Bayesian Comparative Statics*

Teddy Mekonnen † René Leal Vizcaíno ‡

Current Draft: July 25, 2018§

Abstract

We study how the precision of information about an unknown state of the world affects equilibria and welfare in Bayesian games and decision problems. For an agent, more precise information leads to a mean-preserving spread of beliefs. We provide necessary and sufficient conditions to obtain either a non-increasing-mean or a non-decreasing-mean spread of actions whenever information precision increases for at least one agent.

We apply our Bayesian comparative statics framework to study informational externalities in strategic environments. In persuasion games, we derive sufficient conditions that lead to extremal disclosure of information. In oligopolistic markets, we characterize the incentives of firms to share information. In macroeconomic models, we show that information not only drives the amplitude of business cycles but also affects aggregate output. Finally, in a novel application, we compare the demand for information in covert and overt information acquisition games.

JEL Codes: C44, C61, D42, D81

*We are indebted to our advisors Jeff Ely, Eddie Dekel, Alessandro Pavan, Bruno Strulovici, and Asher Wolinsky for their constant guidance and encouragement. For very useful comments, we thank Laura Doval, Jay Lu, Ryota Iijima, Ilya Segal, Gabriel Carroll, Harry Di Pei, Federico Echenique, Rob Porter, Phil Reny, Marciano Siniscalchi, Jean Tirole, Zenon Zabinski, Andrei Gomberg, Tridib Sharma, and seminar participants at 2017 Los Angeles Economic Theory Conference, ITAM First Alumni Conference, 2018 Kansas Workshop in Economic Theory, Northwestern theory bag lunch and the students' theory lunch, and Caltech SISL.

†California Institute of Technology, SISL. Contact: mekonnen@caltech.edu

‡Northwestern University, Department of Economics. Contact: renelealv@u.northwestern.edu

§First Draft: November 2016

1 Introduction

The comparative statics of equilibrium welfare with respect to the quality of private or public information has long been of interest in economics. For example, private information could be harmful to agents in an exchange economy (Hirshleifer, 1971) or players in a game of imperfect information (Kamien et al., 1990) but never to a single Bayesian decision maker (Blackwell, 1951, 1953). In auction theory, Milgrom and Weber (1982) find that releasing public information about the common value of an object always increases revenue for the seller without affecting efficiency, while in the context of a Keynesian economy, Morris and Shin (2002) find that releasing a public signal sometimes can have a negative effect on welfare.

More recently, the effect of information on welfare has been studied in Bayesian games through the key concept of *informational externalities* (Angeletos and Pavan, 2007). These externalities are characterized by analyzing the effects that information has on equilibrium actions and then comparing the efficient and the equilibrium use of information.

Interestingly enough, the question of how information affects actions in games of incomplete information has only been partially studied in settings where closed-form solutions to equilibrium actions can be explicitly computed, namely, quadratic games with Gaussian information so that best responses are linear functions of the state and other players' actions.[1]

In this paper, we study how changes to the quality of private information in Bayesian games and decision problems affect equilibrium actions. We consider a general class of payoffs and information structures that nests the familiar linear-quadratic games with Gaussian signals. The comparative statics is a useful tool to understand how the quality of information about economic fundamentals (e.g., demand parameters in oligopolistic competition, or productivity parameters in macroeconomic models) affects economic outcomes (e.g., the dispersion of oligopoly prices, or the volatility of investment and aggregate output). From a normative perspective, this comparative statics is also a useful intermediate step to characterize informational externalities and investigate the welfare effects of information beyond linear-quadratic-Gaussian games.

Our theory of *Bayesian comparative statics* is comprised of three key ingredients: an information order (call it order 1), a stochastic ordering of equilibrium actions (call it order 2) and a class of utility functions. Our main result shows a duality between the order of actions and information: First, if signal A is more precise than signal B according to order 1, then for any preference in the class of utility functions, A induces equilibrium actions that are more

[1]For examples, see the symposium in Pavan and Vives (2015) and references therein.

dispersed according to order 2 than signal B does. Second, if signal A induces more dispersed equilibrium actions than signal B does for all preferences in the class of utility functions, then A is necessarily more precise than B according to order 1.

We illustrate the usefulness of our approach through several examples and applications. In persuasion games, we characterize conditions for which extremal information (either full revelation or no information) is the optimal persuasion policy. We also extend the industrial organization literature on information sharing in oligopolies to non-linear-quadratic environments.[2] In macroeconomic models, we show how information precision affects the amplitude of the business cycle, and emphasize that the effect of information on the expected aggregate output is important for studying welfare.

In a novel application, we compare the demand for information in two games of information acquisition, one in which information acquisition is a covert action and another in which it is overt. We apply our theory of Bayesian comparative statics to give a taxonomy of the demand for information in these games,[3] as well as analyze the role of information acquisition as a barrier to entry in oligopolistic competition.

To concretely motivate our comparative statics question and illustrate our approach to the normative and positive effects of information, we first consider the following example of monopoly production with an uncertain cost parameter.

1.1 A Simple Example

A monopolist faces a demand curve $P(q) = 1 - q$ and a cost function $c(\theta, q) = (1 - \theta)q + q^2/2$, where q is the quantity produced and θ is a cost parameter. However, the cost parameter is an unobserved random variable that is uniformly distributed on the unit interval.

The monopolist instead observes a signal such that with probability $\rho \in [0, 1]$, the signal realization s matches the realized cost parameter ($s = \theta$), and with probability $1 - \rho$, the signal realization s is uniformly and independently drawn from the unit interval. The quality of the signal is increasing in ρ: the signal is uninformative when $\rho = 0$ and fully revealing when $\rho = 1$.

A standard normative question in this example is, *"how does ρ affect consumer surplus?"* The signal quality affects the monopolist's production decision which in turn affects consumer welfare. Thus, in order to identify the consumer welfare effects, we must first answer the positive

[2] For example, Raith (1996) and Myatt and Wallace (2015).

[3] The taxonomy of overt vs covert information acquisition is also connected to the seminal work of Fudenberg and Tirole (1984) and Bulow et al. (1985) on capacity investment in the context of entry, accommodation and exit in oligopolistic markets.

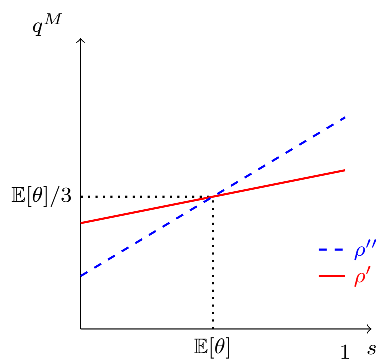
question, *"how does ρ affect the monopolist's production decision?"*

From an interim perspective, a monopolist that observes a signal realization s when the signal quality is ρ optimally produces

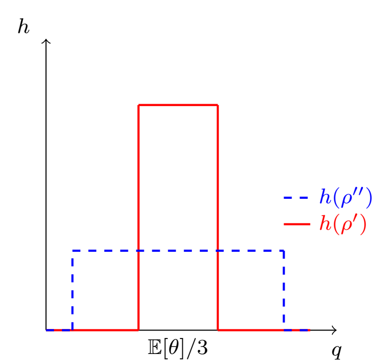
$$q^M(s;\rho) = \underbrace{\frac{\mathbb{E}[\theta]}{3}}_{\substack{\text{Based only}\\ \text{on prior}}} + \underbrace{\rho\left(\frac{s-\mathbb{E}[\theta]}{3}\right)}_{\text{Based on signal}}.$$

From an ex-ante perspective, the optimal quantity is a random variable whose distribution is given by $H(z;\rho) = \mathbb{P}(\{s : q^M(s;\rho) \leq z\})$, the probability that the monopolist optimally produces at most z units given a signal of quality ρ. *Our goal in this paper is to characterize how $H(\cdot;\rho)$ changes when ρ increases.*

In Figure 1a, the rotation of the solid line, $q^M(\cdot;\rho')$, to the dashed line, $q^M(\cdot;\rho'')$, captures the more "extreme" production decision when signal quality increases from ρ' to ρ''. The monopolist produces more when she observes "good news" ($s > \mathbb{E}[\theta]$) from ρ'' than from ρ' because good news from ρ'' is a stronger evidence of high values of θ (low marginal cost). Symmetrically, the monopolist produces less when she observes "bad news" ($s < \mathbb{E}[\theta]$) from ρ'' than from ρ' because bad news from ρ'' is a stronger evidence of low values of θ (high marginal cost). The rotation of q^M induces a mean-preserving spread in the distribution H, as shown by the density function h, in Figure 1b.



(a) Quantity produced



(b) PDF of production decision

Figure 1: Monopolist signal quality and production decision

From the demand function, we can derive the consumer surplus function $CS(q) = 0.5q^2$. The convexity implies that, ex-ante, consumers benefit from a more dispersed distribution of quantity, i.e., $\int CS(z)dH(z;\rho'') > \int CS(z)dH(z;\rho')$ for any $\rho'' > \rho'$. Thus, we have an answer to our positive and normative question.

Claim 1 *In a monopoly with linear demand and quadratic cost, an increase in signal quality induces a mean-preserving spread of quantities which in turn increases expected consumer welfare. In other words, the social value of information exceeds the monopolist's private value of information.*

The result, however, makes heavy use of the quadratic profit function and the "truth-or-noise" signal. This paper develops the tools so we can address such normative and positive questions for a general class of utility functions and information structures (signals). We revisit the monopoly problem in Section 4.1 to characterize the environments where a Pigouvian informational subsidy is desirable in a monopolistic market.

We now proceed to give a detailed description of the paper.

1.2 The Theory of Bayesian Comparative Statics

We first analyze the case of a single-agent Bayesian decision problem and characterize how the quality of the agent's signal affects the induced distribution of her optimal action. We consider a setting in which the agent has a supermodular utility function—the agent prefers to take higher actions for higher states of the world.

There are three main ingredients to the comparative statics result: an order over the distributions of optimal actions that captures changes in the mean and dispersion, an order over information structures that captures quality, and a class of utility functions that leads to a "duality" between the two orders.

An information structure ρ induces a distribution of optimal actions $H(\rho)$. For two information structures ρ'' and ρ', we say the agent is more *responsive with a higher mean* under ρ'' than ρ' if $H(\rho'')$ dominates $H(\rho')$ in the increasing convex order. Alternatively, we say the agent is more *responsive with a lower mean* under ρ'' than ρ' if $H(\rho'')$ dominates $H(\rho')$ in the decreasing convex order.[4]

To compare the quality of information, we first restrict attention to information structures in which higher signal realizations lead to first-order stochastic shifts in posterior beliefs. For two information structures ρ'' and ρ', we say ρ'' dominates ρ' in the supermodular stochastic order if, loosely speaking, the signals from ρ'' are more correlated with the state of the world than are the signals from ρ'.

[4] $H(\rho'')$ dominates $H(\rho')$ in the decreasing convex order if, and only if, $H(\rho')$ second-order stochastically dominates $H(\rho'')$.

Our main result shows that an agent, whose marginal utility function is supermodular and convex (in actions), is more responsive with a higher mean under ρ'' than under ρ' if ρ'' dominates ρ' in the supermodular stochastic order. Furthermore, we show that if every agent with supermodular and convex marginal utility is more responsive with a higher mean under ρ'' than under ρ', then ρ'' necessarily dominates ρ' in the supermodular stochastic order.

We also present symmetric results linking responsiveness with a lower mean to preferences with a submodular and concave marginal utility. Furthermore, we provide an example in which a higher quality of information does not lead to a more dispersed distribution of actions when the conditions on the agent's marginal utility function are violated.

We then extend our comparative statics results to Bayesian games with strategic complementarities. The players receive private signals of varying quality about the underlying state of the world before playing a game. Similar to the single agent case, under supermodularity and convexity conditions (resp., submodularity and concavity) on the players' marginal utilities, we show that a higher quality of information for any one player makes all players more responsive with a higher (resp., lower) mean, i.e., a more dispersed distribution of Bayesian Nash equilibrium actions along with an increase (resp., decrease) in the mean equilibrium actions for all players.

Our analysis points out a more intricate interaction between a player's equilibrium strategy and the quality of information than has been previously studied. First, we generalize the observation in linear-quadratic games that a player's distribution of best-responses becomes more dispersed when that player's own signal becomes more informative. Furthermore, even when the quality of information is held fixed, we show that a player's distribution of best-responses becomes more dispersed if another player's distribution of actions becomes more dispersed. Our main result shows that the combination of these effects is that players are not only responsive to changes in the quality of their own signals but also to changes in the quality of their opponent's signals.

We present several examples—generalized beauty contests, joint ventures with uncertain returns, and network games with random graphs—in which our result can be readily applied to study informational externalities. We also present several applications of our comparative statics results. A reader who is more interested in these applications may skip ahead to Section 4. As an application of the comparative statics in single-agent decision problems, we reconsider the monopolist example from Section 1.1 in a more general setting and study how a social planner should regulate the quality of the monopolist's information. Additionally, in a Bayesian persuasion framework, we derive sufficient conditions under which extremal information disclosure is optimal.

As an application of the comparative statics in games, we derive sufficient conditions on payoffs for which full information sharing between players in a Bayesian game is optimal, thereby extending the literature on information sharing in oligopolies beyond linear-quadratic payoffs and Gaussian signals. Additionally, we consider a novel approach to studying the strategic effect of information by comparing two different classes of endogenous information acquisition: one in which information acquisition is a covert activity (a player cannot observe how much information her opponents acquire) and another in which information acquisition is an overt activity. We provide a taxonomy of the strategic effect, which we call the *value of transparency*, and explore its connection to the demand and value of information in overt and covert games as well as analyze the strategic role of information acquisition as a barrier to entry in oligopolistic competition.

1.3 Related Literature

From a methodological point of view, this paper contributes to the literature on the theory of monotone comparative statics (Milgrom and Shannon, 1994; Milgrom and Roberts, 1994; Athey, 2002; Quah and Strulovici, 2009). Athey (2002) and Quah and Strulovici (2009) show that optimal actions increase as beliefs become more favorable. We take the next step and show how the distribution of optimal actions change as the distribution over beliefs changes.[5]

Our work also relates to literature on the value of information: Blackwell (1951, 1953), Lehmann (1988), Persico (2000), Quah and Strulovici (2009), and Athey and Levin (2017). In particular, Athey and Levin show that in the class of supermodular payoff functions, an agent values more information if, and only if, information quality is increasing in the supermodular stochastic order. For payoffs that additionally exhibit supermodular and convex (or submodular and concave) marginal utilities, we show that the agent's optimal actions are more dispersed if and only if information quality is increasing in the supermodular stochastic order.

When we move to Bayesian games, the references on comparative statics of equilibria include Vives (1990), Milgrom and Roberts (1994), Villas-Boas (1997), Van Zandt and Vives (2007). The value of information in Bayesian games with complementarities has also been recently studied by Amir and Lazzati (2016). Amir and Lazzati show that in the class of games with supermodular payoff functions, the value of information is increasing and convex in the

[5]In the context of our motivating example, Athey (2002) provides comparative statics results on $q^M(s;\rho)$ as a function of the signal realization s for a fixed ρ. We instead provide comparative statics results for the entire mapping $q^M(\cdot;\rho)$ as a function of ρ.

supermodular stochastic order. For payoffs that additionally exhibit supermodular and convex (or submodular and concave) marginal utilities, we show that the equilibrium actions for all players become more dispersed if information quality for any of the players increases in the supermodular stochastic order.

As we have mentioned in the introduction, this paper also relates to the vast literature on the use and social value of information, going back at least to Radner (1962) and Hirshleifer (1971). More recently Morris and Shin (2002) and Angeletos and Pavan (2007) fostered renewed interest and Ui and Yoshizawa (2015) gave a complete characterization of the social value of information in quadratic games with normally distributed public and private signals.

Hellwig and Veldkamp (2009) studied the problem of information acquisition within the framework of quadratic games and noticed the inheritance of the complementarity in actions to information acquisition. Colombo et al. (2014) study how the social value of public information is affected by private information acquisition decisions in a more flexible quadratic framework, and Myatt and Wallace (2011) notably allow for endogenously determined public information in a similar quadratic game of information acquisition.

We also contribute to the industrial organization literature on information sharing in oligopoly surveyed in Raith (1996) and recently picked up in Angeletos and Pavan (2007), Bergemann and Morris (2013) and more directly in Myatt and Wallace (2015). The comparative statics of equilibrium welfare and price or quantity dispersion are both of normative and positive importance for antitrust authorities, and we explore the robustness of the results to the assumption of quadratic economies.

Two papers that are closely related to ours but do not fit in the previous literatures are Jensen (2018) and Lu (2016). Jensen (2018) considers a decision-maker who has complete information about the state of the world. His paper characterizes how changes in the distribution over the state of the world affect the induced distribution over optimal actions.[6] Moreover, in the application to games, Jensen only considers exogenous changes to the distribution of independent private types.

Lu (2016) studies how the quality of information affects the value of a menu. In particular, he shows that increasing the quality of information in Blackwell's order implies the cummulative distribution of the interim value of the menu becomes more dispersed (increases in the increasing-convex-order). We instead show that the choice from within a menu becomes more dispersed as the quality of information increases.

[6]In the context of our motivating example, the monopolist observes the state θ and optimally produces quantity $q^M(\theta)$. Jensen characterizes how different distributions over θ affect the distribution of $q^M(\theta)$.

Finally, our analysis of the value of *transparency* in Bayesian games is related to the characterization of strategic investment in sequential versus simultaneous games of complete information in Fudenberg and Tirole (1984) and Bulow, Geanakoplos, and Klemperer (1985). We defer a detailed discussion of the relationship to Section 4.4.

The remainder of the paper is structured as follows: In Section 2, we present the single agent framework and provide sufficient and necessary conditions for an agent to become more responsive as information quality increases. We extend the analysis to Bayesian games with strategic complementarities in Section 3. In Section 4, we present our four main applications. Section 5 concludes. Proofs that are not presented in the text are in the Appendix (Section 6).

2 Single-agent Model

2.1 Preliminary Definitions and Notation

Let $X \triangleq \times_{i=1}^{m} X_i$ be a compact subset of $\mathbb{R}^m$, and let $X_{-i} \triangleq \times_{j \neq i} X_j$. For $x'', x' \in X$, let $x'' \geq$ (resp., $>$)x' if $x_i'' \geq$ (resp., $>$)x_i' for $i = 1, 2, \ldots, m$.

We say a function $g : X \to \mathbb{R}$ is increasing in x_i if, for all $x_{-i} \in X_{-i}$, $x_i'' > x_i'$ implies $g(x_i'', x_{-i}) \geq g(x_i', x_{-i})$. We say g has increasing (resp., decreasing, or constant) differences in $(x_{-i}; x_i)$ if for any $x_{-i}'' \geq x_{-i}'$, $g(x_i, x_{-i}'') - g(x_i, x_{-i}')$ is increasing (resp., decreasing, or constant) in x_i. We emphasize that any references to "increasing/decreasing," "increasing/decreasing differences," or "concave/convex" are in the weak sense.

If g is a differentiable function, we write g_{x_i} as a shorthand for $\frac{\partial}{\partial x_i} g(x)$ and $g_{x_i x_j}$ for $\frac{\partial^2}{\partial x_i x_j} g(x)$. If g is differentiable and has increasing (resp., decreasing, or constant) differences in $(x_{-i}; x_i)$, then $g_{x_i x_j} \geq 0$ (resp., $g_{x_i x_j} \leq 0$, or $g_{x_i x_j} = 0$) for each $j \neq i$.

2.2 Setup

Let $A \triangleq [\underline{a}, \bar{a}]$ be the action space and let $\Theta \triangleq [\underline{\theta}, \bar{\theta}]$ represent the state space. We denote the random state variable by $\tilde{\theta}$ and the realization by θ. Let $\Delta(\Theta)$ denote the set of all Borel probability measures on Θ. An agent (she) has to choose an action $a \in A$ before observing the realized state of the world. The agent's prior belief is denoted by the measure $\mu^0 \in \Delta(\Theta)$. We allow for beliefs to be either discrete or absolutely continuous measures. Payoffs are given by the function $u : \Theta \times A \to \mathbb{R}$ such that

(A.1) $u(\theta, a)$ is uniformly bounded, measurable in θ, and twice differentiable in a, (A.2) for all $\theta \in \Theta$, $u(\theta, \cdot)$ is strictly concave in a with $u_{aa}(\theta, \cdot) < 0$,

(A.3) for all $\theta \in \Theta$, there exists an action $a \in A$ such that $u_a(\theta, a) = 0$, and

(A.4) $u(\theta, a)$ has increasing differences in $(\theta; a)$.

Increasing differences (ID) implies that the agent prefers a high action when the state is high and a low action when the state is low. Assumptions (A.1)-(A.3) allow us to characterize the optimal actions by their first order conditions.[7]

Given any belief $\mu \in \Delta(\Theta)$, define

$$a^*(\mu) = \arg\max_{a \in A} \int_\Theta u(\theta, a)\mu(d\theta).$$

The continuity and strict concavity of the utility function along with the compactness of A guarantee that a unique and measurable solution exists. Furthermore, (A.4) implies $a^*(\mu_2) \geq a^*(\mu_1)$ whenever μ_2 first-order stochastically dominates μ_1 (Athey, 2002).[8]

Prior to decision-making, the agent can observe an informative random signal $\tilde{s}$ about the unknown state. We denote the signal realization by s to distinguish it from the random signal. Signals are generated by an information structure $\Sigma_\rho \triangleq \langle S, F(\cdot, \cdot; \rho)\rangle$ where $S \subseteq \mathbb{R}$ is the signal space, $F(\cdot, \cdot; \rho) : \Theta \times S \to [0, 1]$ is a joint probability distribution over $(\tilde{\theta}, \tilde{s})$, and ρ is an index that is useful when comparing different signal structures.

We denote the marginal distribution of $\tilde{\theta}$ by $F_\Theta(\cdot; \rho) : \Theta \to [0, 1]$. However, any information structure Σ_ρ induces the same marginal $F_\Theta(\theta; \rho) = F_\Theta(\theta) = \int_{\underline{\theta}}^{\theta} \mu^0(d\omega)$ which depends only on the prior.

Similarly, we denote the marginal distribution of $\tilde{s}$ by $F_S(\cdot; \rho) : S \to [0, 1]$. Without loss of generality, we assume that all information structures induce the same marginal on $\tilde{s}$, i.e., $F_S(s; \rho) = F_S(s)$ for all $s \in S$. Moreover, F_S has a positive bounded density f_S.[9]

[7] In Section 6.2.1, we discuss the difficulties that arise when some of these assumptions are violated.

[8] We say that μ_2 first-order stochastically dominates μ_1, denoted $\mu_2 \succeq_{FOSD} \mu_1$, if for any increasing function $g : \Theta \to \mathbb{R}$, $\int_\Theta g(\theta)\mu_2(d\theta) \geq \int_\Theta g(\theta)\mu_1(d\theta)$.

[9] The assumption is without loss of generality: we can apply the integral probability transform to any random signal $\tilde{s}$ with a continuous marginal distribution and create a new signal which is uniformly distributed on the unit interval. If the marginal distribution of $\tilde{s}$ is discontinuous at $\tilde{s} = s^*$ with $F_S(s^*; \rho) = q$, then, as noted by Lehmann (1988), we can construct a new signal, $\tilde{s}'$, where $\tilde{s}' = \tilde{s}$ if $\tilde{s} < s^*$, $\tilde{s}' = \tilde{s} + q\tilde{t}$ if $\tilde{s} = s^*$, and $\tilde{s}' = \tilde{s} + q$ if $\tilde{s} > s^*$, where $\tilde{t} \sim U(0, 1)$. The new signal $\tilde{s}'$ is equally informative as $\tilde{s}$ and has a continuous and strictly increasing marginal distribution.

2.3 Order 1: Actions

From an interim perspective, the agent first observes a signal realization $s \in S$ from an information structure Σ_ρ, updates her beliefs to a posterior $\mu(\cdot|s;\rho) \in \Delta(\Theta)$ via Bayes rule, and then chooses the optimal action $a^*\big(\mu(\cdot|s;\rho)\big)$. Define the measurable function $a(\rho) : S \to A$ by $a(s;\rho) = a^*\big(\mu(\cdot|s;\rho)\big)$.

From an ex-ante perspective, the signal realizations are yet to be observed. Therefore, $a(\rho)$ is a random variable that is distributed according to the CDF $H(\cdot;\rho)$ which is defined as

$$H(z;\rho) \triangleq \int_{\{s:a(s;\rho)\leq z\}} dF_S(s)$$

for $z \in \mathbb{R}$.

Given two information structures $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$, we say that $a(\rho'')$ dominates $a(\rho')$ in the *increasing convex order* if

$$\int_{-\infty}^{\infty} \varphi(z)dH(z;\rho'') \geq \int_{-\infty}^{\infty} \varphi(z)dH(z;\rho')$$

for any increasing convex function $\varphi : \mathbb{R} \to \mathbb{R}$. Alternatively, we say that $a(\rho'')$ dominates $a(\rho')$ in the *decreasing convex order* if

$$\int_{-\infty}^{\infty} \phi(z)dH(z;\rho'') \geq \int_{-\infty}^{\infty} \phi(z)dH(z;\rho')$$

for any decreasing convex function $\phi : \mathbb{R} \to \mathbb{R}$.[10] If $a(\rho'')$ dominates $a(\rho')$ in both the increasing convex and decreasing convex order, then $a(\rho'')$ is a mean-preserving spread of $a(\rho')$.

Definition 1 (Responsiveness) *Given two information structures $\Sigma_{\rho''}$ and $\Sigma_{\rho'}$, we say that*

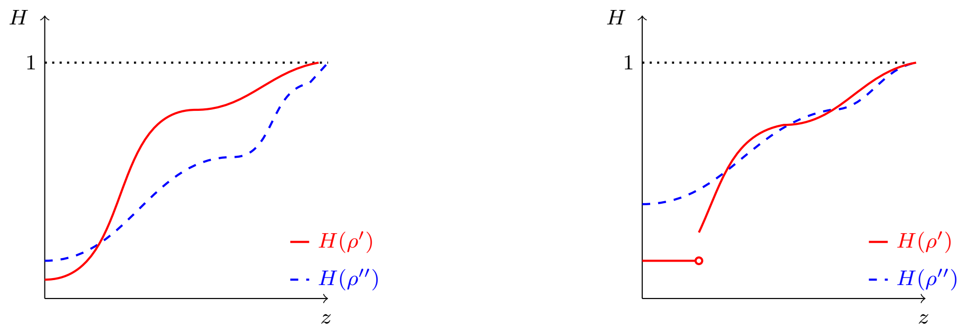
*i. an agent is **more responsive with a higher mean** under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$ if $a(\rho'')$ dominates $a(\rho')$ in the **increasing convex order**, and*

*ii. an agent is **more responsive with a lower mean** under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$ if $a(\rho'')$ dominates $a(\rho')$ in the **decreasing convex order**.*

Lemma 1 (Appendix) provides an equivalent characterization of responsiveness based on the CDF $H(\cdot;\rho)$. Figure 2 plots the distribution over actions induced by two information

[10] $a(\rho'')$ dominates $a(\rho')$ in the decreasing convex order if, and only if, $a(\rho')$ second-order stochastically dominates $a(\rho'')$.

structures $\Sigma_{\rho''}$ and $\Sigma_{\rho'}$. In Figure 2a, integrating $H(\cdot;\rho'') - H(\cdot;\rho')$ right to left always yields a negative value which, by Lemma 1, implies responsiveness with a higher mean.[11] In contrast, in Figure 2b, integrating $H(\cdot;\rho'') - H(\cdot;\rho')$ left to right always yields a positive value which, by Lemma 1, implies responsiveness with a lower mean.



(a) Responsiveness with a higher mean (b) Responsiveness with a lower mean

Figure 2: CDF of Optimal Actions and Responsiveness

2.4 Order 2: Information

The next step is to determine an appropriate way to compare different information structures. We first restrict attention to information structures in which higher signal realizations lead to a first-order stochastic increase in beliefs. This assumption is weaker than the monotone likelihood ratio property commonly assumed in settings with complementarities.

(A.5) For any given information structure Σ_ρ, $s' > s$ implies $\mu(\cdot|s';\rho) \succeq_{FOSD} \mu(\cdot|s;\rho)$.

Definition 2 (Supermodular Stochastic Order) *Given two information structures $\Sigma_{\rho''}$ and $\Sigma_{\rho'}$, we say that $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the supermodular stochastic order, denoted $\rho'' \succeq_{spm} \rho'$, if $F(\theta, s;\rho'') \geq F(\theta, s;\rho')$ for all $(\theta, s) \in \Theta \times S$.*

Intuitively, $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the supermodular stochastic order if $\tilde{\theta}$ and $\tilde{s}$ are more positively correlated under $\Sigma_{\rho''}$. By (A.5), low signal realizations are evidence of low states. The agent considers a signal $\tilde{s} \leq s$ from $\Sigma_{\rho''}$ as a stronger evidence of a low state could be low

[11] Additionally, the area between the y-axis and $H(\cdot;\rho'')$ is bigger than the area between the y-axis and $H(\cdot;\rho')$ which implies that $\Sigma_{\rho''}$ induces optimal actions with a higher mean than $\Sigma_{\rho'}$.

(than a signal $\tilde{s} \leq s$ from $\Sigma_{\rho'}$). Thus, $\mathbb{P}(\tilde{\theta} \leq \theta | \tilde{s} \leq s; \rho'') \geq \mathbb{P}(\tilde{\theta} \leq \theta | \tilde{s} \leq s; \rho')$. By assumption (WLOG), the marginal on the signals are the same for both $\Sigma_{\rho''}$ and $\Sigma_{\rho'}$. Hence,

$$F(\theta, s; \rho'') = \mathbb{P}(\tilde{\theta} \leq \theta | \tilde{s} \leq s; \rho'') F_S(s) \geq \mathbb{P}(\tilde{\theta} \leq \theta | \tilde{s} \leq s; \rho') F_S(s) = F(\theta, s; \rho').$$

For example, the class of "truth-or-noise" information structures we considered in Section 1.1 are ordered by the supermodular stochastic order. Another example is the class of Gaussian information structures such that $\tilde{\theta}$ and $\tilde{s}$ are both normally distributed with mean θ_0, variance σ^2, and have a correlation coefficient of $\rho \in [0, 1]$. In both cases, $\rho'' \succeq_{spm} \rho'$ if $\rho'' > \rho'$.

In the Online Appendix (Mekonnen and Leal-Vizcaíno, 2018), we elaborate that given (A.5), the supermodular stochastic order nests the familiar Blackwell informativeness (Blackwell, 1951, 1953) and the Lehmann (accuracy) order (Lehmann, 1988).[12] We also provide an example of non-parametric information structures that can be ranked by the supermodular stochastic order but not by either the Blackwell or the Lehmann order.

2.5 Preferences and Main Result

The main contribution of this paper is to identify a class of decision problems for which the agent becomes more responsive when information quality increases according to the supermodular stochastic order.

Let $\mathcal{U}^I$ be the class of payoff functions $u : \Theta \times A \to \mathbb{R}$ that satisfy (A.1)-(A.4) and have a marginal utility $u_a(\theta, a)$ that

(i) is convex in a for all $\theta \in \Theta$, and (ii) has increasing differences in $(\theta; a)$.

In other words, a utility function $u \in \mathcal{U}^I$ exhibits a marginal utility that diminishes at a *diminishing rate* and an *increasing* increasing differences in $(\theta; a)$. Below, we show that an agent with a payoff function $u \in \mathcal{U}^I$ becomes more responsive with a higher mean as information quality increases in the supermodular stochastic order.

Similarly, let $\mathcal{U}^D$ be the class of payoff functions $u : \Theta \times A \to \mathbb{R}$ that satisfy (A.1)-(A.4) and have a marginal utility $u_a(\theta, a)$ that

[12]See Persico (2000) and Jewitt (2006) for detailed description and applications.

(*i*) is concave in a for all $\theta \in \Theta$, and (*ii*) has decreasing differences in $(\theta; a)$.

In other words, a utility function $u \in \mathcal{U}^D$ exhibits a marginal utility that diminishes at an *accelerating rate* and a *decreasing* increasing differences in $(\theta; a)$. Below, we show that an agent with a payoff function $u \in \mathcal{U}^D$ becomes more responsive with a lower mean as information quality increases in the supermodular stochastic order.[13]

Theorem 1 *Consider two information structures $\Sigma_{\rho''}$ and $\Sigma_{\rho'}$ that satisfy* (A.5)*. Any agent with payoff $u \in \mathcal{U}^I$ $\big($resp., $u \in \mathcal{U}^D\big)$ is more responsive with a higher (resp., lower) mean under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$ if, and only if, $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the supermodular stochastic order.*

When information quality increases, the distribution over the agent's posterior beliefs becomes more dispersed. Theorem 1 provides the conditions on the agent's utility function under which we can map the more dispersed distribution of posterior beliefs to a more dispersed distribution of actions that incorporates monotone changes to the average optimal action.

The mechanism behind Theorem 1 is best understood through Proposition 1 which shows that when $u \in \mathcal{U}^I$ $\big($resp.,$u \in \mathcal{U}^D\big)$, optimal actions are "convex" (resp., "concave") in the agent's posterior belief.

Proposition 1 *Let $\mu_1, \mu_2 \in \Delta(\Theta)$ be any two beliefs with $\mu_2 \succeq_{FOSD} \mu_1$. If $u \in \mathcal{U}^I$, then for any $\lambda \in [0, 1]$*

$$a^*\big(\lambda\mu_1 + (1 - \lambda)\mu_2\big) \leq \lambda a^*(\mu_1) + (1 - \lambda)a^*(\mu_2)$$

If $u \in \mathcal{U}^D$, the opposite inequality holds.

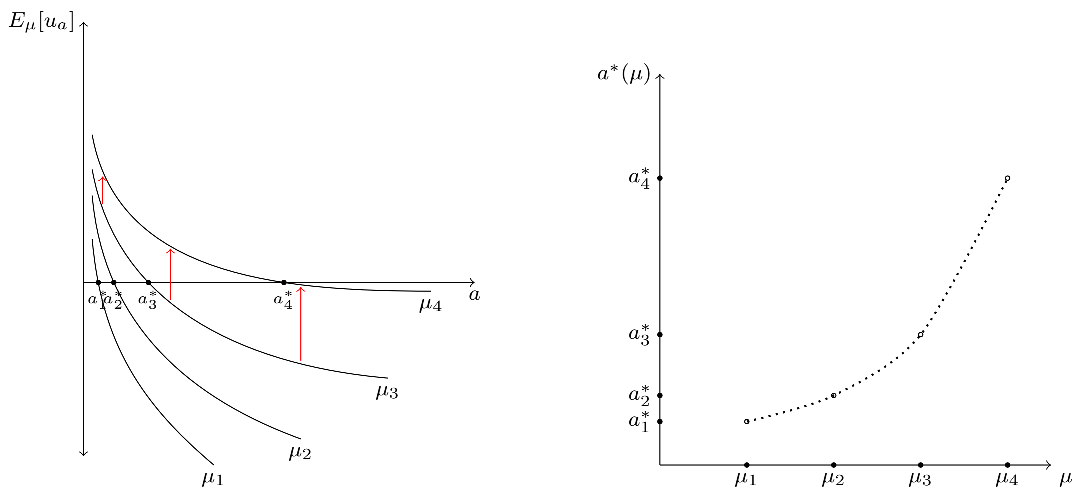
Henceforth, we focus on payoffs in $\mathcal{U}^I$ but the arguments we provide can be symmetrically applied to payoffs in $\mathcal{U}^D$.

For a simple visual representation, let the state space be $\Theta = \{\underline{\theta}, \bar{\theta}\}$ with $\bar{\theta} > \underline{\theta}$. With some abuse of notation, let $\mu \in [0, 1]$ represent the agent's belief that $\tilde{\theta} = \bar{\theta}$. Consider four different beliefs $\{\mu_n\}_{n=1,2,3,4}$ such that, $\mu_n = n\delta$ for some $\delta \in (0, 1/4)$. Figure 3a below plots out the expected marginal utility of a payoff function $u \in \mathcal{U}^I$ for the different beliefs. The optimal action $a_n = a^*(\mu_n)$ is given by the action at which the expected marginal utility under belief μ_n intersects the x-axis. Since $\mu_4 \succeq_{FOSD} \mu_3 \succeq_{FOSD} \mu_2 \succeq_{FOSD} \mu_1$ and $u(\theta, a)$ satisfies ID, $a_4 \geq a_3 \geq a_2 \geq a_1$.

[13]The class of functions $\mathcal{U}^I$ (resp., $\mathcal{U}^D$) is a superset of ultramodular (resp., inframarginal) functions. See Marinacci and Montrucchio (2005) for an analysis of ultra/inframodular functions and the connection to cooperative game theory.

ID of $u_a(\theta, a)$ implies that the gap between the expected marginal utilities of μ_{n+1} and μ_n is widening as the action increases (the height of the red arrows increases left to right). In such a case, for a small $\epsilon > 0$, the agent's benefit from increasing a_2 to $a_2 + \epsilon$ when beliefs increase from μ_2 to μ_3 is larger than her benefit from increasing a_1 to $a_1 + \epsilon$ when beliefs increase from μ_1 to μ_2, and so on.

In contrast, concavity of $u(\theta, a)$ in a implies that the agent's benefit from increasing a_2 to $a_2 + \epsilon$ is less than her benefit from increasing a_1 to $a_1 + \epsilon$ for any fixed belief, and so on. Thus, there are two opposing forces at work. However, when $u_a(\theta, a)$ is convex in a, the marginal utility does not diminish too quickly. This *diminishing* diminishing marginal utility is captured in Figure 3a by the convex marginal utilities curves. All these properties combined result in $a_4 - a_3 > a_3 - a_2 > a_2 - a_1$. Figure 3b depicts this "convexity" property as described in Proposition 1.



(a) Marginal utilities

(b) Optimal action

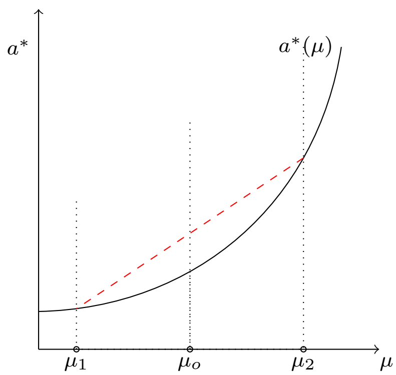
Figure 3: Convexity for $u \in \mathcal{U}^I$

To see how the "convexity" of the optimal action is related to responsiveness, let us continue with the above simplified setting. Let Figure 4a represent the convex optimal action (as a function of posteriors) of some agent with utility $u \in \mathcal{U}^I$. Let $\mu_0 \in (0, 1)$ be the agent's prior belief.

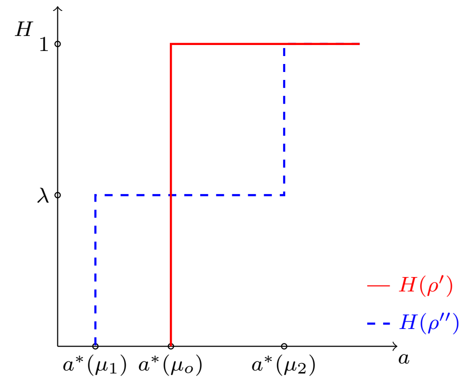
Let $\Sigma_{\rho'}$ be a completely uninformative information structure which induces $a^*(\mu_0)$ with probability one. Let $\Sigma_{\rho''}$ be a more informative structure that induces two posteriors $\{\mu_1, \mu_2\}$ with probability $\{\lambda, 1 - \lambda\}$. Hence, it induces $a^*(\mu_1)$ with probability λ and $a^*(\mu_2)$ with probability $1 - \lambda$. Bayes-consistency implies $\mu_0 = \lambda\mu_1 + (1 - \lambda)\mu_2$. Since $u \in \mathcal{U}^I$, from Proposition 1,

$\lambda a^*(\mu_1) + (1-\lambda)a^*(\mu_2) \geq a^*(\lambda\mu_1 + (1-\lambda)\mu_2) = a^*(\mu_0)$, i.e., $\Sigma_{\rho''}$ induces a higher average optimal action than $\Sigma_{\rho'}$.

Figure 4b maps the induced distributions of actions, $H(\cdot;\rho'')$ (the dashed line) and $H(\cdot;\rho')$ (the solid line). The integral $\int_x^\infty H(z;\rho'') - H(z;\rho')dz \leq 0$ for all $x \in \mathbb{R}$ which, by Lemma 1, implies that the agent is more responsive with a higher mean under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$.



(a) Optimal action



(b) Induced distribution

Figure 4: Convexity of a^* and responsiveness with higher mean

Corollary 1 *Let $\Sigma_{\rho''}$ be an information structure that satisfies* (A.5). *Let $\Sigma_{\rho'}$ be any* garbling *of $\Sigma_{\rho''}$. If an agent has utility $u \in \mathcal{U}^I$ $\left(\text{resp., } u \in \mathcal{U}^D\right)$, then the agent is more responsive with a higher (resp., lower) mean.*

Remark 1 *Proposition 1 directly implies Corollary 1, which shows that the agent becomes more responsive when information quality increases in the Blackwell order. While the result appears to be an implication of Theorem 1, there is a subtle difference—the garbling $\Sigma_{\rho'}$ does not have to satisfy* (A.5).[14] *The proof is in Mekonnen and Leal-Vizcaíno (2018).*

Remark 2 *Whenever $u \notin \mathcal{U}^I$, there exist beliefs $\mu_1, \mu_2 \in \Delta(\Theta)$ for which Proposition 1 is violated. Hence, we can find a prior $\mu^0 \in \Delta(\Theta)$ and information structures $\Sigma_{\rho''}, \Sigma_{\rho'}$ with $\rho'' \succeq_{spm} \rho'$ such that the agent is not more responsive with a higher mean under ρ''. In this sense, the class of preferences $\mathcal{U}^I$ is not only sufficient but also necessary for responsiveness with a higher mean. We present such an example in Section 6.2.1.*

[14] $\Sigma_{\rho'}$ is a garbling of $\Sigma_{\rho''}$ if there exist stochastic maps $\{\xi(\cdot|\hat{s})\}_{\hat{s}\in S}$ with $\xi(\cdot|\hat{s}) : S \to [0,1]$ such that $F(\theta, s; \rho') = \int_{[\underline{\theta},\theta]\times S} \xi(s|\hat{s})dF(\omega, \hat{s}; \rho'')$ for each $(\theta, s) \in \Theta \times S$.

3 Games

In this section, we extend our results from the single-agent framework to games of incomplete information with strategic complementarities. This class of games includes beauty contests and quadratic games, oligopolistic competition, games with network effects, search models, and investment games, among others (see Milgrom and Roberts (1990)).

3.1 Setup

There are n players with $N \triangleq \{1, 2, \ldots, n\}$ denoting the set of players. Let $\Theta_i \triangleq [\underline{\theta}_i, \bar{\theta}_i]$ be the state space for player i and define $\Theta \triangleq \times_{i \in N} \Theta_i$ and $\Theta_{-i} \triangleq \times_{j \neq i} \Theta_j$. Let $\tilde{\theta} = (\tilde{\theta}_i, \tilde{\theta}_{-i})$ denote the random state variables, and let $\theta = (\theta_i, \theta_{-i})$ denote the realizations. The players hold a common prior $\mu^0 \in \Delta(\Theta)$. Once again, we allow for beliefs to be either discrete or absolutely continuous measures. Let F_{Θ_i} be the marginal distribution of $\tilde{\theta}_i$ induced by μ^0. Similarly, let $F_{\Theta_{-i}}(\cdot|\theta_i)$ be the joint distribution of $\tilde{\theta}_{-i}$ conditional on $\tilde{\theta}_i = \theta_i$. We assume that

(A.6) for all $i \in N$, $\theta_i' > \theta_i$ implies $F_{\Theta_{-i}}(\cdot|\theta_i') \succeq_{FOSD} F_{\Theta_{-i}}(\cdot|\theta_i)$,

which is a weaker assumption than affiliation.

Let $A_i \triangleq [\underline{a}_i, \bar{a}_i]$ be the action space of player i. Let $A \triangleq \times_{i \in N} A_i$ and $A_{-i} \triangleq \times_{j \neq i} A_j$. The payoff for each player $i = 1, ..., n$ is given by a utility function $u^i : \Theta \times A \to \mathbb{R}$ such that

(A.7) $u^i(\theta, a)$ is uniformly bounded, measurable in θ, continuous and twice differentiable in a,

(A.8) for all $(\theta, a_{-i}) \in \Theta \times A_{-i}$, $u^i(\theta, a_{-i}, \cdot)$ is strictly concave in a_i,

(A.9) for all $(\theta, a_{-i}) \in \Theta \times A_{-i}$, there exists an action $a_i \in A_i$ such that $u^i_{a_i}(\theta, a_{-i}, a_i) = 0$, and

(A.10) $u^i(\theta, a)$ has increasing differences in $(\theta, a_{-i}; a_i)$.

Similar to the single-agent framework, (A.10) implies that there are complementarities between the state of the world and a player's action. Additionally, there are strategic complementarities between the players' actions. Thus, when player j takes a higher action, player i wants to do the same.

Following the terminology introduced by Bergemann and Morris (2016), we decompose the entire game of incomplete information into two components: the basic game and the information structure. The basic game $G \triangleq (N, \{A_i, u^i\}_{i \in N}, \mu^0)$ is composed of (i) a set of players N, (ii) for each player $i \in N$, an action space A_i along with a payoff function $u^i : \Theta \times A \to \mathbb{R}$, and (iii)

a common prior $\mu^0 \in \Delta(\Theta)$. The setting is general enough to accommodate private or common values as well as independence or affiliation.

The second component of the game is the information structure: each player $i \in N$ observes a signal $\tilde{s}_i$ about $\tilde{\theta}_i$ from an information structure $\Sigma_{\rho_i} \triangleq \big(S_i, F(\cdot,\cdot;\rho_i)\big)$.[15] $S_i \subseteq \mathbb{R}$ is the signal space, $F(\cdot,\cdot;\rho_i) : \Theta_i \times S_i \to [0,1]$ is a joint probability distribution over $(\tilde{\theta}_i, \tilde{s}_i)$, and ρ_i is an index. Once again, we assume, without loss of generality, that any information structure Σ_{ρ_i} induces the same marginal on $\tilde{\theta}_i$, F_{Θ_i}, and the same marginal on $\tilde{s}_i$, F_{S_i}, with a positive and bounded density f_{S_i}.

Let $S \triangleq \times_{i \in N} S_i$ and $\Sigma_\rho \triangleq (\Sigma_{\rho_1}, \ldots, \Sigma_{\rho_n})$. An information structure Σ_ρ induces a joint distribution $\boldsymbol{F}(\cdot,\cdot;\rho) : \Theta \times S \to [0,1]$ over $(\tilde{\theta}, \tilde{s})$. The following are working assumptions for this section:

(A.11) For all $s \in S$ and $\theta \in \Theta$, $\boldsymbol{F}(s|\theta;\rho) = \prod_{i \in N} F(s_i|\theta_i;\rho_i)$.

(A.12) For all players $i \in N$, $s_i' > s_i$ implies $\mu(\cdot|s_i';\rho_i) \succeq_{FOSD} \mu(\cdot|s_i;\rho_i)$.

(A.13) For all players $i \in N$, $\theta_i' > \theta_i$ implies $F(\cdot|\theta_i';\rho_i) \succeq_{FOSD} F(\cdot|\theta_i;\rho_i)$.

Assumption (A.11) implies that player i can directly learn only about $\tilde{\theta}_i$, not $(\tilde{\theta}_{-i}, \tilde{s}_{-i})$. Assumption (A.12) is an extension of (A.5) and implies that higher signal realizations lead to a first-order increase in a player's belief. Assumption (A.13) implies the converse: higher states are likely to lead to higher signal realizations. A distribution over the state and signal space that satisfies the monotone likelihood ratio property jointly satisfies (A.12)-(A.13).

The full game of incomplete information is given by $\mathcal{G}_\rho \triangleq (\Sigma_\rho, G)$. Both components of the game are common knowledge. First, each player $i \in N$ privately observes a signal realization $s_i \in S_i$ generated from Σ_{ρ_i}. Then, the players participate in the basic game G by simultaneously choosing an action.

Momentarily ignoring existence issues, let $a^\star(\rho) = \big(a_1^\star(\rho), a_2^\star(\rho), \ldots, a_n^\star(\rho)\big)$ be a profile of pure strategy actions that constitute a Bayesian Nash equilibrium (BNE) of the game $\mathcal{G}_\rho$, and let $a_{-i}^\star(\rho)$ be the profile of BNE strategies excluding player i. For each player $i \in N$, $a_i^\star(\rho) : S_i \to A_i$ is a measurable function. We interpret $a_i^\star(s_i;\rho)$ as the solution to

$$\max_{a_i \in A_i} \int_{\Theta \times S_{-i}} u^i\big(\theta, a_{-i}^\star(s_{-i};\rho), a_i\big) d\boldsymbol{F}(\theta, s_{-i}|s_i;\rho).$$

[15] There is an implicit assumption in the setup that player i can directly learn only about $\tilde{\theta}_i$. We make this assumption explicit in (A.11).

In other words, $a_i^\star(s_i;\rho)$ is the action player i takes in an equilibrium of the game $\mathcal{G}_\rho$ when she observes signal realization s_i and her opponents use strategies $a_{-i}^\star(\rho)$. Fixing the basic game G, we are interested in how a change in the information structure from $\Sigma_{\rho'}$ to $\Sigma_{\rho''}$ affects the BNEs of the full game $\mathcal{G}_{\rho'} \triangleq (\Sigma_{\rho'}, G)$ and $\mathcal{G}_{\rho''} \triangleq (\Sigma_{\rho''}, G)$.

We restrict our attention to monotone BNEs, i.e., each player's equilibrium action, $a_i^\star(s_i;\rho)$ is increasing in the signal s_i.[16] The existence of monotone pure strategy BNE has long been established by the literature on supermodular Bayesian games. In particular, the existence result of Van Zandt and Vives (2007) is noteworthy in our setting; their existence result does not require players to have atomless posterior beliefs when they participate in the basic game.

3.2 Order 1: Bayesian Nash Equilibrium Actions

We parallel the single-agent framework as closely as possible. We first extend the responsiveness definition into a multi-player setting.

Definition 3 (Equilibrium Responsiveness) *Given two Bayesian games, $\mathcal{G}_{\rho''} \triangleq (\Sigma_{\rho''}, G)$ and $\mathcal{G}_{\rho'} \triangleq (\Sigma_{\rho'}, G)$, we say that*

- *players are **more responsive with a higher mean** under $\mathcal{G}_{\rho''}$ than $\mathcal{G}_{\rho'}$ if for each monotone BNE $a^\star(\rho')$ of $\mathcal{G}_{\rho'}$, there exists a monotone BNE $a^\star(\rho'')$ of $\mathcal{G}_{\rho''}$ such that $a_i^\star(\rho'')$ dominates $a_i^\star(\rho')$ in the **increasing convex order** for all $i \in N$, and*

- *players are **more responsive with a lower mean** under $\mathcal{G}_{\rho''}$ than $\mathcal{G}_{\rho'}$ if for each monotone BNE $a^\star(\rho'')$ of $\mathcal{G}_{\rho''}$, there exists a monotone BNE $a^\star(\rho')$ of $\mathcal{G}_{\rho'}$ such that $a_i^\star(\rho'')$ dominates $a_i^\star(\rho')$ in the **decreasing convex order** for all $i \in N$.*

The definition for responsiveness in the Bayesian game setting is more involved than the single-agent case because we have to take into account the possibility of multiple BNE outcomes. However, if we focus on a particular equilibrium selection, then we can restore the simpler definition of responsiveness used in the single-agent setting.

[16] By assumptions (A.6), (A.10), and (A.12), player i's best response is monotone in s_i when her opponents use monotone strategies. While restricting attention to monotone BNEs may be with loss of generality, extremal equilibria are nonetheless monotone. Specifically, the least and the greatest pure strategy monotone BNEs of a supermodular Bayesian game bound all other BNEs (Milgrom and Roberts (1990); Van Zandt and Vives (2007)).

3.3 Order 2: Information

We then extend the supermodular stochastic order from a single-agent framework into a setting with multiple information structures.

Definition 4 (Supermodular Stochastic Order in Games) *Given two profile of information structures $\Sigma_{\rho''} \triangleq (\Sigma_{\rho''_1}, \Sigma_{\rho''_2}, \ldots, \Sigma_{\rho''_n})$ and $\Sigma_{\rho'} \triangleq (\Sigma_{\rho'_1}, \Sigma_{\rho'_2}, \ldots, \Sigma_{\rho'_n})$, we say $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the supermodular stochastic order, denoted $\rho'' \succeq_{spm} \rho'$, if $\Sigma_{\rho''_i}$ dominates $\Sigma_{\rho'_i}$ in the supermodular stochastic order for all $i \in N$.*

3.4 Preferences and Main Result for Games

Let Γ^I be the class of payoff functions $u : \Theta \times A \to \mathbb{R}$ that satisfy (A.7)-(A.10) and have a marginal utility $u_{a_i}(\theta, a)$ that, for all $j \in N$,

(i) is convex in a_j for all $(\theta, a_{-j}) \in \Theta \times A_{-j}$, (ii) has increasing differences in $(\theta, a_{-j}; a_j)$.

Below, we show that payoffs in Γ^I are linked to responsiveness with a higher mean.[17]

Let Γ^D be the class of payoff functions $u : \Theta \times A \to \mathbb{R}$ that satisfy (A.7)-(A.10) and have a marginal utility $u_{a_i}(\theta, a)$ that, for all $j \in N$,

(i) is concave in a_j for all $(\theta, a_{-j}) \in \Theta \times A_{-j}$, (ii) has decreasing differences in $(\theta, a_{-j}; a_j)$.

Below, we show that payoffs in Γ^D are linked to responsiveness with a lower mean.

Theorem 2 *Consider two Bayesian games $\mathcal{G}_{\rho''} \triangleq (\Sigma_{\rho''}, G)$ and $\mathcal{G}_{\rho'} \triangleq (\Sigma_{\rho'}, G)$ in which $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the supermodular stochastic order. If $u^i \in \Gamma^I$ $\big($ resp., $u^i \in \Gamma^D\big)$ for all $i \in N$, then players are more responsive with a higher (resp., lower) mean under $\mathcal{G}_{\rho''}$ than $\mathcal{G}_{\rho'}$.*

The proof for Theorem 2 can be found in the Online Appendix (Mekonnen and Leal-Vizcaíno, 2018). Here, we provide a brief sketch which proceeds in four steps. Suppose $u^i \in \Gamma^I$ for all $i \in N$, and consider a profile of information structures $\Sigma_{\rho''}$ and $\Sigma_{\rho'}$. Fix a player $i \in N$.

1. Holding all else fixed, a higher quality of own information leads to a more dispersed distribution of best-responses.

[17] Note that $\Gamma^I \subseteq \mathcal{U}^I$. Furthermore, if $u(\theta, a)$ is independent of (θ_{-i}, a_{-i}) and $u \in \mathcal{U}^I$, then $u \in \Gamma^I$.

- Suppose $\rho''_i \succeq_{spm} \rho'_i$, and $\Sigma_{\rho''_j} = \Sigma_{\rho'_j}$ for all $j \neq i$. Fix a monotone strategy for all players $j \neq i$. Then player i's best-reply under $\Sigma_{\rho''}$ dominates her best-reply under $\Sigma_{\rho'}$ in the increasing convex order. This is an extension of Theorem 1 from the single-agent setting.

2. Holding all else fixed, a higher quality of an opponent's information leads to a more dispersed distribution of best-responses.

   - Suppose $\rho''_j \succeq_{spm} \rho'_j$ for some $j \neq i$, and $\Sigma_{\rho''_k} = \Sigma_{\rho'_k}$ for all $k \neq j$. Fix the monotone strategies of players $k \neq i$. Then player i's best-reply under $\Sigma_{\rho''}$ dominates her best-reply under $\Sigma_{\rho'}$ in the increasing convex order.[18] As player j's information quality increases, $\tilde{s}_j$ becomes more correlated to $\tilde{\theta}_j$, which in turn is (weakly) correlated to $\tilde{\theta}_i$.[19] Thus, by increasing the quality of information for player j, the signals $\tilde{s}_i$ and $\tilde{s}_j$ indirectly become more correlated. Hence, player i can better predict player j's random action and match it.

3. Holding all else fixed, a more dispersed distribution of an opponent's actions leads to a more dispersed distribution of best-responses.

   - Suppose $\Sigma_{\rho''} = \Sigma_{\rho'}$. For some player $j \neq i$, consider two monotone strategies α''_j and α'_j such that α''_j dominates α'_j in the increasing convex order. Fix the monotone strategies of players $k \neq j, i$. Then player i's best-reply to α''_j dominates her best-reply to α'_j in the increasing convex order. It is of similar spirit to the result that strategic complementarities between (a_j, a_i) imply that player i's best-reply is in monotone strategies whenever player j uses a monotone strategy.

4. Finally, we show that the combination of the three aforementioned effects is that each player's distribution of BNE outcomes becomes more dispersed if at least one player gets a higher quality of information.

While the requirements placed on payoff functions may be rather restrictive, we present some examples of applications in which they are satisfied.

Example 1 (Generalized Beauty Contests)

Let $g_i : \Theta \times A_{-i} \to \mathbb{R}$ and $h_i : A_{-i} \to \mathbb{R}$ be bounded and measurable functions, and let

[18] The dominance can be in the weak sense, i.e., it is possible for the best-reply to not change under the two information structures.

[19] By weakly correlated, we mean that we allow for $\tilde{\theta}_i$ to be independent of $\tilde{\theta}_j$.

$\beta_i \in (0,1)$. Let

$$u^i(\theta, a) = -\beta_i \Big(g_i(\theta, a_{-i}) - a_i \Big)^2 - (1-\beta_i) \Big(h_i(a_{-i}) - a_i \Big)^2.$$

Then $u^i \in \Gamma^I$ $\left(resp., u^i \in \Gamma^D\right)$ if $g_i(\theta, a_{-i})$ and $h_i(a_{-i})$ (i) are increasing, (ii) are twice differentiable and convex (resp., concave) in a_j for all $j \neq i$, and (iii) have increasing (resp., decreasing) differences in $(\theta, a_{-j}; a_j)$ for all $j \neq i$.

The example generalizes the canonical beauty contest model (Keynes, 1936; Morris and Shin, 2002) which assumes a normally distributed (common value) state variable, normally distributed signals, and a symmetric payoff with $g_i(\theta, a_{-i}) = \theta$ and $h_i(a_{-i}) = \frac{1}{n-1} \sum_{j \neq i} a_j$.

The assumptions in the canonical model, and more generally, games with linear best-responses (Angeletos and Pavan, 2007; Bergemann and Morris, 2013) make it tractable to compute explicit solutions to the player's optimization problems. From these closed-form solutions, it is straightforward to show that an increase in information leads to a mean-preserving spread of the equilibrium distribution of actions.

Since $u^i \in \Gamma^{\uparrow} \cap \Gamma^D$ in the canonical case, from Theorem 2, we can maintain that more information leads to a mean-preserving spread of the equilibrium distribution of actions without the assumption of normally distributed states and signals. Furthermore, the wider class of payoffs we consider allows us to characterize changes to the equilibrium distribution of actions even when best-responses are no longer linear and explicit solutions are not easily computable.

Example 2 (Joint Projects)
Let $A_i = [0,1]$ for all $i \in N$. Let $v_i : \Theta \to \mathbb{R}$ and $c_i : A_i \to \mathbb{R}$ be bounded and measurable functions. Let

$$u^i(\theta, a) = \prod_{j=1}^{n} a_j v_i(\theta) - c_i(a_i).$$

Then $u^i \in \Gamma^I$ if (i) $v_i(\theta)$ is a non-negative and increasing function, (ii) $c_i(a_i)$ is a convex, increasing, and twice differentiable function, and (iii) $c_i'(a_i)$ is concave in a_i (which is satisfied if the player has quadratic cost).

The example is a variant of the "moral hazard in teams" model (Holmstrom, 1982): each player i exerts effort a_i at cost $c_i(a_i)$. The probability of success is $\prod_{j=1}^{n} a_j$, in which case player i gets a (possibly common-value) payoff $v_i(\theta)$. Each player privately observes a signal about the value of the project before exerting effort.

We can also incorporate an adverse selection component to the example: additionally assume that $\Theta_i = [0,1]$ for all $i \in N$ and $v_i(\theta) = \nu_i \prod_{j=1}^{n} \theta_j$. A player's productivity is given by $\theta_i a_i$ where θ_i represents the player's ability and a_i represents effort. The total probability of success is $\prod_{j=1}^n \theta_j a_j$, in which case player i gets a value of $\nu_i > 0$. Each player privately observes a signal about her productivity before exerting effort.

Example 3 (Network Games with Incomplete Information)
Let $A_i = [0, \bar{a}_i]$ for all $i \in N$. Let $\beta_i : \Theta \to \mathbb{R}$ and $c_i : A_i \to \mathbb{R}$ be bounded and measurable functions. Let $g : \Theta \to \mathbb{R}^{n \times n}$ be the graph of a network with $g_{i,i}(\theta) = 0$ for all $\theta \in \Theta$, i.e., $g(\theta)$ is an $n \times n$ zero-diagonal matrix. Let

$$u^i(\theta, a) = \beta_i(\theta)a_i + \sum_{j=1}^{n} g_{i,j}(\theta)a_i a_j - c_i(a_i).$$

Then $u^i \in \Gamma^I$ if (i) $\beta_i(\theta)$ is an increasing function, (ii) $g_{i,j}(\theta)$ is a non-negative and increasing function for all $j \neq i$, (iii) $c_i(a_i)$ is a convex, increasing, and twice differentiable function, and (iv) $c_i'(a_i)$ is concave in a_i.

A complete information version of this game has been used to study peer effects in social networks (Ballester et al., 2006) as well as monopoly pricing in the presence of network externalities (Candogan et al., 2012).

The example can be used to study peer effects in education: if a student with ability θ_i spends a_i hours studying, she incurs an opportunity cost of $c_i(a_i)$ but improves her educational outcomes (test scores, earnings, etc.) by $\beta_i(\theta_i)a_i$. Holding fixed the number of hours spent studying, the higher the student's ability, the higher her outcome.

Additionally, there are (positive) peer effects between student i and student $j \neq i$ captured by $g_{i,j}(\theta)a_i a_j$. Holding fixed the number of hours spent studying, the higher any student's ability, the more positively the student affects her peers. In particular, if we assume that $g_{i,j}(\theta) = \max\{\theta_i, \theta_j\}$, smart students have a multiplier effect on the rest of their peers. If we instead assume $g_{i,j}(\theta) = \min_{k \in N} \theta_k$, peer effects are only as strong as the weakest student in the class.

Example 4 (Sentiments, Business Cycles and Aggregate Output)
Consider an "island economy" (Lucas Jr, 1972) in which island $i \in \mathcal{I} = [0, 1]$ has an equal probability of being matched with any other island $j \in \mathcal{I}$. After the match, each island first observes some information concerning the island's productivity $\tilde{\theta}_i$, and then trades with its partner. The reduced form of the model is summarized by the best response function

$$y_i = (1 - \alpha)\mathbb{E}_i[\theta_i] + \alpha\mathbb{E}_i[h(y_j, Y)]$$

where y_i is the output in island i and $Y = \int_0^1 y_j dj$ is the aggregate output conditional on all information. We depart from the classical setup by letting the aggregator $h(y_j, Y)$ also depend on Y.

Angeletos and La'O (2013) embed the model above into a dynamic setting to study how business cycles are driven by "sentiment" shocks. Their main innovation is the information structure which captures correlation in beliefs: In each period $t = 1, 2, \ldots$, island i receives signals $x_{i1t} = \theta_{it} + \varepsilon_{i1t}$, $x_{i2t} = x_{j1t} + \varepsilon_{i2t}$, $x_{i3t} = x_{j2t} + \xi_t + \varepsilon_{i3t}$, where $\varepsilon_{i1t}, \varepsilon_{i2t}, \varepsilon_{i3t}$ are idiosyncratic noise terms distributed iid (across islands), Normal with mean 0, and variance $\sigma_1^2, \sigma_2^2, \sigma_3^2$. The sentiment shock ξ_t, which captures the correlation in beliefs, is common to all islands and distributed $N(0, \sigma_\xi^2)$.

If h is increasing and convex in each argument, and has increasing differences in $(y_j; Y)$, then the game corresponds to one of the generalized beauty contests described in Example 1. Increasing the precision $1/\sigma_1$ of signal x_{i1t} will increase the dispersion of output $\{y_{jt}\}_{j \in \mathcal{I}}$ across islands and also leads to a higher level of average output Y_t in each period t.

Furthermore, Angeletos and La'O show that whenever $\sigma_2^2 > 0$ and $\sigma_\xi^2 > 0$, output y_{it} and Y_t vary with the sentiment ξ_t. Therefore, the economy displays business cycles triggered by "exuberant" or "gloomy" beliefs. This amounts to aggregate output $Y_t = Y(x_t, \xi_t) = \int_0^1 y_{it} di$ having more dispersion relative to an economy without x_{i3t}. Interestingly enough, when $h(y_j, Y)$ has increasing differences, the aggregate output has a higher trend, $\bar{Y} = \mathbb{E}(Y_t)$, in the business cycle equilibrium. In particular, business cycles might shift the trend of output upwards and therefore allow for higher average investment and capital accumulation.

4 Applications

We consider two application of our main result in the single-agent setting, and two applications of our result in Bayesian games.

4.1 Application: Pigouvian Subsidies and Monopoly Production

In the example from Section 1.1, we considered the effect of information quality on a monopolist's production decision in a highly stylized example. In this subsection, we consider the example in a more general setting as follows: a monopolist who produces $q \in [0, \bar{q}]$ faces a downward sloping inverse demand curve $P(q)$ and a cost function $c(\theta, q)$ where the parameter $\theta \in \Theta$ is unknown. The monopolist holds a prior $\mu^0 \in \Delta(\Theta)$. As θ increases, the marginal cost declines, i.e. $-c(\theta, q)$ has increasing differences in $(\theta; q)$. We assume that the monopolist's profit $\pi(\theta, q) = qP(q) - c(\theta, q)$ is strictly concave in q and admits an interior solution for each $\theta \in \Theta$.

Prior to making any production decisions, the monopolist can acquire information from $\mathcal{P}$, a set of information structures that satisfy (A.5). For any $\Sigma_{\rho''}, \Sigma_{\rho'} \in \mathcal{P}$, either $\rho'' \succeq_{spm} \rho'$ or vice versa. Let $\kappa : \mathcal{P} \to \mathbb{R}$ be the cost of acquiring information with $\kappa(\rho'') \geq \kappa(\rho')$ when $\rho'' \succeq_{spm} \rho'$.

Consider a social planner who is unable to regulate prices or quantities. *Under what conditions does the social planner demand more information than the monopolist?*[20]

Let $q^M(s; \rho)$ be the optimal quantity the monopolist produces when she observes a signal realization $s \in S$ from an information structure $\Sigma_\rho \in \mathcal{P}$. The monopolist's ex-ante problem is to choose an information structure that maximizes

$$\int_{\Theta \times S} \pi\big(\theta, q^M(s; \rho)\big) dF(\theta, s; \rho) - \kappa(\rho).$$

In contrast, the social planner takes the consumer surplus into account. Let $CS(q)$ be the consumer surplus when the monopolist produces q. The planner's ex-ante payoff is given by

$$\int_{\Theta \times S} \pi\big(\theta, q^M(s; \rho)\big) dF(\theta, s; \rho) + \int_S CS\big(q^M(s; \rho)\big) dF_S(s) - \kappa(\rho).$$

Thus, the planner has a higher demand for information than the monopolist when a higher quality of information increases the expected consumer surplus, i.e., when information is a positive externality on the consumers.

Proposition 2 *Let $-qP''(q)/P'(q) \leq 1$, and let the profit function $\pi \in \mathcal{U}^I$. Then the social planner has a* higher *demand for information than the monopolist.*

Intuitively, $-qP''(q)/P'(q) \leq 1$ implies that as the quantity produced increases, the consumers capture more and more of the welfare gains than does the monopolist. Therefore, the consumer surplus is a convex function of the quantity which in turn implies that social planner is "more risk-loving" than the monopolist, i.e., consumers (and the planner) benefit when the monopolist becomes more responsive with a higher mean as quality of information increases. From Theorem 1, we get the desired responsiveness behavior when $\pi \in \mathcal{U}^I$.

[20] Athey and Levin (2017) consider a similar problem. However, in their application, the planner can regulate prices/quantities as well as the quality of information.

4.2 Application: Information Disclosure

In the Bayesian persuasion game of Kamenica and Gentzkow (2011), a sender (he) has full flexibility in what information to disclose to a receiver (she) in order to persuade the receiver to take an action that is desirable to the sender. Kamenica and Gentzkow provide a tool to solve the sender's problem: first, characterize the sender's interim value as a function of the receiver's posterior belief, and then take the concave closure of the sender's interim value function.

However, the concavification approach requires a closed form solution to the receiver's optimization strategy. Usually, this is only possible when the set of actions is finite or when the optimal strategy of the receiver is just a function of the posterior mean.

In order to consider a richer set of preferences, we restrict the sender to choose information structures that can be ranked by the supermodular order. Using the comparative statics of Theorem 1, we then characterize under what conditions the sender prefers to disclose a higher/lower quality of information.

Let the sender's payoff be given by $v : \Theta \times A \to \mathbb{R}$ which is continuous in a for all $\theta \in \Theta$. The receiver's payoff is given by $u : \Theta \times A \to \mathbb{R}$ which satisfies (A.1)-(A.4). Unlike the canonical persuasion problem, we assume that the sender is restricted to $\mathcal{P}$,a set of information structures that satisfy (A.5). Additionally, for any $\Sigma_{\rho''}, \Sigma_{\rho'} \in \mathcal{P}$ either $\rho'' \succeq_{spm} \rho'$ or vice versa.

The sender's problem is given by

$$\max_{\Sigma_\rho \in \mathcal{P}} \; V(\rho) = \int_{\Theta \times S} v(\theta, a(s,\rho)) dF(\theta, s; \rho) \qquad s.t.$$
$$a(s;\rho) = \arg\max_{a \in A} \int_{\Theta} u\big(\theta, a\big)\mu(d\theta | s; \rho) \quad \forall \Sigma_\rho \in \mathcal{P}, \forall s \in S.$$

Proposition 3 *Assume $v(\theta, a)$ satisfies increasing differences (resp., decreasing differences) in $(\theta; a)$, and suppose one of the following holds:*

i. $u \in \mathcal{U}^I$ and $v(\theta, a)$ is increasing and convex (resp., decreasing and concave) in a,

ii. $u \in \mathcal{U}^D$ and $v(\theta, a)$ is decreasing and convex (resp., increasing and concave) in a, or

iii. $u \in \mathcal{U}^I \cap \mathcal{U}^D$ and $v(\theta, a)$ is convex (resp., concave) in a.

For information structures $\Sigma_{\rho''}, \Sigma_{\rho'} \in \mathcal{P}$, $V(\rho'') \geq V(\rho')$ $\Big($resp., $V(\rho'') \leq V(\rho')\Big)$ if $\rho'' \succeq_{spm} \rho'$.

To see the value in Proposition 3, consider the Rothschild and Stiglitz (1971) portfolio choice problem. There are two assets: money that yields a zero rate of return and a risky asset that yields a random rate of return of $\tilde{x}$. The random return on the risky asset is drawn from a support in $\underline{x} < 0 < \bar{x}$ according to an absolutely continuous distribution function G_θ with denisty g_θ. Changes in the parameter θ capture changes to the underlying "riskiness" of the risky asset. Suppose for $\theta'' > \theta'$,

$$\int_{\underline{x}}^{z} x\big[dG_{\theta''}(x) - dG_{\theta'}(x)\big] \geq 0, \forall z \in [\underline{x}, \bar{x}] \tag{RS}$$

with equality when $z = \bar{x}$. Rothschild and Stiglitz show that all risk-averse agents invest more in a risky asset distributed according to $G_{\theta''}$ than $G_{\theta'}$ if, and only if, (RS) holds.

We alter their model into a portfolio management problem between a risk-neutral financial adviser (the sender) and a risk-averse investor (the receiver) with a Bernoulli utility $\vartheta : \mathbb{R} \to \mathbb{R}$ which is continuous, strictly increasing, and strictly concave. Suppose the financial adviser gets a share $\pi \in (0, 1)$ of the return on the risky asset (e.g., money management fees). Hence, if the investor places a fraction $a \in [0, 1]$ of her wealth $W > 0$ in the risky asset, her ex-post payoff is

$$u(\theta, a) = \int_{\underline{x}}^{\bar{x}} \vartheta\Big((1-a)W + aW\big(1 + x(1-\pi)\big)\Big) dG_\theta(x),$$

whereas the financial adviser's ex-post payoff is given by

$$v(\theta, a) = aW\pi \int_{\underline{x}}^{\bar{x}} x dG_\theta(x).$$

Ex-ante, the value of θ is unknown, and both the sender and receiver hold a common prior $\mu^0 \in \Delta(\Theta)$. The financial adviser chooses what information to disclose to the investor in order to influence how much is invested in the risky asset. *When is the financial adviser better off disclosing more information about the risky asset?*

The investor's optimal strategy is not characterized by a cutoff in her posterior beliefs, and it depends on higher moments of her posterior (not just the posterior mean). Thus, the example does not fit the simplifying assumptions often made in the persuasion literature.

Nonetheless, in our portfolio management example, (RS) implies that $u(\theta, a)$ has increasing differences in $(\theta; a)$, and that the financial adviser has a payoff $v(\theta, a)$ which is state-independent, linear, and increasing in a. We can readily apply Proposition 3 and conclude that the financial adviser prefers to provide the investor a higher (resp., lower) quality of information if $u \in \mathcal{U}^I$ (resp., $u \in \mathcal{U}^D$). For instance, when the investor's Bernoulli utility satisfies the relative prudence condition[21]

$$-\frac{\vartheta'''(x)}{\vartheta''(x)}x \geq 1,$$

it is straightforward to show that $u \in \mathcal{U}^I$ (using the second mean value theorem). Thus, the financial adviser prefers to disclose all information to the investor.

Corollary 2 *Let $|supp(\mu^0)| = 2$. Suppose the conditions from Proposition 3 on $v(\theta, a)$ and $u(\theta, a)$ hold. Then, full information revelation (resp., no information) is the optimal persuasion policy.*

Remark 3 *When the sender has full flexibility, he may prefer some other information structure that does not satisfy (A.5), and hence, cannot be ordered by the supermodular order. However, when there are only two relevant state variables, i.e., $|supp(\mu^0)| = 2$, (A.5) is trivially satisfied. Hence, Corollary 2 is a characterization of the optimal persuasion policy over all unrestricted information structures.*

4.3 Application: Information Sharing in Supermodular Games

Consider a two-player common value Bayesian game with $\tilde{\theta}_1 = \tilde{\theta}_2 = \tilde{\theta}$. The basic game is given by $G \triangleq (\{A_i, u^i\}_{i=1,2}, \mu^0)$ where the payoff $u^i : \Theta \times A \to \mathbb{R}$ for $i = 1, 2$ satisfies (A.7)-(A.10), and the common prior $\mu^0 \in \Delta(\Theta)$ trivially satisfies (A.6).

Prior to playing the basic game, each player i observes a signal from an information structure $\Sigma_{\rho_i} \in \mathcal{P}_i$, where $\mathcal{P}_i$ denotes the set of information structures. We assume that each $\Sigma_\rho \triangleq (\Sigma_{\rho_1}, \Sigma_{\rho_2}) \in \mathcal{P}_1 \times \mathcal{P}_2$ satisfies (A.11)-(A.13). Furthermore, for any $i = 1, 2$ and any two information structures $\Sigma_{\rho''_i}$, $\Sigma_{\rho'_i} \in \mathcal{P}_i$, either $\rho''_i \succeq_{spm} \rho'_i$ or vice versa. Let $\Sigma_{\bar{\rho}_i} \in \mathcal{P}_i$ represent the full-information structure, i.e., an information structure that perfectly correlates the signal and the state.

Suppose player 1 is exogenously endowed with $\Sigma_{\rho_1} = \Sigma_{\bar{\rho}_1}$, i.e., player 1 observes the realization of $\tilde{\theta}$. In contrast, player 2 does not observe an exogenous signal. Instead, player 1 chooses an information structure $\Sigma_{\rho_2} \in \mathcal{P}_2$ for player 2. In other words, prior to the learning the state, player 1 commits to how much information she will share with player 2 by choosing a "disclosure" policy.[22] Each choice of Σ_{ρ_2} defines a Bayesian game $\mathcal{G}_\rho \triangleq (\Sigma_{\bar{\rho}_1}, \Sigma_{\rho_2}, G)$ as outlined in Figure 5.

[21]See Kimball (1990) for an analysis of the relative prudence coefficient and its effect on precautionary savings.

[22]Another interpretation is player 1 plays the role of the "sender" and player 2 plays the role of the "receiver" in a Bayesian persuasion game as in subsection 4.2. The only difference here is that both the sender and the receiver take an action.

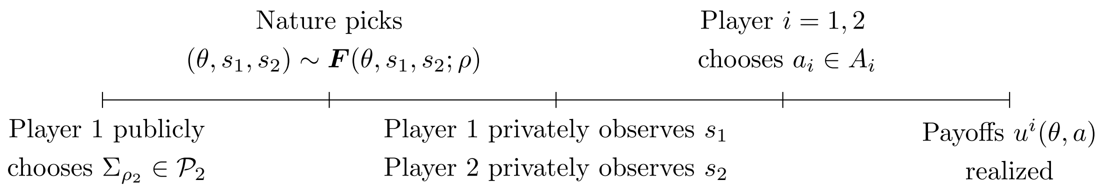


Figure 5: Timing of information sharing game

For each Bayesian game $\mathcal{G}_\rho$, we assume that the players can coordinate on the maximal monotone BNE $a^\star(\rho) = (a_1^\star(\rho), a_2^\star(\rho))$ with $a_i^\star(\cdot; \rho) : S_i \to A_i$. Since $\tilde{s}_1$ is perfectly correlated to $\tilde{\theta}$, with some abuse of notation, player 1's BNE payoff is given by

$$U_1(\rho) = \int_{\Theta \times S_2} u^1\big(\theta, a_1^\star(\theta; \rho), a_2^\star(s_2; \rho)\big) dF(\theta, s_2; \rho_2).$$

How much information, if any, would player 1 want to share with player 2? Such a question about information sharing in Bayesian games has been explored within the context of firm competition starting with Novshek and Sonnenschein (1982), Clarke (1983), Vives (1984), Gal-Or (1985), and Raith (1996). The literature overall shows that full information disclosure is optimal for the case of firm competition with strategic complements (e.g., differentiated Bertrand competition). More recently, Bergemann and Morris (2013) provide a comprehensive analysis of information sharing in beauty contests, and similarly show full information disclosure is optimal when strategic complementarities exist between players.

However, the previous literature has focused on linear-quadratic games and normally distributed states and signals. In this application, we instead use the comparative statics developed in Theorem 2 to extend the optimality of full information disclosure beyond games with linear best-responses.

Proposition 4 *Suppose $u^1 : \Theta \times A \to \mathbb{R}$ satisfies increasing differences in $(\theta, a_1; a_2)$, and one of the following holds:*

i. $u^i \in \Gamma^I$ for $i = 1, 2$ and u^1 is increasing and convex in a_2,

ii. $u^i \in \Gamma^D$ for $i = 1, 2$ and u^1 is decreasing and convex in a_2, or

iii. $u^i \in \Gamma^I \cap \Gamma^D$ for $i = 1, 2$ and u^1 is convex in a_2.

Then, it is optimal for player 1 to choose $\Sigma_{\rho_2} = \Sigma_{\bar{\rho}_2}$.

The joint project game from Example 2 and the network game in Example 3 satisfy the first sufficient condition, and the standard differentiated Bertrand competition model with linear demand (e.g., Raith (1996)) satisfies the third sufficient condition.[23] Hence, in all three cases, we can readily apply Proposition 4 to conclude that it is ex-ante optimal for player 1 to fully share her information with player 2.

It is worth noting that, even with the generalization from the linear-quadratic games, the application in this subsection is a special case of the standard information sharing model. Player 1 observes everything about the (common value) state while player 2 does not. Thus, only player 1 is in a position to share information. In Leal-Vizcaíno and Mekonnen (2018), we generalize the result in Proposition 4 to a setting in which each player receives an exogenous signal and decides how much information to share with her opponent.[24] We establish that the full-information sharing result is robust to different specifications of information structures and payoffs. Moreover, we also show that it is a dominant strategy and, therefore, the unique Nash outcome of the information sharing game.

4.4 Information Acquisition and the Value of Transparency

Oligopolists are affected by many variables they cannot observe or estimate precisely: their own cost function, the cost function of their rivals, the demand in a particular market on a given date, etc. To the extent that these pieces of information are private and subject to learning, we must envision the process of gathering information as a game of information acquisition.

Just as fixed costs or increasing returns might generate an imperfectly competitive market structure by limiting entry, superior information by an incumbent firm might also constitute a barrier to entry. In principle, the case of information is not different to the classical treatment of capital or capacity investment when studying entry, accommodation and exit in oligopolistic markets. However, we illustrate how investing in information differs from other types of investment, such as capacity, learning by doing, advertising, etc. (Bulow et al., 1985).

[23] Linear differentiated Bertrand: for each player $i \in N$, profit function is given by

$$u^i(\theta, a) = (a_i - c_i)\left(\alpha_i(\theta) + \sum_{j \neq i} \beta_{ij} a_j - \beta_{ii} a_i\right),$$

where a is the price vector, $\alpha_i(\theta)$ is a demand shifter with $\alpha_i'(\cdot) \geq 0$, $c_i > 0$ is the marginal cost, and $\beta_{ij} \geq 0 > \beta_{ii}$ $\forall j \neq i$.

[24] The generalization requires a stronger order over information structures. Therefore, the results in Leal-Vizcaíno and Mekonnen (2018) are not an immediate application of Theorem 2.

We focus on the analysis of entry-accommodation,[25] and decompose the impact of information acquisition on the incumbent's profits into two effects: a direct effect (which is always non-negative (Blackwell, 1951, 1953)) from improving the incumbent's decision making, and an indirect effect (which can be positive or negative) stemming from the response of the entrant adjusting her strategy to the incumbent's information. We call the indirect effect *the value of transparency* and we show that it is positive or negative depending on (i) the responsiveness of the entrant to changes in the incumbent's information quality, and (ii) the sign of the externality imposed on the incumbent by the entrant's responsiveness.

The analysis of entry accommodation and the value of transparency is formally equivalent to characterizing the demand for information in overt and covert information acquisition games—the difference in the value of information in these two games is precisely the value of transparency. Understanding what drives the difference between the overt and covert demands for information is of independent interest to theorists studying the value of information, who more often than not restrict attention to one of the two games (covert or overt) for technical simplicity.

4.4.1 Setup

We consider a two-player Bayesian game composed of two stages: an information acquisition stage followed by a basic game $G \triangleq (\{A_i, u^i\}_{i=1,2}, \mu^0)$ where the payoff $u^i : \Theta \times A \to \mathbb{R}$ for $i = 1, 2$ satisfies (A.7)-(A.10) and the common prior $\mu^0 \in \Delta(\Theta)$ satisfies (A.6).

In the information acquisition stage, player 2 has an exogenously given information structure Σ_{ρ_2}. On the other hand, player 1 is allowed to choose an information structure from a set $\mathcal{P}_1$ such that for any $\Sigma_{\rho_1} \in \mathcal{P}_1$, $\Sigma_\rho \triangleq (\Sigma_{\rho_1}, \Sigma_{\rho_2})$ satisfies (A.11)-(A.13). Additionally, we assume that for any two information structures $\Sigma_{\rho_1''}, \Sigma_{\rho_1'} \in \mathcal{P}_1$, either $\rho_1'' \succeq_{spm} \rho_1'$ or vice versa. Let $\kappa : \mathcal{P}_1 \to \mathbb{R}$ be the cost of acquiring information with $\kappa(\rho_1'') \geq \kappa(\rho_1')$ when $\rho_1'' \succeq_{spm} \rho_1'$.

Throughout this section, we only consider information acquisition in pure strategies in the first stage.[26] We also assume that players coordinate on the maximal pure-strategy monotone BNE in the second stage.

To better understand the difference between overt and covert information acquisition, suppose initially that player 1 is endowed with information structure $\Sigma_{\rho_1'}$ and this is common

[25] In the face of an entry threat three kinds of behavior by the incumbent will be possible: entry might be blockaded, deterred or accommodated. See Tirole (1988) textbook.

[26] For overt information acquisition, this is without loss as player 2 observes the chosen information structure before the second stage. Hence, player 1 randomizes only when she is indifferent.

knowledge, i.e., both players know the Bayesian game is $\mathcal{G}_{\rho'} \triangleq (\Sigma_{\rho'_1}, \Sigma_{\rho_2}, G)$. Let $\big(a_1^\star(\rho'), a_2^\star(\rho')\big)$ be the resulting BNE of $\mathcal{G}_{\rho'}$. Consider the following two scenarios as a thought experiment.

In the first scenario, player 1 is allowed to either keep $\Sigma_{\rho'_1}$ or switch to $\Sigma_{\rho''_1}$. Player 2 observes whether or not player 1 switches. This scenario mirrors the the overt information acquisition game. If player 1 switches to $\Sigma_{\rho''_1}$, the game changes from $\mathcal{G}_{\rho'}$ to $\mathcal{G}_{\rho''} \triangleq (\Sigma_{\rho''_1}, \Sigma_{\rho_2}, G)$ and the resulting BNE is $\big(a_1^\star(\rho''), a_2^\star(\rho'')\big)$.

In the second scenario, player 1 can again switch to $\Sigma_{\rho''_1}$ but player 2 is neither aware that player 1 can switch nor observes player 1's choice. This scenario mirrors the covert information acquisition game. If player 1 switches, player 2 will naively believe that the game is still $\mathcal{G}_{\rho'}$ and continues to play $a_2^\star(\rho')$. On the other hand, player 1 best-replies to $a_2^\star(\rho')$ by playing the strategy $a_1^{BR}(a_2^\star(\rho'), \rho'')$.

Since we wish to distinguish between player 1's choice of information and player 2's beliefs, we denote the actual outcome of the information acquisition stage by $\rho = (\rho_1, \rho_2)$ and player 2's belief of the outcome of the information acquisition stage by $\hat{\rho} = (\hat{\rho}_1, \rho_2)$. We say player 2 has correct beliefs when $\hat{\rho}_1 = \rho_1$ (which must be the case in any equilibrium).

Given actual first stage outcome ρ and player 2's belief $\hat{\rho}$, let player 1's ex-ante payoff in the covert game (second scenario) be $U_1(\rho; \hat{\rho}) - \kappa(\rho_1)$ where

$$U_1(\rho; \hat{\rho}) = \int_{\Theta \times S} u^1\big(\theta, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2; \hat{\rho})\big) d\boldsymbol{F}(\theta, s; \rho).$$

In the overt game (first scenario), player 2 has correct beliefs. Hence, given actual first stage outcome ρ, player 1's payoff in the overt game is $U_1(\rho; \rho) - \kappa(\rho_1)$ with

$$\begin{aligned} U_1(\rho; \rho) &= \int_{\Theta \times S} u^1\big(\theta, a_1^{BR}(s_1; a_2^\star(\rho), \rho), a_2^\star(s_2; \rho)\big) d\boldsymbol{F}(\theta, s; \rho) \\ &= \int_{\Theta \times S} u^1\big(\theta, a_1^\star(s_1; \rho), a_2^\star(s_2; \rho)\big) d\boldsymbol{F}(\theta, s; \rho), \end{aligned}$$

where the equality follows from $a_1^{BR}(a_2^\star(\rho), \rho) = a_1^\star(\rho)$ by the definition of a BNE.

Definition 5 *Given actual first stage outcome ρ and player 2's belief $\hat{\rho}$, the value of transparency is given by:*

$$VT(\rho; \hat{\rho}) = U_1(\rho; \rho) - U_1(\rho; \hat{\rho}).$$

In other words, $VT(\rho; \hat{\rho})$ represents the gain/loss to player 1 from disclosing to player 2 her actual first stage choice, Σ_{ρ_1}, instead of letting player 2 incorrectly believe that the first stage choice is $\Sigma_{\hat{\rho}_1}$. The value of transparency does not capture any direct substantive advantages of information; player 1's chosen information structure in both cases is Σ_{ρ_1}. Instead, it captures the indirect effects of information stemming from a change in player 2's beliefs and, therefore, her strategic response.[27]

4.4.2 Value and Demand for Information

Before we discuss how to characterize the value of transparency, we present why it is an interesting economic concept. In particular, we show that the value of transparency is helpful in answering the following questions: *When is a higher quality of costless but overt information acquisition always beneficial to player 1? Does player 1 acquire more information when information acquisition is overt or when it is covert?*

In covert games, information only has a direct effect, i.e., more information allows player 1 to make better decisions in the second stage. Therefore, the value of costless information is never negative (Neyman, 1991).

While information has the same beneficial direct effect in overt games, there are also strategic effects; player 2 observes how much information player 1 acquires, and responds to it in the second stage. If player 2 finds it optimal to choose an unfavorable action (punish player 1) in the equilibrium of the second stage whenever player 1 acquires more information, then the value of information in overt games may be negative (Kamien et al., 1990). Nonetheless, we show that the value of *overt* information cannot be negative if player 1 benefits from disclosing to player 2 that a higher quality of information has been acquired.

Proposition 5 *Let κ be a constant function. For any two information structures $\Sigma_{\rho_1}, \Sigma_{\hat{\rho}_1} \in \mathcal{P}_1$, suppose $\rho_1 \succeq_{spm} \hat{\rho}_1$ implies $VT(\rho;\hat{\rho}) \geq 0$. Then, $U_1(\rho;\rho) \geq U_1(\hat{\rho};\hat{\rho})$.*

Proof. For two information structures $\Sigma_{\rho_1}, \Sigma_{\hat{\rho}_1} \in \mathcal{P}_1$, we can write

$$U_1(\rho;\rho) - U_1(\hat{\rho};\hat{\rho}) = \underbrace{U_1(\rho;\rho) - U_1(\rho;\hat{\rho})}_{=VT(\rho;\hat{\rho})} + \underbrace{U_1(\rho;\hat{\rho}) - U_1(\hat{\rho};\hat{\rho})}_{\text{value of covert information}}.$$

Amir and Lazzati (2016) (Proposition 7) show that the second term is non-negative when $\rho_1 \succeq_{spm} \hat{\rho}_1$, i.e., the value of covert information is non-negative when quality of information

[27] Our treatment of the value of transparency is loosely connected to the *expectations conformity* conditions in Tirole (2015). Expectations conformity implies that player 1 is more willing to acquire Σ_{ρ_1} over $\Sigma_{\hat{\rho}_1}$ when player 2 believes that player 1 will acquire Σ_{ρ_1}. It is straightforward to show that expectations conformity is equivalent to $VT(\rho;\hat{\rho}) + VT(\hat{\rho};\rho) \geq 0$.

increases. Hence, if $VT(\rho;\hat{\rho}) \geq 0$, we can conclude that the value of overt information is also non-negative when quality of information increases. ■

To answer the second question about the demand of information, let $\Sigma_{\rho_1^c}$ and $\Sigma_{\rho_1^o}$ denote the information structures acquired in a pure strategy Nash equilibrium (PSNE) of covert and overt games.[28] Specifically, $\Sigma_{\rho_1^c}$ is a solution to

$$\max_{\Sigma_{\rho_1}\in\mathcal{P}} U_1(\rho;\rho^c) - \kappa(\rho_1).$$

In other words, given player 2 believes player 1 chooses $\Sigma_{\rho_1^c}$ in equilibrium, it is indeed optimal for player 1 to choose $\Sigma_{\rho_1^c}$. In contrast, $\Sigma_{\rho_1^o}$ solves

$$\max_{\Sigma_{\rho_1}\in\mathcal{P}} U_1(\rho;\rho) - \kappa(\rho_1).$$

In other words, $\Sigma_{\rho_1^o}$ is optimal for player 1 after taking into account that player 2 will observe the chosen information structure in the first stage and will respond to it in the second stage. We show that whenever the value of transparency is non-negative, player 1 acquires more information in overt games than in covert games, regardless of the cost function.

Proposition 6 *For any two information structures $\Sigma_{\rho_1}, \Sigma_{\hat{\rho}_1} \in \mathcal{P}_1$, let $VT(\rho;\hat{\rho}) \geq 0$ if, and only if, $\rho_1 \succeq_{spm} \hat{\rho}_1$. Then, $\rho_1^o \succeq_{spm} \rho_1^c$.*

Proof. Suppose $\Sigma_{\rho_1^c} \neq \Sigma_{\rho_1^o}$ (otherwise, it is trivial).[29] By definition,

$$U_1(\rho^c;\rho^c) - \kappa(\rho_1^c) \geq U_1(\rho^o;\rho^c) - \kappa(\rho_1^o)$$
$$U_1(\rho^o;\rho^o) - \kappa(\rho_1^o) \geq U_1(\rho^c;\rho^c) - \kappa(\rho_1^c).$$

Combining the two inequalities, we get $U_1(\rho^o;\rho^o) - U_1(\rho^o;\rho^c) = VT(\rho^o;\rho^c) \geq 0 \Leftrightarrow \rho_1^o \succeq_{spm} \rho_1^c$.
■

[28] We have made an implicit assumption that a PSNE exists in the covert information acquisition game. Establishing such an equilibrium exists is beyond the scope of this section. However, when κ is a constant function, $U_1(\rho;\hat{\rho}) - \kappa(\rho_1)$ satisfies single crossing in $(\rho_1;\hat{\rho}_1)$, i.e., given $\rho_1'' \succeq_{spm} \rho_1'$ and $\hat{\rho}_1'' \succeq_{spm} \hat{\rho}_1'$, $U_1(\rho'';\hat{\rho}') - \kappa(\rho_1'') \geq U_1(\rho';\hat{\rho}') - \kappa(\rho_1') \implies U_1(\rho'';\hat{\rho}'') - \kappa(\rho_1'') \geq U_1(\rho';\hat{\rho}'') - \kappa(\rho_1')$. Then, with appropriate assumptions on $\mathcal{P}$, we can use Milgrom and Shannon (1994) and Athey (2001) to establish existence of PSNE of the covert game.

[29] The implicit assumption of unique equilibrium outcomes in the result above is only made to simplify exposition. The antecedent of Proposition 6 implies $VT(\hat{\rho};\hat{\rho}) = 0$ and $VT(\rho;\hat{\rho}) \geq 0$ for any $\rho_1 \succeq_{spm} \hat{\rho}_1$. We can therefore apply familiar monotone comparative statics tools for single-crossing functions to show that the solution set for overt equilibrium maximization problem dominates the solution set for covert equilibrium.

4.4.3 Characterizing the Value of Transparency

We now characterize the value of transparency which depends on the responsiveness of player 2 and the externality player 2's responsiveness imposes on player 1.

Theorem 3 *Suppose either the basic game G is one of independent private values, or $u^1(\theta, a)$ has increasing differences in $(\theta, a_1; a_2)$. Additionally, suppose one of the following holds:*

i. $u^i \in \Gamma^I$ *for* $i = 1, 2$ *and* u^1 *is increasing and convex in* a_2,

ii. $u^i \in \Gamma^D$ *for* $i = 1, 2$ *and* u^1 *is decreasing and convex in* a_2, *or*

iii. $u^i \in \Gamma^I \cap \Gamma^D$ *for* $i = 1, 2$ *and* u^1 *is convex in* a_2.

Then for any two information structures $\Sigma_{\rho_1}, \Sigma_{\hat{\rho}_1} \in \mathcal{P}_1$, $VT(\rho; \hat{\rho}) \geq 0$ *if, and only if,* $\rho_1 \succeq_{spm} \hat{\rho}_1$.

The joint project game in Example 2, the network game in Example 3, and the standard differentiated Bertrand models (Raith, 1996) all satisfy the conditions of Theorem 3. Hence, applying Proposition 6, we can conclude that the demand for information in these examples is higher when information acquisition is overt.

To gain some intuition, recall that $VT(\rho; \hat{\rho}) = U_1(\rho; \rho) - U_1(\rho; \hat{\rho})$ is given by

$$\int_{\Theta \times S} \Big[u^1\big(\theta, a_1^\star(s_1; \rho), a_2^\star(s_2; \rho)\big) - u^1\big(\theta, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2; \hat{\rho})\big) \Big] d\boldsymbol{F}(\theta, s; \rho).$$

Consider the case of independent private values, and let $S_2 = [0, 1]$. By taking a first-order Taylor expansion, we can approximate the value of transparency as

$$\approx \underbrace{\int_{\Theta \times S} u^1_{a_1}\big(\theta_1, a_1^\star(s_1; \rho), a_2^\star(s_2; \rho)\big)\big(a_1^\star(s_1; \rho) - a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho)\big) d\boldsymbol{F}(\theta, s; \rho)}_{=0 \text{ by optimality in second stage}}$$

$$+ \int_0^1 \underbrace{\left[\int_{\Theta_1 \times S_1} u^1_{a_2}\big(\theta_1, a_1^\star(s_1; \rho), a_2^\star(s_2; \rho)\big) dF(\theta_1, s_1; \rho_1)\right]}_{\triangleq \zeta(s_2)} \Big(a_2^\star(s_2; \rho) - a_2^\star(s_2; \hat{\rho})\Big) dF_{S_2}(s_2).$$

The conditions in Theorem 3 connect the sign for the value of transparency to player 2's responsiveness, $a_2^\star(\rho) - a_2^\star(\hat{\rho})$, the type of externality player 2's action imposes on player 1, $\text{sign}(u^1_{a_2})$, and player 1's "risk" attitude towards player 2's action, $\text{sign}(u^1_{a_2 a_2})$.

For example, suppose condition *i.* of Theorem 3 holds. As $u^1(\theta_1, a)$ is increasing and convex in a_2, $\zeta(s_2)$ is non-negative and increasing in s_2. Additionally, from Theorem 2, $u^i \in \Gamma^I$ for

$i = 1, 2$ implies that player 2 becomes more responsive with a higher mean as the quality of player 1's information increases. From Lemma 1,

$$\rho_1 \succeq_{spm} \hat{\rho}_1 \implies \int_t^1 \Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big) dF_{S_2}(s_2) \geq 0$$

for all $t \in [0, 1]$. From the second mean value theorem, there exists some $t^* \in [0, 1]$ such that

$$\begin{aligned} VT(\rho;\hat{\rho}) \approx & \int_0^1 \zeta(s_2)\Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big) dF_{S_2}(s_2) \\ =& \zeta(1) \int_{t^*}^1 \Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big) dF_{S_2}(s_2) \geq 0. \end{aligned}$$

For the independent private values case, Theorem 3 can be generalized into the taxonomy provided in Figure 6. The first two columns describe how player 2 responds when the information structure changes from $\Sigma_{\hat{\rho}}$ to Σ_ρ. The next two columns are assumptions placed on player 1's utility function. The last column presents the resulting sign on the value of transparency. The first, third, and fifth rows of Figure 6 correspond to condition i, ii, and iii of Theorem 3 respectively. For instance, the fifth row of Figure 6 states that if a change from $\Sigma_{\hat{\rho}_1}$ to Σ_{ρ_1} leads to a mean-preserving spread in player 2's actions (cst stands for constant mean), and if player 1's utility is convex in a_2 (without any more restrictions on $\text{sign}(u^1_{a_2})$), then the value of transparency $VT(\rho;\hat{\rho})$ is non-negative.

4.4.4 Relation to Strategic Effects of Investment in Firm Competition

The characterization of the value of transparency is related to the taxonomy of strategic behavior in firm competition studied by Fudenberg and Tirole (1984), and Bulow et al. (1985).[30] Here we follow the textbook treatment of Tirole (1988) and only consider the case of entry accommodation in a duopoly under complete information.

There are two periods and two firms, an incumbent (firm 1) and an entrant (firm 2). In the first period, the incumbent chooses a level of investment $K_1 \in \mathbb{R}$, which the entrant observes. The term investment is used in a broad sense and can represent, for example, investment in R&D that lowers the incumbent's marginal costs or advertising that captures a share of the market.

In the second period, both firms compete either in quantities (strategic substitutes) or

[30] For a thorough treatment of different examples and applications, we recommend Shapiro (1989). For a more recent treatment using the tools of supermodular games, see Vives (2001).

$a_2(\rho) - a_2(\hat{\rho})$		Externality		Transparency
responsiveness	mean	sign$(u^1_{a_2})$	sign$(u^1_{a_2 a_2})$	$VT(\rho;\hat{\rho})$
$\nearrow$	$\nearrow$	$+$	$+$	$+$
$\nearrow$	$\nearrow$	$-$	$-$	$-$
$\nearrow$	$\searrow$	$-$	$+$	$+$
$\nearrow$	$\searrow$	$+$	$-$	$-$
$\nearrow$	cst	$\cdot$	$+$	$+$
$\nearrow$	cst	$\cdot$	$-$	$-$
$\searrow$	$\nearrow$	$+$	$-$	$+$
$\searrow$	$\nearrow$	$-$	$+$	$-$
$\searrow$	$\searrow$	$-$	$-$	$+$
$\searrow$	$\searrow$	$+$	$+$	$-$
$\searrow$	cst	$\cdot$	$-$	$+$
$\searrow$	cst	$\cdot$	$+$	$-$

Figure 6: A taxonomy of the value of transparency for independently private values.

prices (strategic complements). Let $\big(a_1^\star(K_1), a_2^\star(K_1)\big)$ be the resulting Nash equilibrium of the second period after the incumbent chose K_1 in the first period. The incumbent's payoff from choosing an investment level K_1 is given by $U_1\big(K_1, a_1^\star(K_1), a_2^\star(K_1)\big)$.

Fudenberg and Tirole (1984) show that the total marginal effect on the incumbent's payoff from increasing investment can be decomposed into

$$\frac{dU_1}{dK_1} = \underbrace{\underbrace{\frac{\partial U_1}{\partial K_1}}_{\text{direct effect}} + \underbrace{\frac{\partial U_1}{\partial a_1}\frac{da_1^\star}{dK_1}}_{\substack{=0 \\ \text{by Envelope theorem}}}}_{\text{value of ``covert'' investment}} + \underbrace{\frac{\partial U_1}{\partial a_2}\frac{da_2^\star}{dK_1}}_{\text{strategic effect}} .$$

Increasing the level of investment has a direct effect on the incumbent's payoff, for example, by reducing the marginal cost. It also affects the incumbent's optimal action choice in the second period, captured by $\frac{da_1^\star}{dK_1}$. If the entrant was unable to observe the incumbent's investment choice, these would be the only marginal effects to account for when the incumbent increases investment.

However, since the entrant observes the incumbent's first period choice of K_1, the investment also has strategic effects; the entrant's production/pricing decision is indirectly affected by K_1. This strategic effect depends on the entrant's equilibrium response to an increase in the level of investment, represented by $\frac{da_2^\star}{dK_1}$, and on the externality the entrant's actions impose on the incumbent's payoff, represented by $\frac{\partial U^1}{\partial a_2}$.

In our model, the game is one of incomplete information, player 1 is the incumbent, player 2 is the entrant, and the investment level K_1 corresponds to the quality of the player 1's information structure ρ_1. The total effect of overtly increasing investment in information from Σ_{ρ_1} to $\Sigma_{\hat{\rho}_1}$ can be similarly decomposed into

$$U_1(\rho;\rho) - U_1(\hat{\rho};\hat{\rho}) = \underbrace{U_1(\rho;\hat{\rho}) - U_1(\hat{\rho};\hat{\rho})}_{\text{value of covert investment}} + \underbrace{U_1(\rho;\rho) - U_1(\rho;\hat{\rho})}_{\text{strategic effect}}.$$

The value of covert investment (value of covert information) captures how the player 1's payoff increase by her ability to make better informed decisions while holding the player 2's strategy fixed. The strategic effect in our model corresponds to the value of transparency. It captures how player 1's payoff changes when the player 2's strategy is indirectly affected by the change in information quality.

From the first-order Taylor expansion, we have shown that the strategic effect of information depends on player 2's responsiveness, $a_2^\star(\rho) - a_2^\star(\hat{\rho})$, the externality player 2's action imposes on player 1, $u^1_{a_2}$, and additionally, player 1's "risk" attitude towards player 2's action, $u^1_{a_2 a_2}$. Our characterization of the value of transparency can hence be thought of as a stochastic extension to the characterization of strategic effects of investment by Fudenberg and Tirole (1984).

5 Conclusion

We have provided a framework to study how changes in the quality of private information affect equilibrium outcomes and welfare by extending the theory of monotone comparative statics to Bayesian games and Bayesian decision problems. Our theory of Bayesian Comparative Statics is comprised of three key components: an information order, a stochastic ordering of actions, and a class of utility functions. Our main theorem proves that for a subclass of supermodular utility functions, there is a duality between the order of actions and the information order: equilibrium outcomes become more dispersed in the stochastic ordering of actions if, and only if, signal quality increases in the information order.

From the perspective of positive economics, the comparative static results developed in this paper provide simple and convenient tools to study how changes in the quality of information about market fundamentals affect, for example, price dispersion in industrial economics, or the volatility of investment and aggregate output in macroeconomics. We provide several examples where our comparative statics can be readily applied, and an application of monopoly production in which a higher quality of information increases both the volatility and the average level of production.

From a normative perspective, we show that characterizing the comparative statics of actions with respect to information is a useful intermediate step to investigate informational externalities. We apply our results to study welfare effects of information: In Bayesian persuasion games, we characterize conditions under which the extremal disclosure of information is optimal. We also extend the industrial organization literature on information sharing in oligopolies to environments without linear best-responses.

Finally, we study the process of entry accommodation in oligopolistic markets where an incumbent can invest in information acquisition. The analysis of the indirect effect of information on the incumbent's profit through the induced behavior of the entrant (*the value of transparency*) is formally equivalent to characterizing the difference between the overt and covert demands for information. Leveraging our theory of Bayesian comparative statics, we characterize the sign of the value of transparency depending on the signs of the entrant's responsiveness and externality.

We expect the theory of Bayesian comparative statics will be useful to generalize many of the insights developed for quadratic economies to a broader class of utility functions. One avenue for future research is to study the efficient and equilibrium use of information and the over or under-coordination of equilibrium outcomes (Angeletos and Pavan, 2007) in non-linear environments. Another area of further study is how a central planner should intervene in markets with imperfect information. For example,Angeletos and Pavan (2009) identify policies that can improve the decentralized use of dispersed information without requiring the government to observe this information, while Lorenzoni (2010) and Angeletos and La'O (2011) study optimal monetary policy with uncertain fundamentals and dispersed information. An interesting open question is how to determine the optimal welfare improving policy in non-linear environments.

More generally, we hope our framework is valuable for future research on stochastic comparative statics that address the effects of *public* information, exogenous changes in market fundamentals and risk attitudes on equilibrium outcomes and welfare.

References

Amir, R. and Lazzati, N. . Endogenous information acquisition in bayesian games with strategic complementarities. *Journal of Economic Theory*, 163:684–698, 2016.

Angeletos, G.-M. and La'O, J. . Optimal monetary policy with informational frictions. Technical report, National Bureau of Economic Research, 2011.

Angeletos, G.-M. and La'O, J. . Sentiments. *Econometrica*, 81(2):739–779, 2013.

Angeletos, G.-M. and Pavan, A. . Efficient use of information and social value of information. *Econometrica*, 75(4):1103–1142, 2007.

Angeletos, G.-M. and Pavan, A. . Policy with dispersed information. *Journal of the European Economic Association*, 7(1):11–60, 2009.

Athey, S. . Single crossing properties and the existence of pure strategy equilibria in games of incomplete information. *Econometrica*, 69(4):861–889, 2001.

Athey, S. . Monotone comparative statics under uncertainty. *The Quarterly Journal of Economics*, 117(1):187–223, 2002.

Athey, S. and Levin, J. . The value of information in monotone decision problems. *Research in Economics*, 2017. ISSN 1090-9443.

Ballester, C. , Calvó-Armengol, A. , and Zenou, Y. . Who's who in networks. wanted: The key player. *Econometrica*, 74(5):1403–1417, 2006.

Bergemann, D. and Morris, S. . Robust predictions in games with incomplete information. *Econometrica*, 81(4):1251–1308, 2013.

Bergemann, D. and Morris, S. . Bayes correlated equilibrium and the comparison of information structures in games. *Theoretical Economics*, 11(2):487–522, 2016.

Blackwell, D. . Comparison of experiments. In *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pages 93–102, Berkeley, Calif., 1951. University of California Press.

Blackwell, D. . Equivalent comparisons of experiments. *Ann. Math. Statist.*, 24(2):265–272, 06 1953.

Bulow, J. , Geanakoplos, J. D. , and Klemperer, P. D. . Multimarket oligopoly: Strategic substitutes and complements. *Journal of Political Economy*, 93(3):488–511, 1985.

Candogan, O. , Bimpikis, K. , and Ozdaglar, A. . Optimal pricing in networks with externalities. *Operations Research*, 60(4):883–905, 2012.

Clarke, R. N. . Collusion and the incentives for information sharing. *The Bell Journal of Economics*, pages 383–394, 1983.

Colombo, L. , Femminis, G. , and Pavan, A. . Information acquisition and welfare. *Review of Economic Studies*, 81(4):1438–1483, 2014.

Fudenberg, D. and Tirole, J. . The fat-cat effect, the puppy-dog ploy, and the lean and hungry look. *The American Economic Review*, 74(2):361–366, 1984.

Gal-Or, E. . Information sharing in oligopoly. *Econometrica: Journal of the Econometric Society*, pages 329–343, 1985.

Hellwig, C. and Veldkamp, L. . Knowing what others know: Coordination motives in information acquisition. *The Review of Economic Studies*, 76(1):223–251, 2009.

Hirshleifer, J. . The private and social value of information and the reward to inventive activity. *The American Economic Review*, 61(4):561–74, 1971.

Holmstrom, B. . Moral hazard in teams. *The Bell Journal of Economics*, pages 324–340, 1982.

Jensen, M. K. . Distributional comparative statics. *The Review of Economic Studies*, 85(1): 581–610, 2018.

Jewitt, I. . Information order in decision and agency problems. *mimeo*, 01 2006.

Kamenica, E. and Gentzkow, M. . Bayesian persuasion. *American Economic Review*, 101(6): 2590–2615, 2011.

Kamien, M. I. , Tauman, Y. , and Zamir, S. . On the value of information in a strategic conflict. *Games and Economic Behavior*, 2(2):129 – 153, 1990. ISSN 0899-8256.

Keynes, J. M. . The general theory of money, interest and employment. *Reprinted in The Collected Writings of John Maynard Keynes*, 7, 1936.

Kimball, M. S. . Precautionary saving in the small and in the large. *Econometrica: Journal of the Econometric Society*, pages 53–73, 1990.

Leal-Vizcaíno, R. and Mekonnen, T. . Information sharing in supermodualr games. 2018.

Lehmann, E. L. . Comparing location experiments. *Ann. Statist.*, 16(2):521–533, 06 1988.

Lorenzoni, G. . Optimal monetary policy with uncertain fundamentals and dispersed information. *The Review of Economic Studies*, 77(1):305–338, 2010.

Lu, J. . Random choice and private information. *Econometrica*, 84(6):1983–2027, 2016. ISSN 1468-0262.

Lucas Jr, R. E. . Expectations and the neutrality of money. *Journal of economic theory*, 4(2): 103–124, 1972.

Marinacci, M. and Montrucchio, L. . Ultramodular functions. *Mathematics of Operations Research*, 30(2):311–332, 2005.

Mekonnen, T. and Leal-Vizcaíno, R. . Bayesian comparative statics: Online appendix. 2018.

Milgrom, P. and Roberts, J. . Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica: Journal of the Econometric Society*, pages 1255–1277, 1990.

Milgrom, P. and Roberts, J. . Comparing equilibria. *The American Economic Review*, pages 441–459, 1994.

Milgrom, P. and Shannon, C. . Monotone comparative statics. *Econometrica*, 62(1):157–180, 1994. ISSN 00129682, 14680262.

Milgrom, P. R. and Weber, R. J. . A Theory of Auctions and Competitive Bidding. *Econometrica*, 50(5):1089–1122, September 1982.

Morris, S. and Shin, H. S. . Social value of public information. *The American Economic Review*, 92(5):1521–1534, 2002.

Müller, A. and Stoyan, D. . *Comparison methods for stochastic models and risks*, volume 389. Wiley, 2002.

Myatt, D. P. and Wallace, C. . Endogenous information acquisition in coordination games. *The Review of Economic Studies*, 79(1):340–374, 2011.

Myatt, D. P. and Wallace, C. . Cournot competition and the social value of information. *Journal of Economic Theory*, 158:466 – 506, 2015. ISSN 0022-0531. Symposium on Information, Coordination, and Market Frictions.

Neyman, A. . The positive value of information. *Games and Economic Behavior*, 3(3):350–355, 1991.

Novshek, W. and Sonnenschein, H. . Fulfilled expectations cournot duopoly with information acquisition and release. *The Bell Journal of Economics*, pages 214–218, 1982.

Pavan, A. and Vives, X. . Information, coordination, and market frictions: An introduction. *Journal of Economic Theory*, 158:407 – 426, 2015. ISSN 0022-0531. Symposium on Information, Coordination, and Market Frictions.

Persico, N. . Information acquisition in auctions. *Econometrica*, 68(1):135–148, 2000. ISSN 1468-0262.

Quah, J. K.-H. and Strulovici, B. . Comparative statics, informativeness, and the interval dominance order. *Econometrica*, 77(6):1949–1992, 2009. ISSN 1468-0262.

Radner, R. . Team decision problems. *Ann. Math. Statist.*, 33(3):857–881, 09 1962.

Raith, M. . A general model of information sharing in oligopoly. *Journal of Economic Theory*, 71(1):260–288, 1996.

Rothschild, M. and Stiglitz, J. E. . Increasing risk ii: Its economic consequences. *Journal of Economic theory*, 3(1):66–84, 1971.

Shaked, M. and Shanthikumar, J. G. . *Stochastic orders*. Springer Science & Business Media, 2007.

Shapiro, C. . Theories of oligopoly behavior. *Handbook of industrial organization*, 1:329–414, 1989.

Tchen, A. H. . Inequalities for distributions with given marginals. *The Annals of Probability*, pages 814–827, 1980.

Tirole, J. . *The theory of industrial organization*. MIT press, 1988.

Tirole, J. . Cognitive games and cognitive traps. 2015.

Ui, T. and Yoshizawa, Y. . Characterizing social value of information. *Journal of Economic Theory*, 158:507 – 535, 2015. ISSN 0022-0531. Symposium on Information, Coordination, and Market Frictions.

Van Zandt, T. and Vives, X. . Monotone equilibria in bayesian games of strategic complementarities. *Journal of Economic Theory*, 134(1):339–360, 2007.

Villas-Boas, J. M. . Comparative statics of fixed points. *Journal of Economic Theory*, 73(1): 183–198, 1997.

Vives, X. . Duopoly information equilibrium: Cournot and bertrand. *Journal of Economic Theory*, 34(1):71–94, 1984.

Vives, X. . Nash equilibrium with strategic complementarities. *Journal of Mathematical Economics*, 19(3):305–321, 1990.

Vives, X. . *Oligopoly pricing: old ideas and new tools*. MIT press, 2001.

6 Appendix

6.1 Preliminary Lemmas

We provide two equivalent characterizations of responsiveness, one using the CDF $H(\cdot;\rho)$ and another usnig the quantile function defined as $\hat{a}(q;\rho) = \inf\{z : q \leq H(z;\rho)\}$ for $q \in (0,1)$.

Lemma 1 *[Shaked and Shantikumar, 2007; Theorem 4.A.2-A.3]*
Given two information structures $\Sigma_{\rho''}$ and $\Sigma_{\rho'}$, the following are equivalent:

i. An agent is more responsive with higher mean under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$.

ii. For all $x \in \mathbb{R}$,

$$\int_x^\infty H(z;\rho'')dz \leq \int_x^\infty H(z;\rho')dz.$$

iii. For all $t \in [0,1]$,

$$\int_t^1 \hat{a}(q;\rho'')dq \geq \int_t^1 \hat{a}(q;\rho')dq.$$

Similarly, the following are equivalent:

iv. An agent is more responsive with lower mean under $\Sigma_{\rho''}$ *than under* $\Sigma_{\rho'}$.

v. For all $x \in \mathbb{R}$,

$$\int_{-\infty}^x H(z;\rho'')dz \geq \int_{-\infty}^x H(z;\rho')dz.$$

vi. For all $t \in [0,1]$,

$$\int_0^t \hat{a}(q;\rho'')dq \leq \int_0^t \hat{a}(q;\rho')dq.$$

The following characterization of the supermodular stochastic order will prove useful for the proof of Theorem 1.

Lemma 2 *Given two information structures* $\Sigma_{\rho''}$ *and* $\Sigma_{\rho'}$, $\rho'' \succeq_{spm} \rho'$ *if, and only if, for all integrable functions* $\psi : \Theta \times S \to \mathbb{R}$ *that satisfy increasing differences (ID) in* $(\theta; s)$,

$$\int_{\Theta\times S} \psi(\theta,s)dF(\theta,s;\rho'') \geq \int_{\Theta\times S} \psi(\theta,s)dF(\theta,s;\rho')$$

Proof. Recall that all information structures induce the same marginal distribution of $\tilde{\theta}$ as it corresponds to the agent's prior. We have also assumed (WLOG) that all information structures induce the same marginal distribution of $\tilde{s}$. The result follows from (Theorem 3.8.2 of Müller and Stoyan (2002) or Tchen (1980). ■

Some of our results also make use of the following result from Lemma 1 of Quah and Strulovici (2009)

Lemma 3 *Let* $g : [x', x''] \to \mathbb{R}$ *and* $h : [x', x''] \to \mathbb{R}$ *be integrable functions.*

1. *If* g *is increasing and* $\int_x^{x''} h(t)dt \geq 0$ *for all* $x \in [x', x'']$, *then* $\int_{x'}^{x''} g(t)h(t)dt \geq g(x') \int_{x'}^{x''} h(t)dt$
2. *If* g *is decreasing and* $\int_{x'}^{x} h(t)dt \geq 0$ *for all* $x \in [x', x'']$, *then* $\int_{x'}^{x''} g(t)h(t)dt \geq g(x'') \int_{x'}^{x''} h(t)dt$

6.2 Single-Agent

Proof of Theorem 1

Proof. ($\Longrightarrow$) The payoff $u(\theta, a)$ satisfies ID in $(\theta; a)$ and the information structure Σ_ρ has the property that $s > s'$ implies $\mu(\cdot|s;\rho) \succeq_{FOSD} \mu(\cdot|s';\rho)$. From monotone comparative statics, the optimal action $a(\rho) : S \to A$ is a monotone function of s. Hence, from an ex-ante perspective, the optimal action coincdes with the quantile function we used to define responsiveness in Lemma 1, i.e., $a(\rho) = \hat{a}(\rho)$ almost surely.

Without loss of generality, we assume that the marginal on signals is uniformly distributed on the unit interval.[31] For any two information structures $\rho'' \succeq_{spm} \rho'$ and any signal realization $s \in [0, 1]$, the first order conditions imply that

$$\int_\Theta u_a(\theta, a(s;\rho''))\mu(d\theta|s;\rho'') - \int_\Theta u_a(\theta, a(s;\rho'))\mu(d\theta|s;\rho') = 0$$

which we rewrite as

$$\int_\Theta \Big(u_a(\theta, a(s;\rho'')) - u_a(\theta, a(s;\rho'))\Big)\mu(d\theta|s;\rho'') + \int_\Theta u_a(\theta, a(s;\rho'))\Big(\mu(d\theta|s;\rho'') - \mu(d\theta|s;\rho')\Big) = 0$$

If $u \in \mathcal{U}^I$, then $u_a(\theta, a)$ is convex in a for all θ. Thus,

$$u_a(\theta, a(s;\rho'')) - u_a(\theta, a(s;\rho')) \geq u_{aa}(\theta, a(s;\rho'))\big(a(s;\rho'') - a(s;\rho')\big)$$

and

$$\Big(a(s;\rho'') - a(s;\rho')\Big)\int_\Theta u_{aa}(\theta, a(s;\rho'))\mu(d\theta|s;\rho'') + \int_\Theta u_a(\theta, a(s;\rho'))\Big(\mu(d\theta|s;\rho'') - \mu(d\theta|s;\rho')\Big) \leq 0.$$

[31] As mentioned in the text, we can apply the integral probability transformation to signals.

For each $t \in [0,1]$,

$$\begin{aligned}
&\int_t^1 \big(a(s;\rho') - a(s;\rho'')\big)ds \\
&\leq \int_t^1 \underbrace{\left(-\int_\Theta u_{aa}(\theta, a(s;\rho'))\mu(d\theta|s;\rho'') \right)^{-1}}_{\triangleq B(s)} \int_\Theta u_a(\theta, a(s;\rho'))\Big(\mu(d\theta|s;\rho') - \mu(d\theta|s;\rho'')\Big)ds \\
&= \int_{\Theta\times[0,1]} u_a(\theta, a(s;\rho'))B(s)\mathbb{1}_{[s\geq t]}\Big(dF(\theta,s;\rho') - dF(\theta,s;\rho'')\Big),
\end{aligned}$$

where $\mathbb{1}_{[s\geq t]}$ is the indicator function that equals 1 if $s \geq t$ and 0 otherwise.

Define $\psi(\theta,s;t) \triangleq u_a(\theta,a(s;\rho'))B(s)\mathbb{1}_{[s\geq t]}$. For any $\theta'' > \theta'$, $\psi(\theta'',s;t) - \psi(\theta',s;t) = 0$ for $s < t$ and

$$\psi(\theta'',s;t) - \psi(\theta',s;t) = B(s)\Big(u_a(\theta'',a(s;\rho')) - u_a(\theta',a(s;\rho'))\Big) \geq 0$$

for $s \geq t$. The inequality follows from ID of u in $(\theta;a)$ and the strict concavity of u in a. Since $u \in \mathcal{U}^I$, u_a also satisfies ID in $(\theta;a)$, i.e., $u_a(\theta'',a) - u_a(\theta',a)$ is increasing in a. Since $a(s;\rho')$ is increasing in s, $u_a(\theta'',a(s;\rho')) - u_a(\theta',a(s;\rho'))$ is also increasing in s.

Additionally, $u \in \mathcal{U}^I$ implies that $-u_a$ satisfies decreasing differences in $(\theta;a)$ and is concave in a. Hence, $-u_{aa}(\theta,a)$ is decreasing in both θ and a. Since higher signal realizations lead to higher actions and to first-order stochastic shifts in beliefs,

$$-\int_\Theta u_{aa}(\theta,a(s;\rho'))\mu(d\theta|s;\rho'')$$

is a decreasing function of s. Thus $B(s)$ is increasing in s. We can therefore conclude that $\psi(\theta'',s;t) - \psi(\theta',s;t)$ is increasing in s. In other words, $\psi(\theta,s;t)$ satisfies ID in $(\theta;s)$. Thus, for each $t \in [0,1]$,

$$\begin{aligned}
&\int_t^1 \big(a(s;\rho') - a(s;\rho'')\big)ds \\
&\leq \int_{\Theta\times[0,1]} \psi(\theta,s;t)\Big(dF(\theta,s;\rho') - dF(\theta,s;\rho'')\Big) \leq 0
\end{aligned}$$

where the last inequality follows from Lemma 2.

($\Longleftarrow$) By definition, if $\rho'' \not\geq_{spm} \rho'$, there exists a $(\theta^*, s^*) \in \Theta \times [0,1]$ such that

$$F(\theta^*, s^*; \rho'') < F(\theta^*, s^*; \rho').$$

Define a payoff function

$$u(\theta, a) = -\frac{1}{2}\Big(\bar{a} - \mathbb{1}_{[\theta \leq \theta^*]}(\bar{a} - \underline{a}) - a\Big)^2.$$

The payoff $u(\theta, a)$ satisfies (A.1)-(A.4): It is continuous, twice differentiable, and strictly concave in a for each $\theta \in \Theta$. It satisfies ID in $(\theta; a)$. For each $\theta \in \Theta$, the optimal action is easily computed from the first order conditions so that the optimal action under complete information is $\underline{a}$ if $\theta \leq \theta^*$ and $\bar{a}$ otherwise. Furthermore, the marginal utility $u_a(\theta, a) = \bar{a} - \mathbb{1}_{[\theta \leq \theta^*]}(\bar{a} - \underline{a}) - a$ is

i. linear in a for all $\theta \in \Theta$, and

ii. has constant differences in $(\theta; a)$.

Therefore, $u \in \mathcal{U}^I \bigcap \mathcal{U}^{\downarrow}$.

For any given Σ_ρ,

$$\begin{aligned} a(s;\rho) =& \bar{a} - (\bar{a} - \underline{a}) E\left[\mathbb{1}_{[\tilde{\theta} \leq \theta^*]} | s; \rho\right] \\ =& \bar{a} - (\bar{a} - \underline{a}) \int_{\underline{\theta}}^{\theta^*} \mu(d\omega | s; \rho). \end{aligned}$$

Then given $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$,

$$\begin{aligned} &\int_0^{s^*} \big(a(s;\rho'') - a(s;\rho')\big) dF_S(s) \\ =& (\bar{a} - \underline{a}) \Big(F(\theta^*, s^*; \rho') - F(\theta^*, s^*; \rho'')\Big) > 0. \end{aligned}$$

Therefore, the agent is not more responsive with a lower mean under $\Sigma_{\rho''}$ than $\Sigma_{\rho'}$. Notice that for any Σ_ρ,

$$E[a(\rho)] = \bar{a} - (\bar{a} - \underline{a}) \int_0^1 \int_{\underline{\theta}}^{\theta^*} \mu(d\omega | s; \rho) dF_S(s) = \bar{a} - (\bar{a} - \underline{a}) \int_{\underline{\theta}}^{\theta^*} \mu^0(d\omega),$$

which is independent of ρ. Thus,

$$\int_{s^*}^{1} \big(a(s;\rho'') - a(s;\rho')\big) dF_S(s)$$

$$= \underbrace{\int_{0}^{1} \big(a(s;\rho'') - a(s;\rho')\big) dF_S(s)}_{\substack{=E[a(\rho'')]-E[a(\rho')] \\ =0}} - \left(\underbrace{\int_{0}^{s^*} \big(a(s;\rho'') - a(s;\rho')\big) dF_S(s)}_{>0} \right) < 0.$$

Therefore, the agent is not more responsive with a higher mean under $\Sigma_{\rho''}$ than $\Sigma_{\rho'}$. ■

Proof of Proposition 1

Proof. Let $a_i = a^*(\mu_i)$ for $i = 1, 2$, $a_\lambda = \lambda a_1 + (1-\lambda)a_2$, and $\mu_\lambda = \lambda\mu_1 + (1-\lambda)\mu_2$. By the first order condition, we have that $\int_\Theta u_a(\theta, a_i)\mu_i(d\theta) = 0$. Let $u \in \mathcal{U}^I$.

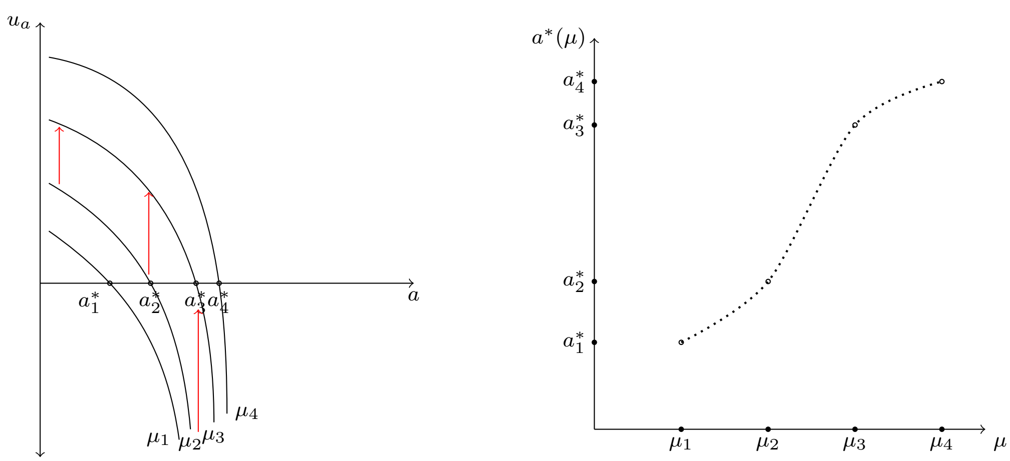
$$\begin{aligned}
\int_\Theta u_a(\theta, a_\lambda)\mu_\lambda(d\theta) &\leq \lambda \int_\Theta u_a(\theta, a_1)\mu_\lambda(d\theta) + (1-\lambda) \int_\Theta u_a(\theta, a_2)\mu_\lambda(d\theta) \\
&= \lambda^2 \int_\Theta u_a(\theta, a_1)\mu_1(d\theta) + (1-\lambda)^2 \int_\Theta u_a(\theta, a_2)\mu_2(d\theta) \\
&+ \lambda(1-\lambda) \left[\int_\Theta u_a(\theta, a_2)\mu_1(d\theta) + \int_\Theta u_a(\theta, a_1)\mu_2(d\theta) \right] \\
&= \lambda(1-\lambda) \int_\Theta \left[u_a(\theta, a_1) - u_a(\theta, a_2)\right] (\mu_2(d\theta) - \mu_1(d\theta)) \\
&\leq 0
\end{aligned}$$

where the first inequality follows from the convexity of u_a. As already noted, ID of the utility $u(\theta, a)$ in $(\theta; a)$ along with $\mu_2 \succeq_{FOSD} \mu_1$ implies $a_2 \geq a_1$. By ID of the marginal utility u_a in $(\theta; a)$, we have $u_a(\theta, a_1) - u_a(\theta, a_2)$ is a decreasing function of θ. The last inequality then follows from the definition of first-order stochastic dominance. Since the marginal value of a_λ is non-positive at μ_λ, we must have $a^*(\mu_\lambda) \leq a_\lambda$. A symmetric argument establishes that if $u \in \mathcal{U}^D$, then $a^*(\mu_\lambda) \geq a_\lambda$. ■

6.2.1 When Responsiveness Fails

In this section, we explore why a higher quality of information may not lead to more dispersed optimal actions when $u \notin \mathcal{U}^I \cup \mathcal{U}^D$. Once again, let the state space be $\Theta = \{\underline{\theta}, \bar{\theta}\}$. Consider four different beliefs $\{\mu_n\}_{n=1,2,3,4}$ such that $\mu_n = n\delta$ for some $\delta \in (0, 1/4)$. Beliefs are ordered by first-order stochastic dominance with $\mu_4 \succeq_{FOSD} \mu_3 \succeq_{FOSD} \mu_2 \succeq_{FOSD} \mu_1$.

In Figure 7a, we plot the expected marginal utilities of some payoff function u. Notice that $u(\theta, a)$ satisfies ID in $(\theta; a)$—the expected marginal utility of μ_{n+1} lies above the expected marginal utility of μ_n. Thus, $a_{n+1} \geq a_n$. Furthermore, $u_a(\theta, a)$ also satisfies ID in $(\theta; a)$—the height of the red arrows increases left to right. However, the marginal utilities are now concave which implies that the marginal utility diminishes at an accelerating rate. Therefore, $u \notin \mathcal{U}^I$. Furthermore, $a_4 - a_3 < a_3 - a_2$ whereas $a_3 - a_2 > a_2 - a_1$. Figure 7b depicts this "non-convexity" of the optimal action as a function of beliefs.



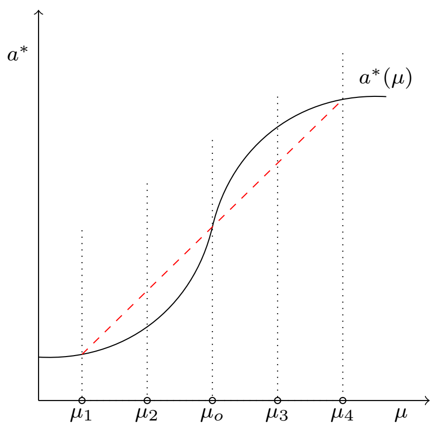
(a) Marginal utilities

(b) Optimal action

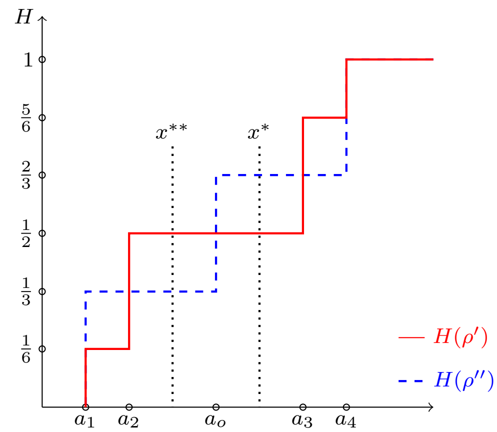
Figure 7: Non-convexity for $u \notin \mathcal{U}^I$

Figure 8 illustrates why the agent may not be responsive to an increase in the quality of information when the optimal action is neither convex nor concave, as in Figure 7b. Let $\Sigma_{\rho''}$ be an information structure that induces three posteriors $\{\mu_1, \mu_o, \mu_4\}$ with probabilities $\{1/3, 1/3, 1/3\}$ such that $\mu_4 \succeq_{FOSD} \mu_o \succeq_{FOSD} \mu_1$. Let $\Sigma_{\rho'}$ induce posteriors $\{\mu_1, \mu_2, \mu_3, \mu_4\}$ with probability $\{1/6, 1/3, 1/3, 1/6\}$ where $\mu_2 = 0.5\mu_1 + 0.5\mu_o$ and $\mu_3 = 0.5\mu_4 + 0.5\mu_o$. Then, $\mu_4 \succeq_{FOSD} \mu_3 \succeq_{FOSD} \mu_2 \succeq_{FOSD} \mu_1$. Notice that $\Sigma_{\rho'}$ is a equivalent to getting information from $\Sigma_{\rho''}$ with probability 0.5 and no information with probability 0.5. Thus, $\rho'' \succeq_{spm} \rho'$.

Let $a^*(\mu)$ be neither convex nor concave and let the average action under $\Sigma_{\rho''}$ equal the average action under $\Sigma_{\rho'}$. In Figure 8a below, this corresponds to the point of intersection of the dashed line and the solid curved line at μ_o. Figure 8b maps the distribution over optimal actions. $\Sigma_{\rho''}$ induces the dashed line while $\Sigma_{\rho'}$ induces the solid line.



(a) Optimal action



(b) Induced distribution

Figure 8: Non-convexity/concavity and non-responsiveness

If we start integrating from the right, then $\int_x^\infty H(z;\rho'') - H(z;\rho')dz \leq 0$ for all $x > a^*(\mu_3)$ but the sign changes at some point $x^* \in (a^*(\mu_o), a^*(\mu_3))$. Thus, the agent is not more responsive with a higher mean under $\Sigma_{\rho''}$. If we instead integrate from the left, then $\int_{-\infty}^x H(z;\rho'') - H(z;\rho')dz \geq 0$ for all $x < a^*(\mu_2)$ but the sign changes at some point $x^{**} \in (a^*(\mu_2), a(\mu_o))$. Thus, the agent is not more responsive with a lower mean under $\Sigma_{\rho''}$.

In fact, as the average action under $\Sigma_{\rho''}$ equals the average action under $\Sigma_{\rho'}$, we can conclude that $a(\rho'')$ and $a(\rho')$ cannot be ordered by most univariate stochastic variability orders such as second-order stochastic dominance, mean-preserving spreads, Lorenz order, dilation order, and dispersive order.[32]

Another reason why a higher quality of information may not lead to more responsive behavior is when the interior solution assumption, (A.3), is violated. Suppose the upper limit on the action space, $\bar{a}$, is a binding constraint for the prior, i.e., $a^*(\mu_o) = \bar{a}$. Let $\Sigma_{\rho'}$ be a completely uninformative information structure. Then, $\Sigma_{\rho'}$ induces $\bar{a}$ with probability one, thereby first-order stochastically dominating the distribution over actions induced by any other information structure $\Sigma_{\rho''}$, even if $\rho'' \succeq_{spm} \rho'$.

[32] Shaked and Shanthikumar (2007) provide a thorough treatment of these orders.

6.3 Applications

Proof of Proposition 2

Proof. $-qP''(q)/P'(q) \leq 1$ implies that $CS(q) = \int_0^q P(t)dt - qP(q)$ is an increasing convex function. If $\pi \in \mathcal{U}^I$, then for two information structures $\Sigma_{\rho''}$ and $\Sigma_{\rho'}$ with $\rho'' \succeq_{spm} \rho'$, $q^M(\rho'')$ dominates $q^M(\rho')$ in the increasing convex order, i.e., the monopolist is more responsive with a higher mean under $\Sigma_{\rho''}$. By definition, $E[CS(q^M(\rho''))] \geq E[CS(q^M(\rho'))]$. ■

Proof of Proposition 3

Proof. Take two information structures $\Sigma_{\rho''}, \Sigma_{\rho'}$ with $\rho'' \succeq_{spm} \rho'$. The sender's ex-ante payoff difference is given by

$$V(\rho'') - V(\rho') \tag{1}$$

$$= \int_{\Theta\times S} v(\theta, a(s;\rho''))\big[dF(\theta,s;\rho'') - dF(\theta,s;\rho')\big] + \int_{\Theta\times S}\big[v(\theta,a(s;\rho'')) - v(\theta,a(s;\rho'))\big]dF(\theta,s;\rho').$$

When $v(\theta,a)$ has ID in $(\theta;a)$ and $a(s;\rho)$ is increasing in s (which follows from $u(\theta,a)$ satisfying ID in $(\theta;a)$ and posteriors increasing in FOSD as s increases), $v(\theta,a(s;\rho))$ has ID in $(\theta;s)$. Thus, by Lemma 2, the first integral term is non-negative.

When $v(\theta,a)$ is differentiable[33] and convex in a for all $\theta \in \Theta$, the second integral term satisfies

$$\int_{\Theta\times S}\big[v(\theta,a(s;\rho'')) - v(\theta,a(s;\rho'))\big]dF(\theta,s;\rho')$$
$$\geq \int_0^1 [a(s;\rho'') - a(s;\rho')] \underbrace{\int_\Theta v_a(\theta,a(s,\rho'))\mu(d\theta|s;\rho')}_{=\mathbb{E}_\Theta[v_a(\tilde{\theta},a(s,\rho'))|s;\rho']} ds.$$

When $v(\theta,a)$ is both convex in a and has ID in $(\theta;a)$, and posterior beliefs increase in FOSD as s increases, the term $\mathbb{E}_\Theta[v_a(\tilde{\theta},a(s,\rho'))|s;\rho']$ is an increasing function of s.

Case I: $u \in \mathcal{U}^I$ and v is increasing in a.

From Theorem 1, $u \in \mathcal{U}^I$ implies that

$$\int_t^1 [a(s;\rho'') - a(s;\rho')]ds \geq 0, \forall t \in [0,1].$$

[33] If v is not differentiable, we can uniformly approximate it by a convex analytic function.

From Lemma 3 and $v(\theta, a)$ increasing in a,

$$\int_0^1 [a(s;\rho'') - a(s;\rho')]\mathbb{E}_\Theta[v_a(\tilde{\theta}, a(s,\rho'))|s;\rho']ds$$

$$\geq \underbrace{\int_0^1 [a(s;\rho'') - a(s;\rho')]ds}_{\geq 0 \text{ by } u\in\mathcal{U}^I} \mathbb{E}_\Theta[v_a(\tilde{\theta}, a(0,\rho'))|0;\rho'] \geq 0.$$

Hence, the second integral term in (1) is also non-negative. In other words, $V(\rho'') \geq V(\rho')$.

Case II: $u \in \mathcal{U}^D$ and v is decreasing in a.
From Theorem 1, $u \in \mathcal{U}^D$ implies that

$$\int_0^t [a(s;\rho') - a(s;\rho'')]ds \geq 0, \forall t \in [0,1].$$

From Lemma 3 and $v(\theta, a)$ decreasing in a,

$$\int_0^1 [a(s;\rho'') - a(s;\rho')]\mathbb{E}_\Theta[v_a(\tilde{\theta}, a(s,\rho'))|s;\rho']ds$$

$$= \int_0^1 [a(s;\rho') - a(s;\rho'')]\mathbb{E}_\Theta[-v_a(\tilde{\theta}, a(s,\rho'))|s;\rho']ds$$

$$\geq \underbrace{\int_0^1 [a(s;\rho') - a(s;\rho'')]ds}_{\geq 0 \text{ by } u\in\mathcal{U}^D} \mathbb{E}_\Theta[-v_a(\tilde{\theta}, a(1,\rho'))|1;\rho'] \geq 0.$$

Once again, the second integral term in (1) is also non-negative. Therefore, $V(\rho'') \geq V(\rho')$.

Case III: $u \in \mathcal{U}^D \cap \mathcal{U}^D$.
From Theorem 1, $u \in \mathcal{U}^I \cap \mathcal{U}^D$ implies that

$$\int_t^1 [a(s;\rho'') - a(s;\rho')]ds \geq 0, \forall t \in [0,1],$$

with equality at $t = 0$. From Lemma 3

$$\int_0^1 [a(s;\rho'') - a(s;\rho')] \int_\Theta v_a(\theta, a(s,\rho'))\mu(d\theta|s;\rho')ds$$

$$\geq \underbrace{\int_0^1 [a(s;\rho'') - a(s;\rho')]ds}_{=0} \int_\Theta v_a(\theta, a(0,\rho'))\mu(d\theta|0;\rho')ds = 0.$$

Hence, $V(\rho'') \geq V(\rho')$.

By setting the sender's payoff in the above arguments to $-v(\theta, a)$, we get the corresponding statements for preferences that satisfy decreasing differences and concavity. ■

Proof of Corollary 2

Proof. Let $\Sigma_{\bar{\rho}}$ and $\Sigma_{\underline{\rho}}$ be the full-information and no-information structures respectively. Any information structure Σ_ρ is Blackwell dominated by $\Sigma_{\bar{\rho}}$ and Blackwell dominates $\Sigma_{\underline{\rho}}$.

When $|supp(\mu^0)| = 2$, it is without loss of generality to restrict attention to information structures that induce posteriors ordered by FOSD. Furthermore, the Blackwell order is a subset of supermodular order (see the online appendix Mekonnen and Leal-Vizcaíno (2018)). Hence, for any information structure Σ_ρ, $\bar{\rho} \succeq_{spm} \rho \succeq_{spm} \underline{\rho}$. We get the desired result by applying Proposition 3. ■

Proof of Proposition 4

Proof. Take any two information structures $\Sigma_{\rho''} \triangleq (\Sigma_{\bar{\rho}_1}, \Sigma_{\rho_2''})$, $\Sigma_{\rho'} \triangleq (\Sigma_{\bar{\rho}_1}, \Sigma_{\rho_2'})$ with $\rho_2'' \succeq_{spm} \rho_2'$. Then,

$$U_1(\rho'') - U_1(\rho')$$

$$= \int_{\Theta \times S_2} u^1\big(\theta, a_1^\star(\theta;\rho''), a_2^\star(s_2;\rho'')\big)dF(\theta, s_2;\rho_2'') - \int_{\Theta \times S_2} u^1\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(s_2;\rho')\big)dF(\theta, s_2;\rho_2'),$$

which can be written as

$$\int_{\Theta\times S_2}\Big[u^1\big(\theta,a_1^\star(\theta;\rho''),a_2^\star(s_2;\rho'')\big)-u^1\big(\theta,a_1^\star(\theta;\rho'),a_2^\star(s_2;\rho'')\big)\Big]dF(\theta,s_2;\rho'') \tag{2}$$
$$+\int_{\Theta\times S_2}\Big[u^1\big(\theta,a_1^\star(\theta;\rho'),a_2^\star(s_2;\rho'')\big)-u^1\big(\theta,a_1^\star(\theta;\rho'),a_2^\star(s_2;\rho')\big)\Big]dF(\theta,s_2;\rho'')$$
$$+\int_{\Theta\times S_2}u^1\big(\theta,a_1^\star(\theta;\rho'),a_2^\star(s_2;\rho')\big)\Big[dF(\theta,s_2;\rho'')-dF(\theta,s_2;\rho')\Big].$$

The first term of (2) is non-negative as $a_1^\star(\rho'')$ is player 1's best response to $a_2^\star(\rho'')$ and information structure $\Sigma_{\rho''}$. For the third term of (2), take $s_2''>s_2'$ which implies $a_2^\star(s_2'';\rho')\geq a_2^\star(s_2';\rho')$ and note that

$$u^1\big(\theta,a_1^\star(\theta;\rho'),a_2^\star(s_2'';\rho')\big)-u^1\big(\theta,a_1^\star(\theta;\rho'),a_2^\star(s_2';\rho')\big)$$

is increasing in θ because u^1 has ID $(\theta,a_1;a_2)$. Hence, $u^1\big(\theta,a_1^\star(\theta;\rho'),a_2^\star(s_2;\rho')\big)$ has ID in $(\theta;s_2)$. By Lemma 2,

$$\int_{\Theta\times S_2}u^1\big(\theta,a_1^\star(\theta;\rho'),a_2^\star(s_2;\rho')\big)\Big[dF(\theta,s_2;\rho'')-dF(\theta,s_2;\rho')\Big]\geq 0$$

and the third term of (2) is also non-negative.

By convexity and differentiability of u^1 in a_2, the second term of (2) can be rewritten as

$$\int_{S_2}\int_{\Theta}\Big[u^1\big(\theta,a_1^\star(\theta;\rho'),a_2^\star(s_2;\rho'')\big)-u^1\big(\theta,a_1^\star(\theta;\rho'),a_2^\star(s_2;\rho')\big)\Big]\mu(d\theta|s_2;\rho'')ds_2$$
$$\geq\int_{S_2}\big(a_2^\star(s_2;\rho'')-a_2^\star(s_2;\rho')\big)\int_{\Theta}u^1_{a_2}\big(\theta,a_1^\star(\theta;\rho'),a_2^\star(s_2;\rho')\big)\mu(d\theta|s_2;\rho'')ds_2.$$

Since u^1 has ID in $(\theta,a_1;a_2)$, $a_1^\star(\theta;\rho')$ is increasing in θ, $a_2^\star(s_2;\rho')$ is increasing in s_2, and assumption (A.12),

$$\int_{\Theta}u^1_{a_2}\big(\theta,a_1^\star(\theta;\rho'),a_2^\star(s_2;\rho')\big)\mu(d\theta|s_2;\rho'')$$

is increasing in s_2.

Case I: $u^i\in\Gamma^I$ for $i=1,2$ and u^1 is increasing in a_2.

By Theorem 2, $u^i\in\Gamma^I$ for $i=1,2$ implies $a_2^\star(\rho'')$ dominates $a_2^\star(\rho')$ in the increasing convex order. By Lemma 1,

$$\int_t^1 \Big(a_2^\star(s_2;\rho'') - a_2^\star(s_2;\rho')\Big)ds_2 \geq 0$$

for all $t \in [0,1]$. By Lemma 3, the second term of (2) is greater than

$$\int_{S_2} \big(a_2^\star(s_2;\rho'') - a_2^\star(s_2;\rho')\big) \int_\Theta u^1_{a_2}\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(s_2;\rho')\big)\mu(d\theta|s_2;\rho'')ds_2$$

$$\geq \underbrace{\int_{S_2} \big(a_2^\star(s_2;\rho'') - a_2^\star(s_2;\rho')\big)ds_2}_{\geq 0 \text{ by increasing convex order}} \int_\Theta \underbrace{u^1_{a_2}\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(0;\rho')\big)}_{\geq 0 \text{ as } u^1 \text{ is increasing in } a_2} \mu(d\theta|0;\rho'') \geq 0.$$

Thus, $\rho_2'' \succeq_{spm} \rho_2'$ implies $U_1(\rho'') \geq U_1(\rho')$. As $\bar{\rho}_2 \succeq_{spm} \rho_2$ for all $\Sigma_{\rho_2} \in \mathcal{P}_2$, player 1's ex-ante payoff is maximized by the full-information structure.

Case II: $u^i \in \Gamma^D$ for $i = 1, 2$ and u^1 is decreasing in a_2.

By Theorem 2, $u^i \in \Gamma^D$ for $i = 1, 2$ implies $a_2^\star(\rho'')$ dominates $a_2^\star(\rho')$ in the decreasing convex order. By Lemma 1,

$$\int_0^t \Big(a_2^\star(s_2;\rho') - a_2^\star(s_2;\rho'')\Big)ds_2 \geq 0$$

for all $t \in [0,1]$. By Lemma 3, the second term of (2) is greater than

$$\int_{S_2} \big(a_2^\star(s_2;\rho') - a_2^\star(s_2;\rho'')\big) \int_\Theta -u^1_{a_2}\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(s_2;\rho')\big)\mu(d\theta|s_2;\rho'')ds_2$$

$$\geq \underbrace{\int_{S_2} \big(a_2^\star(s_2;\rho') - a_2^\star(s_2;\rho'')\big)ds_2}_{\geq 0 \text{ by decreasing convex order}} \int_\Theta \underbrace{-u^1_{a_2}\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(1;\rho')\big)}_{\geq 0 \text{ as } u^1 \text{ is decreasing in } a_2} \mu(d\theta|1;\rho'') \geq 0.$$

Thus, $\rho_2'' \succeq_{spm} \rho_2'$ implies $U_1(\rho'') \geq U_1(\rho')$ and player 1's ex-ante payoff is maximized by the full-information structure.

Case III: $u^i \in \Gamma^I \cap \Gamma^D$ for $i = 1, 2$.

By Theorem 2, $u^i \in \Gamma^I \cap \Gamma^D$ for $i = 1, 2$ implies $a_2^\star(\rho'')$ is a mean-preserving spread of $a_2^\star(\rho')$, i.e., $a_2^\star(\rho'')$ dominates $a_2^\star(\rho')$ in both increasing and decreasing convex order. By Lemma 1,

$$\int_0^t \Big(a_2^\star(s_2;\rho'') - a_2^\star(s_2;\rho')\Big)ds_2 \geq 0$$

for all $t \in [0,1]$. By Lemma 3, the second term of (2) is greater than

$$\int_{S_2} \big(a_2^\star(s_2;\rho'') - a_2^\star(s_2;\rho')\big) \int_\Theta u^1_{a_2}\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(s_2;\rho')\big)\mu(d\theta|s_2;\rho'')ds_2$$

$$\geq \underbrace{\int_{S_2} \big(a_2^\star(s_2;\rho'') - a_2^\star(s_2;\rho')\big)ds_2}_{=0 \text{ by mean-preserving spread}} \int_\Theta u^1_{a_2}\big(\theta, a_1^\star(\theta;\rho'), a_2^\star(0;\rho')\big)\mu(d\theta|0;\rho'') = 0.$$

Thus, $\rho_2'' \succeq_{spm} \rho_2'$ implies $U_1(\rho'') \geq U_1(\rho')$ and player 1's ex-ante payoff is maximized by the full-information structure. ■

Proof of Theorem 3

Proof. We only show the proof for the case when $u^i \in \Gamma^I$ for $i = 1, 2$ and $u^1(\theta, a)$ is an increasing and convex function of a_2. The remaining cases can be established by a similar argument.[34]

Take two information structures $\Sigma_{\rho_1}, \Sigma_{\hat{\rho}_1} \in \mathcal{P}$. By definition, $VT(\rho;\hat{\rho}) = U_1(\rho;\rho) - U_1(\rho;\hat{\rho})$ is given by

$$\int_{\Theta\times S} \Big[u^1\big(\theta, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big) - u^1\big(\theta, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2;\hat{\rho})\big)\Big] d\boldsymbol{F}(\theta, s;\rho)$$

$$= \underbrace{\int_{\Theta\times S} \Big[u^1\big(\theta, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big) - u^1\big(\theta, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2;\rho)\big)\Big] d\boldsymbol{F}(\theta, s;\rho)}_{\geq 0 \text{ by optimality}}$$

$$+ \int_{\Theta\times S} \underbrace{\Big[u^1\big(\theta, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2;\rho)\big) - u^1\big(\theta, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2;\hat{\rho})\big)\Big]}_{\substack{\geq u^1_{a_2}(\theta, a_1^{BR}(s_1; a_2^\star(\hat{\rho}),\rho), a_2^\star(s_2;\hat{\rho}))(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})) \\ \text{by convexity of } u^1 \text{ in } a_2}} d\boldsymbol{F}(\theta, s;\rho)$$

$$\geq \int_0^1 \Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big) \int_{\Theta_1\times S_1} u^1_{a_2}\big(\theta, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2;\hat{\rho})\big) d\boldsymbol{F}(\theta, s_1|s_2;\rho) ds_2.$$

Define $\zeta : [0,1] \to \mathbb{R}$ by

$$\zeta(s_2) \triangleq \int_{\Theta_1\times S_1} u^1_{a_2}\big(\theta, a_1^{BR}(s_1; a_2^\star(\hat{\rho}), \rho), a_2^\star(s_2;\hat{\rho})\big) d\boldsymbol{F}(\theta, s_1|s_2;\rho).$$

[34] The reader may also refer to the proof of Proposition 4 which contains a similar proof for all three cases.

So far, we have established that

$$VT(\rho;\hat{\rho}) \geq \int_0^1 \Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big)\zeta(s_2)ds_2.$$

We can also rewrite $VT(\rho;\hat{\rho}) = U_1(\rho;\rho) - U_1(\rho;\hat{\rho})$ as

$$\begin{aligned}
&\int_{\Theta\times S}\Big[u^1\big(\theta, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big) - u^1\big(\theta, a_1^{BR}(s_1; a_2^\star(\hat{\rho}),\rho), a_2^\star(s_2;\hat{\rho})\big)\Big]d\boldsymbol{F}(\theta,s;\rho)\\
&= \int_{\Theta\times S}\underbrace{\Big[u^1\big(\theta, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big) - u^1\big(\theta, a_1^\star(s_1;\rho), a_2^\star(s_2;\hat{\rho})\big)\Big]}_{\substack{\leq -u^1_{a_2}(\theta, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho))(a_2^\star(s_2;\hat{\rho}) - a_2^\star(s_2;\rho)) \\ \text{by concavity of } -u^1 \text{ in } a_2}} d\boldsymbol{F}(\theta,s;\rho)\\
&+ \underbrace{\int_{\Theta\times S}\Big[u^1\big(\theta, a_1^\star(s_1;\rho), a_2^\star(s_2;\hat{\rho})\big) - u^1\big(\theta, a_1^{BR}(s_1; a_2^\star(\hat{\rho}),\rho), a_2^\star(s_2;\hat{\rho})\big)\Big]d\boldsymbol{F}(\theta,s;\rho)}_{\leq 0 \text{ by optimality}}\\
&\leq \int_0^1 \Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big)\int_{\Theta_1\times S_1} u^1_{a_2}\big(\theta, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big)d\boldsymbol{F}(\theta, s_1|s_2;\rho)ds_2.
\end{aligned}$$

Define $\eta : [0,1] \to \mathbb{R}$ by

$$\eta(s_2) \triangleq \int_{\Theta_1\times S_1} u^1_{a_2}\big(\theta, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\big)d\boldsymbol{F}(\theta, s_1|s_2;\rho).$$

Then,

$$\int_0^1 \Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big)\eta(s_2)ds_2 \geq VT(\rho;\hat{\rho}) \geq \int_0^1 \Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big)\zeta(s_2)ds_2.$$

Recall that $u^1(\theta,a)$ is increasing in a_2, i.e., positive externalities. Hence, both $\zeta(s_2)\geq 0$ and $\eta(s_2)\geq 0$ for all $s_2\in[0,1]$. Additionally, $u^1_{a_2}(\theta,a)$ is also increasing in a_2 by convexity. Thus, when we have independent private values, or when $u^1(\theta,a)$ satisfies ID in $(\theta,a_1;a_2)$ (along with (A.6) and (A.11)-(A.13)), then $\zeta(s_2)$ and $\eta(s_2)$ are increasing in s_2.

($\Longrightarrow$) Suppose $\rho_1 \succeq_{spm} \hat{\rho}_1$. From Theorem 2, $u^i\in\Gamma^I$ for $i=1,2$ implies that $a_2^\star(\rho)$ dominates

$a_2^\star(\hat{\rho})$ in the increasing convex order. By Lemma 1,

$$\int_t^1 \Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big) ds_2 \geq 0$$

for all $t \in [0,1]$. Using Lemma 3, we can then conclude that

$$\begin{aligned} VT(\rho;\hat{\rho}) \geq & \int_0^1 \Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big)\zeta(s_2) ds_2 \\ \geq & \zeta(0) \int_0^1 \Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big) ds_2 \\ \geq & 0. \end{aligned}$$

($\Longleftarrow$) Suppose $\rho_1 \not\succeq_{spm} \hat{\rho}_1$. By assumption, $\mathcal{P}$ is a totally ordered set of information structures. Thus, $\hat{\rho}_1 \succeq_{spm} \rho_1$. From Theorem 2, $u^i \in \Gamma^I$ for $i = 1,2$ implies that $a_2^\star(\hat{\rho})$ dominates $a_2^\star(\rho)$ in the increasing convex order. By Lemma 1,

$$\int_t^1 \Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big) ds_2 \leq 0$$

for all $t \in [0,1]$. Using the second mean value theorem, there exists $t^* \in [0,1]$ such that

$$\begin{aligned} VT(\rho;\hat{\rho}) \leq & \int_0^1 \Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big)\eta(s_2) ds_2 \\ = & \eta(1) \int_{t^*}^1 \Big(a_2^\star(s_2;\rho) - a_2^\star(s_2;\hat{\rho})\Big)\eta(s_2) ds_2 \\ \leq & 0. \end{aligned}$$

■