\langle q^S, r^R \rangle} \right),$$

with gradient at p^S

$$\nabla V_S(p^S) = u_S'(\langle p^R, x \rangle) (r^R x - \langle p^R, x \rangle r^R).$$

By Corollary 1, the value of information control is positive if and only if there exists ε , with $p^S + \varepsilon \in \Delta(\Theta)$, such that

$$\langle \nabla V_S(p^S), \varepsilon \rangle < V_S(p^S + \varepsilon) - V_S(p^S). \quad (41)$$

We now show that an ε satisfying (40) also satisfies (41). Since

$$u_S \left(\frac{\langle q^S, r^R x \rangle}{\langle q^S, r^R \rangle} \right) - u_S(\langle p^R, x \rangle) - u_S'(\langle p^R, x \rangle) \left(\frac{\langle q^S, r^R x \rangle}{\langle q^S, r^R \rangle} - \langle p^R, x \rangle \right) = \int_{\langle p^R, x \rangle}^{\frac{\langle q^S, r^R x \rangle}{\langle q^S, r^R \rangle}} \int_{\langle p^R, x \rangle}^t u_S''(\tau) d\tau dt,$$

we can rewrite (41) as

$$u'_S(\langle p^R, x \rangle) \langle \varepsilon, r^R \rangle \Delta_S < \int_{\langle p^R, x \rangle}^{\frac{\langle q^S, r^R x \rangle}{\langle q^S, r^R \rangle}} \int_{\langle p^R, x \rangle}^t u''_S(\tau) d\tau dt.$$

By the mean value theorem, we have

$$\int_{\langle p^R, x \rangle}^{\frac{\langle q^S, r^R x \rangle}{\langle q^S, r^R \rangle}} \int_{\langle p^R, x \rangle}^t u''_S(\tau) d\tau dt \geq -\max |u''_S(a)| \int_{\langle p^R, x \rangle}^{\frac{\langle q^S, r^R x \rangle}{\langle q^S, r^R \rangle}} \int_{\langle p^R, x \rangle}^t d\tau dt = -\frac{1}{2} \max |u''_S(a)| \Delta_S^2.$$

Moreover, if ε satisfies (40) then it also satisfies

$$\langle \varepsilon, r^R \rangle \min u'_S(a) < -\frac{1}{2} \max |u''_S(a)| \Delta_S,$$

implying that ε also satisfies (41) since

$$u'_S(\langle p^R, x \rangle) \langle \varepsilon, r^R \rangle \Delta_S < \langle \varepsilon, r^R \rangle \Delta_S \min u'_S(a) < -\frac{1}{2} \max |u''_S(a)| \Delta_S^2 \leq \int_{\langle p^R, x \rangle}^{\frac{\langle q^S, r^R x \rangle}{\langle q^S, r^R \rangle}} \int_{\langle p^R, x \rangle}^t u''_S(\tau) d\tau dt.$$

For each ε satisfying (40), we now construct a signal that improves the sender's expected utility and that has a realization that induces belief $p^S + \varepsilon$ in the sender. Let v be the excess of the right hand side over the left hand side in (41),

$$v = V_S(p^S + \varepsilon) - V_S(p^S) - \langle \nabla V_S(p^S), \varepsilon \rangle > 0. \quad (42)$$

Consider the signal $\pi(\varepsilon, \delta)$ with $Z = \{\varepsilon^+, \varepsilon^-\}$, such that $\Pr_S[z = \varepsilon^+] = \delta$ and if $z = \varepsilon^+$ then the sender's posterior is $p^S + \varepsilon$. A Taylor series expansion of $V_S(q^S)$ yields

$$V_S(q^S) = V_S(p^S) + \langle \nabla V_S(p^S), q^S - p^S \rangle + L (q^S - p^S), \text{ with } \lim_{t \rightarrow 0} \frac{L(t(q^S - p^S))}{t} = 0. \quad (43)$$

Then the sender's gain from signal $\pi(\varepsilon, \delta)$ is

$$\begin{aligned} \Delta_{\pi(\varepsilon, \delta)} &= \delta (V_S(p^S + \varepsilon) - V_S(p^S)) + (1 - \delta) \left(V_S(p^S - \frac{\delta}{1 - \delta} \varepsilon) - V_S(p^S) \right) \\ &= \delta (v + \langle \nabla V_S(p^S), \varepsilon \rangle) - \delta \langle \nabla V_S(p^S), \varepsilon \rangle + L \left(-\frac{\delta}{1 - \delta} \varepsilon \right) \\ &= \delta \left(v - (1 - \delta) \frac{L(-\delta \varepsilon / (1 - \delta))}{(-\delta / (1 - \delta))} \right). \end{aligned}$$

The convergence to zero of the second term in the parenthesis when δ tends to zero and $v > 0$ guarantees the existence of $\delta > 0$ such that $\Delta_{\pi(\varepsilon, \delta)} > 0$. \blacksquare

Proof of Proposition 8: First, we introduce additional notation. With $l_{r^R}(\varepsilon)$ defined as in (39), define the sets $M(r^R)$ by

$$M(r^R) = \{\varepsilon : l_{r^R}(\varepsilon) < -m, \Delta_S > 0, p^S + \varepsilon \in \Delta(\Theta)\}.$$

Note that r^S and x are negatively collinear if and only if r^R and $r^R x$ are positively collinear. That is, the condition on Proposition 5 could be instead stated in terms of collinearity of r^R and $r^R x$. Moreover, if r^R and $r^R x$ are not collinear then the restriction of $l_{r^R}(\varepsilon)$ to $\{\varepsilon : \langle \varepsilon, 1 \rangle = 0\}$ is surjective and thus the set $M(r^R)$ is non-empty.

Define the function

$$\Psi(\varepsilon, r^R) = \langle \varepsilon, r^R - m f^R \rangle + (\langle \varepsilon, r^R \rangle)^2, \text{ with } f^R = r^R x - \langle p^S, r^R x \rangle,$$

which characterizes $M(r^R)$ since for ε such that $p^S + \varepsilon \in \Delta(\Theta)$, $\Psi(\varepsilon, r^R) \leq 0$ and $\langle \varepsilon, f^R \rangle \geq 0$ if and only if $\varepsilon \in M(r^R)$. Finally, let

$$\gamma = 2 \left(1 + m (\max |x_\theta| + \|x\|) + (4 + m \|x\|) \sup_{r^R \in R} \|r^R\| \right), \quad (44)$$

$$Z = \min_{\varepsilon \in \{p^S + \varepsilon \in \Delta(\Theta)\}, r^R \in R} \Psi(\varepsilon, r^R) \text{ s.t. } \langle \varepsilon, r^R (x - \langle p^S, r^R x \rangle) \rangle \leq 0, r^R \in R. \quad (45)$$

Under the conditions of Proposition 8, $Z < 0$. Finally, define β in (27) as

$$\beta = \frac{|Z|}{\gamma}. \quad (46)$$

Our proof is structured in two steps that show (i) if $\cap_{r^R \in R} M(r^R)$ is non-empty then following Lemma B.1 allows us to design a signal π that increases the sender's expected utility for every receiver's belief in the support of $h(p^R|p^S)$, and (ii) under the conditions of Proposition 8, $\cap_{r^R \in R} M(r^R) \neq \emptyset$.

Step (i) - Suppose that $\varepsilon \in \cap_{r^R \in R} M(r^R)$. Consider v as defined by (42). As v is a continuous function of r^R in the compact set R , it achieves a minimum $\underline{v} = \min_{r^R \in R} v > 0$. Then, define $\underline{\delta}$ as

$$\underline{\delta} = \min \left\{ \delta : \underline{v} + \frac{L\left(-\frac{\delta}{1-\delta}\varepsilon\right)}{\delta} \geq 0 \right\},$$

with the function L given by (43). Now define the signal $\pi(\varepsilon, \delta')$ as in the proof of Lemma B.1, i.e. $Z = \{\varepsilon^+, \varepsilon^-\}$, $q^S(\varepsilon^+) = p^S + \varepsilon$ and $\Pr_S[z = \varepsilon^+] = \delta'$, and set $\delta' = \underline{\delta}$. Then the sender's gain from $\pi(\varepsilon, \delta')$ is positive for any receiver's prior in $Supp(h(p^R|p^S))$.

Step (ii) - Fix $p^{R'}$ with associated likelihood ratio $r^{R'} \in R$. For any $r^R \in R$ with $\eta = r^R - r^{R'}$, we have

$$\Psi(\varepsilon, r^R) - \Psi(\varepsilon, r^{R'}) = \left(1 + m \langle p^S, r^{R'} x \rangle + \langle \varepsilon, r^R + r^{R'} \rangle\right) \langle \varepsilon, \eta \rangle - m \langle \varepsilon, \eta x \rangle + m \langle p^S, \eta x \rangle \langle \varepsilon, r \rangle.$$

The following bounds make use of the Cauchy-Schwartz inequality (in particular the implication that $|\langle \varepsilon, \eta x \rangle| \leq \|\varepsilon\| \|\eta\| \|x\|$, see Steele 2004²¹) and the fact that $\|p^S\| \leq 1$ and $\|\varepsilon\| = \|q^S - p^S\| \leq 2$,

$$\begin{aligned} \left|1 + m \langle p^S, r^{R'} x \rangle + \langle \varepsilon, r^R + r^{R'} \rangle\right| &\leq 1 + m \max x_\theta + 4 \sup_{r^R \in R} \|r^R\|, \\ |m \langle \varepsilon, \eta x \rangle| &\leq m \|\varepsilon\| \|\eta\| \|x\| \leq 2m \|\eta\| \|x\|, \\ |m \langle p^S, \eta x \rangle \langle \varepsilon, r \rangle| &\leq 2m \|\eta\| \|x\| \sup_{r^R \in R} \|r^R\|. \end{aligned}$$

From these bounds, we then obtain the following estimate

$$\begin{aligned} \left|\Psi(\varepsilon, r^R) - \Psi(\varepsilon, r^{R'})\right| &\leq \left|1 + m \langle p^S, r^{R'} x \rangle + \langle \varepsilon, r^R + r^{R'} \rangle\right| \|\varepsilon\| \|\eta\| \\ &\quad + |m \langle \varepsilon, \eta x \rangle| + |m \langle p^S, \eta x \rangle \langle \varepsilon, r \rangle| \\ &\leq 2 \left(1 + m \max x_\theta + 4 \sup_{r^R \in R} \|r^R\|\right) \|\eta\| + 2m \|x\| \|\eta\| \\ &\quad + 2m \|x\| \sup_{r^R \in R} \|r^R\| \|\eta\| \\ &= \gamma \|\eta\|, \end{aligned}$$

where γ is defined by (44). Selecting ε' an $r^{R'}$ that solve the program (45) and noting that $Z < 0$ we then have that for any $r^R \in R$,

$$\Psi(\varepsilon', r^R) = \Psi(\varepsilon', r^{R'}) + \Psi(\varepsilon', r^R) - \Psi(\varepsilon', r^{R'}) \leq Z + \gamma \|\eta\| \leq Z + |Z| = 0.$$

This implies that $\varepsilon' \in M(r^R)$ for all $r^R \in R$. ■

²¹Steele, J. M. (2004) "The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities," *Mathematical Association of America*.