Leal Vizcaíno, René; Mekonnen, Teddy

**Working Paper**

# Bayesian comparative statics

Working Papers, No. 2019-03

**Provided in Cooperation with:**
Bank of Mexico, Mexico City

*Suggested Citation:* Leal Vizcaíno, René; Mekonnen, Teddy (2019) : Bayesian comparative statics, Working Papers, No. 2019-03, Banco de México, Ciudad de México

This Version is available at:
https://hdl.handle.net/10419/240671

Banco de México

Documentos de Investigación

Banco de México

Working Papers

N° 2019-03

# Bayesian Comparative Statics

René Leal Vizcaíno
Banco de México

Teddy Mekonnen
Caltech

March 2019

Documento de Investigación
2019-03

Working Paper
2019-03

# Bayesian Comparative Statics*

René Leal Vizcaíno†
Banco de México

Teddy Mekonnen‡
Caltech

**Abstract:** We study how information affects equilibria and welfare in games. For an agent, more precise information about an unknown state of the world 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.

**Keywords:** Comparative Statics, Information Acquisition, Information Orders, Persuasion, Value of Information, Supermodular Games.
**JEL Classification:** C44, C61, D42, D81

**Resumen:** Estudiamos cómo la información afecta el equilibrio y el bienestar en entornos estratégicos. Para un agente, información más precisa sobre un estado desconocido del mundo conduce a una dispersión de sus creencias que conserva la media. Proporcionamos condiciones necesarias y suficientes para obtener una mayor dispersión de sus acciones en equilibrio, ya sea con una media mayor o menor, cuando la información de al menos un agente en el juego aumenta. Aplicamos nuestro método de estática comparativa bayesiana para estudiar externalidades informativas en entornos estratégicos. En juegos de persuasión, obtenemos condiciones suficientes que conducen a la divulgación extrema de información. En mercados oligopolísticos, caracterizamos los incentivos de las empresas para compartir información. En modelos macroeconómicos, mostramos que la información no solo impacta la amplitud de los ciclos económicos, sino que también afecta el nivel de producción agregada.

**Palabras Clave:** Estática Comparativa, Adquisición de Información, Órdenes de Información, Persuasión, Valor de la Información, Juegos Supermodulares.

---

*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, Jean Tirole, and seminar participants at 2017 Los Angeles Economic Theory Conference, ITAM First Alumni Conference, Northwestern theory bag lunch and Caltech SISL.

† Dirección General de Investigación Económica, Banco de México. Email: rlealv@banxico.org.mx.

‡ Social Sciences. Email: mekonnen@caltech.edu.

# 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 can sometimes 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 first analyzing how information affects equilibrium actions and then comparing the equilibrium actions and payoffs to the efficient outcomes of a game.

Interestingly enough, the question of how information affects actions in games of imperfect information has only been partially studied in some settings where closed-form solutions to equilibrium actions can be explicitly computed, specifically, 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 embeds the familiar linear-quadratic games with Gaussian signals. Our 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, these comparative statics are also a useful intermediate step to characterize informational externalities and investigate the welfare effects of information beyond linear-quadratic-Gaussian games.

[1]For a very recent symposium summarizing the state of the art in these games see Pavan and Vives (2015) and the references therein.

Our theory of *Bayesian comparative statics* is comprised of three key ingredients: a stochastic order of equilibrium actions (call it order 1), an information order (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 2, then for any preference in the class of utility functions, $A$ induces equilibrium actions that are more dispersed according to order 1 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 2.

We illustrate the usefulness of our approach through several examples and applications. In persuasion games, we characterize the minimal and maximal levels of conflict between a sender and a receiver, conditions under which extremal disclosure of information is optimal (either full revelation or no information). 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.

The remainder of the paper is structured as follows. In Section 2, we consider an example of monopoly production with an uncertain cost parameter to concretely motivate our comparative statics question and illustrate our approach to the normative and positive effects of information. We also discuss more thoroughly the theory of Bayesian comparative statics and relate it to

[2]See for example: Raith (1996), Angeletos and Pavan (2007), Bergemann and Morris (2013), 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.

the literature. In Section 3, we present the single agent framework and provide necessary and sufficient conditions for an agent to become more responsive as information quality increases. We extend the analysis to Bayesian games with strategic complementarities in Section 4. In Section 5, we present our four main applications. Section 6 concludes. The proofs for Section 3 are in Appendix A. All the other proofs that are not in the main text are in Appendix B.

# 2 Background on Bayesian Comparative Statics

## 2.1 A Simple Example

A monopolist faces an inverse demand curve $P(q) = 1 - q$ and a cost function $c(\theta, q) = (1 - \theta)\left(q + aq^2/2\right) + q^2/2$, where $q \in [0, 1]$ is the quantity produced. Both $\theta$ and $a$ are cost parameters; $\theta$ is an unobserved random variable that is uniformly distributed on the unit interval, while $a > -1$ is a constant.

The uncertainty about $\theta$ captures the monopolist's uncertainty about her marginal cost. A higher $\theta$ is associated with a lower intercept for the marginal cost curve. The parameter $a$ captures additional uncertainty about the slope of the marginal cost curve: When $a = 0$, there is no additional uncertainty about the slope. When $a > 0$, higher $\theta$ is associated with a flatter marginal cost curve while the opposite holds when $a < 0$.

The monopolist observes a signal $\tilde{s}^\rho$ whose realization $s$ matches the realized cost parameter ($s = \theta$) with probability $\rho \in [0, 1]$ 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 $\rho$: the signal is uninformative when $\rho = 0$ and fully revealing when $\rho = 1$.

*How does the quality of information $\rho$ affect consumer surplus?* The quality of the signal affects the monopolist's production decision which in turn affects consumer welfare. Thus, in order to identify the welfare effects, we must first answer: *How does the quality of information $\rho$ affect the monopolist's production decision?*

From an interim perspective, a monopolist that observes a signal realization $s$ when the

signal quality is $\rho$ optimally produces:

$$q^M(s;\rho,a) = \frac{\rho s + (1-\rho)\mathbb{E}[\theta]}{3 + a(1-\rho s - (1-\rho)\mathbb{E}[\theta])}.$$

Since $q^M$ depends on the signal realization $s$, from an ex-ante perspective, the optimal quantity is a random variable with distribution $H(q;\rho,a) = \mathbb{P}(\{s : q^M(s;\rho,a) \leq q\})$— the probability that the monopolist produces at most $q$ units of output given her signal quality $\rho$. Our goal in this paper is to characterize how $H(\cdot;\rho,a)$ changes when information quality $\rho$ increases.

Let us consider the case when $a = 0$ which simplifies the monopolist's production decision to

$$q^M(s;\rho,0) = \underbrace{\frac{\mathbb{E}[\theta]}{3}}_{\substack{\text{based}\\ \text{on prior}}} + \underbrace{\rho\left(\frac{s-\mathbb{E}[\theta]}{3}\right)}_{\substack{\text{update based}\\ \text{on signal}}}.$$

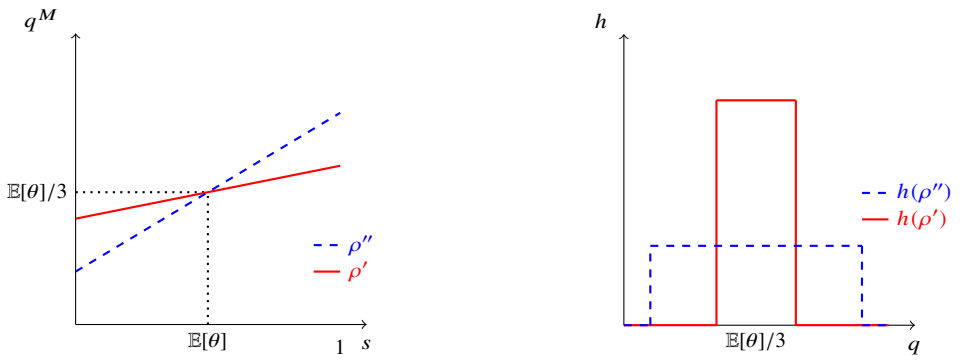
In Figure 1a, we plot $q^M(\cdot;\cdot;0)$ for two different signal qualities, $\rho'$ and $\rho'' > \rho'$. The rotation of the solid line, $q^M(\cdot;\rho',0)$, to the dashed line, $q^M(\cdot;\rho'',0)$, captures the more "dispersed" production decisions when signal quality increases from $\rho'$ to $\rho''$. Intuitively, the monopolist produces more when good news ($s > \mathbb{E}[\theta]$) come from $\rho''$ than when they come from $\rho'$. This is because good news from $\rho''$ are stronger evidence of high values of $\theta$ (lower costs) than good news from $\rho'$. Symmetrically, the monopolist produces less when bad news ($s < \mathbb{E}[\theta]$) come from $\rho''$ than when they come from $\rho'$. The rotation of $q^M(\cdot;\cdot,0)$ induces a mean-preserving spread in the distribution $H(\cdot;\cdot,0)$, as shown by the density function $h$, in Figure 1b.

Observe that the consumer surplus is a convex function of output $CS(q) = \frac{1}{2}q^2$ which implies that consumers benefit from a mean-preserving spread, i.e.,

$$\int CS(q)dH(q;\rho'',0) > \int CS(q)dH(q;\rho',0)$$

for any $\rho'' > \rho'$. Thus, for $a = 0$, we have an answer to our positive and normative questions:

**Claim 1** *In a monopoly with linear demand and quadratic cost with uncertainty only about*




(a) Quantity produced

(b) PDF of production decision

Figure 1: Monopolist signal quality and production decision

*the intercept of the marginal cost ($a = 0$), 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.*

For the case when $a \neq 0$, the cost is no longer a quadratic polynomial in $(\theta, a)$. We later show (see Example 1 and also Section 5.1) that an increase in the quality of the signal leads to a non-decreasing-mean spread of actions when $a > 0$ and a non-increasing-mean spread of actions when $a < 0$. Moreover, the result that the social value of information exceeds the monopolist's private value of information still holds when $a > 0$, while the result is ambiguous for $a < 0$.

These positive and normative results, however, make heavy use of the closed-form solution to $q^M(\cdot; \rho, a)$, which depends on the functional form for profits 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 5.1 to characterize the environments where a Pigouvian subsidy to information is desirable in a monopolistic market.

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

## 2.2 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 $\rho$ induces a distribution of optimal actions $H(\rho)$. For two information structures $\rho''$ and $\rho'$, we say the agent is more *responsive with a higher mean* under $\rho''$ than $\rho'$ if $H(\rho'')$ dominates $H(\rho')$ in the increasing convex order. Alternatively, we say the agent is more *responsive with a lower mean* under $\rho''$ than $\rho'$ 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 $\rho''$ and $\rho'$, we say $\rho''$ dominates $\rho'$ in the supermodular stochastic order if, loosely speaking, the signals from $\rho''$ are more correlated with the state of the world than are the signals from $\rho'$.

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 $\rho''$ than under $\rho'$ if $\rho''$ dominates $\rho'$ 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 $\rho''$ than under $\rho'$, then $\rho''$ necessarily dominates $\rho'$ 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

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

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.

### 2.2.1 Examples and Applications

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 5. As an application of the comparative statics in single-agent decision problems, we reconsider the monopolist example from Section 2.1 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 application comparing two different games of information acquisition: one in which information acquisition is a covert activity (a player cannot observe the quality of her opponents information) and another in which information acquisition is overt. The analysis is formally equivalent to the process of entry accommodation in oligopolistic markets where an incumbent can invest in information acquisition. The difference between the overt and covert demands for information is the indirect effect of information on the incumbent's profit through the induced behavior of the entrant (the value of transparency). We characterize the value of transparency depending on the entrant's responsiveness to the incumbent's information and the sign of the externality imposed on the incumbent by the entrant's responsiveness.

## 2.3 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

[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 $\rho$. We instead provide comparative statics results for the entire mapping $q^M(\cdot;\rho)$ as a function of $\rho$.

(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 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, dating back to Radner (1962) and Hirshleifer (1971). More recently Morris and Shin (2002) fostered renewed interest and Angeletos and Pavan (2007) gave a characterization of the equilibrium and the efficient use of information. Finally, Ui and Yoshizawa (2015) provide necessary and sufficient conditions for welfare to increase with public or private information in quadratic games with normally distributed public and private signals.

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] We instead characterize how increasing information (changes in distributions over beliefs) affects the induced distribution over optimal actions, and we provide a duality between the order in the distribution of actions and the information order. Moreover, in the application to games, Jensen only considers exogenous changes to the distribution of independent private types,

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

while we allow for a richer environment.

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 cumulative 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.

### 2.3.1 Applications

We contribute to the literature on information disclosure (Rayo and Segal, 2010) and Bayesian persuasion (Kamenica and Gentzkow, 2011). We depart from common restrictions in payoffs where the set of actions is discrete or the receiver only cares about the posterior mean.[7] Instead, we restrict the preferences of the receiver to the class that allows unambiguous Bayesian comparative statics (Theorem 1) and we characterize the minimal and maximal levels of conflict between a sender and a receiver, conditions under which extremal disclosure of information is optimal. More details on related papers can be found in subsection 5.2.

We also contribute to the industrial organization literature on information sharing in oligopoly surveyed by Raith (1996) and recently touched upon in Angeletos and Pavan (2007), Bergemann and Morris (2013) and more directly addressed in Myatt and Wallace (2015). We explore the robustness of the results to the assumption of quadratic economies. More details on related papers can be found in subsection 5.3.

In the last application, we study one sided information acquisition and compare the overt and covert demands for information. 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

[7]For example, these conditions are used in Rayo and Segal (2010), Gentzkow and Kamenica (2016), Kolotilin et al. (2017), Dworczak and Martini (2018).

acquisition. Although complementarities are important for our analysis, none of these papers studies the difference between the overt and covert demands for information.

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 5.4.

# 3 Single-agent Model

## 3.1 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 $x_i'' > x_i'$ implies $g(x_i'', x_{-i}) \geq g(x_i', x_{-i})$ for all $x_{-i} \in X_{-i}$. 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 $j \neq i$.

## 3.2 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 $\theta$. Let $\Delta(\Theta)$ denote the set of all Borel probability measures on $\Theta$. 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 $\theta$, and twice differentiable in $a$,

(A.2) for all $\theta \in \Theta$, $u(\theta, \cdot)$ is strictly concave in $a$,

(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.[8]

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 $\mu_2$ first-order stochastically dominates $\mu_1$ (Athey, 2002).[9]

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 $\rho$ is an index that is useful when comparing different signal structures.

We denote the marginal distribution of $\tilde{\theta}$ by $F_\Theta(\cdot; \rho) : \Theta \to [0, 1]$. For Bayesian rationality to hold (Kamenica and Gentzkow, 2011), we assume that any information structure $\Sigma_\rho$ 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.

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 $f_S$.[10]

[8] In Section 7.2.1, we discuss the difficulties that arise when some of these assumptions are violated.

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

[10] The assumption is without loss of generality: we can apply the integral probability transform to any random

## 3.3 Order 1: Actions

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

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

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

for all $x \in \mathbb{R}$. Alternatively, we say that $a(\rho'')$ dominates $a(\rho')$ in the *decreasing convex order* if

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

for all $x \in \mathbb{R}$.[11] 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 $a(\rho'')$*

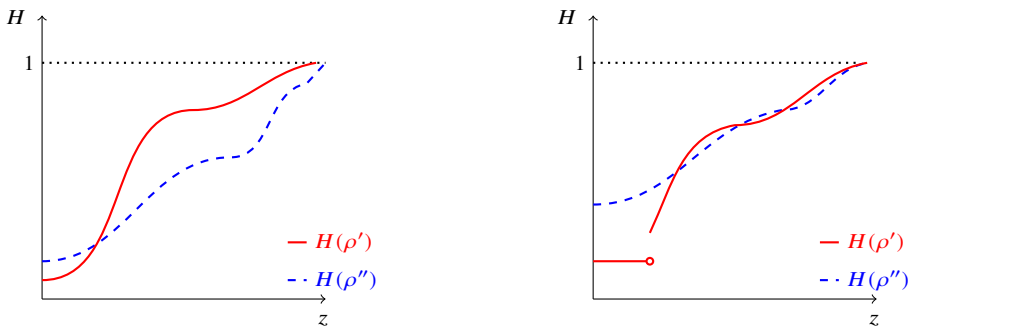
signal $\tilde{s}$ with a continuous marginal distribution and create a new signal which is uniformly distributed on the unit interval. If the marginal distribution of $\tilde{s}$ is discontinuous at $\tilde{s} = s^*$ with $F_S(s^*;\rho) = q$, then, as noted by Lehmann (1988), we can construct a new 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.

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

*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****.*

Figure 2 plots the distribution over actions induced by two information structures $\Sigma_{\rho''}$ and $\Sigma_{\rho'}$. In Figure 2a, the area between the $y$-axis and $H(\cdot;\rho'')$ (the dashed curve) is bigger than the area between the $y$-axis and $H(\cdot;\rho')$ (the solid curve) which implies that $\Sigma_{\rho''}$ induces optimal actions with a higher mean than $\Sigma_{\rho'}$. Furthermore, integrating $H(\cdot;\rho'') - H(\cdot;\rho')$ right to left always yields a negative value which implies responsiveness with a higher mean. In contrast, in Figure 2b, the area between the $y$-axis and $H(\cdot;\rho'')$ (the dashed curve) is smaller than the area between the $y$-axis and $H(\cdot;\rho')$ (the solid curve) which implies that $\Sigma_{\rho''}$ induces optimal actions with a lower mean than $\Sigma_{\rho'}$. Furthermore, integrating $H(\cdot;\rho'') - H(\cdot;\rho')$ left to right always yields a positive value which 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

## 3.4 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 shift in beliefs. This assumption is weaker than the monotone likelihood

ratio property commonly assumed in settings with complementarities (Milgrom and Weber, 1982; Athey, 2002).

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

**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 $\rho'' \geq_{spm} \rho'$, if $F(\theta, s;\rho'') \geq F(\theta, s;\rho')$ for all $(\theta, s) \in \Theta \times S$. (Tchen, 1980).*

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 stronger evidence of a low state (than a signal $\tilde{s} \leq s$ from $\Sigma_{\rho'}$). Thus, $\mathbb{P}(\tilde{\theta} \leq \theta|\tilde{s} \leq s;\rho'') \geq \mathbb{P}(\tilde{\theta} \leq \theta|\tilde{s} \leq s;\rho')$. Without loss of generality, 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 2.1 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 $\theta_0$ and variance $\sigma^2$, and have a correlation coefficient of $\rho \in [0, 1]$. In both cases, $\rho'' \geq_{spm} \rho'$ if $\rho'' > \rho'$.

In Appendix B (Section 8), 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.

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

## 3.5 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 increasing differences in $(\theta; a)$ that increase with $a$, and a marginal utility that diminishes at a *diminishing rate* for every state $\theta$. 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

$(i)$ 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 increasing differences in $(\theta; a)$ that decrease with $a$, and a marginal utility that diminishes at an *accelerating rate*. 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 $u \in \mathcal{U}^I$ $\left(resp., u \in \mathcal{U}^D\right)$ 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

[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.

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 $u \in \mathcal{U}^I$ (resp.,$u \in \mathcal{U}^D$), 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 $\lambda \in [0,1]$*

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

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

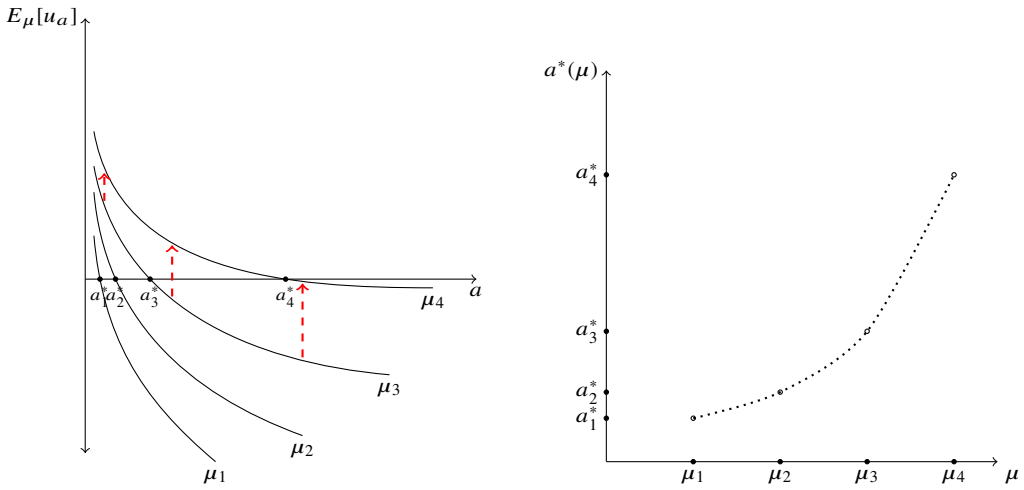
Henceforth, we focus on payoffs in $\mathcal{U}^I$ but the arguments we provide can be symmetrically applied to payoffs in $\mathcal{U}^D$.

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 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 $\mu_n$ 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, $a_4 \geq a_3 \geq a_2 \geq a_1$.

ID of $u_a(\theta, a)$ implies that the gap between the expected marginal utilities of $\mu_{n+1}$ and $\mu_n$ is widening as the action increases (the height of the dashed 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 $\mu_2$ to $\mu_3$ is larger than her benefit from increasing $a_1$ to $a_1 + \epsilon$ when beliefs increase from $\mu_1$ to $\mu_2$, and so on.

In contrast, concavity of $u(\theta, a)$ in $a$ implies that, for any fixed belief, the agent's benefit from increasing $a_2$ to $a_2 + \epsilon$ is less than her benefit from increasing $a_1$ to $a_1 + \epsilon$, 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

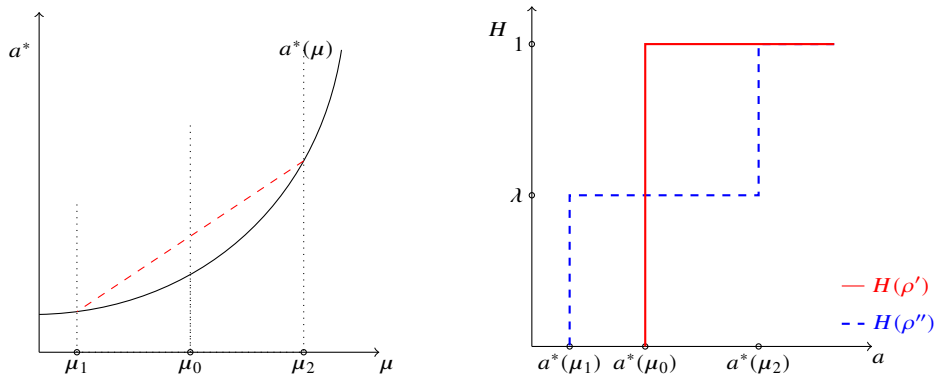
Figure 3: Convexity for $u \in \mathcal{U}^I$

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 $\lambda$ and $a^*(\mu_2)$ with probability $1 - \lambda$. Since posteriors are derived by Bayesian updating, $\mu_0 = \lambda\mu_1 + (1 - \lambda)\mu_2$. From Proposition 1, $u \in \mathcal{U}^I$ implies that $\lambda a^*(\mu_1) + (1 - \lambda)a^*(\mu_2) \geq a^*(\lambda\mu_1 + (1 - \lambda)\mu_2) = a^*(\mu_0)$, 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 implies that the agent is more responsive with a higher mean under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$.

**Corollary 1** *Let $\Sigma_{\rho''}$ be an information structure that satisfies* (A.5). *Let $\Sigma_{\rho'}$ be any* garbling




(a) Optimal action (b) Induced distribution

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

*of $\Sigma_{\rho''}$. If an agent has utility $u \in \mathcal{U}^I$ (resp., $u \in \mathcal{U}^D$), 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]

**Remark 2** *Since $\Theta$ is an ordered set, the full information structure (the signal that perfectly reveals the state) induces posteriors that trivially satisfy (A.5). Furthermore, any other information structure is a garbling of the full information structure. Hence, when $u \in \mathcal{U}^I$ (resp., $u \in \mathcal{U}^D$), Corollary 1 implies that the agent's actions are the most dispersed with the highest (resp., lowest) mean under the full information structure.*

**Remark 3** *Whenever $u \notin \mathcal{U}^I$, there exist beliefs $\mu_1, \mu_2 \in \Delta(\Theta)$ with $\mu_2 \succeq_{FOSD} \mu_1$ and a $\lambda \in (0,1)$ for which Proposition 1 is violated. Consider a prior $\mu^0 = \lambda\mu_1 + (1-\lambda)\mu_2$, an uninformative structure $\Sigma_{\rho'}$ that induces the prior with probability 1, and a more informative structure $\Sigma_{\rho''}$ that induces posteriors $\mu_1$ and $\mu_2$ with probabilities $\lambda$ and $1-\lambda$ respectively. Then $\rho'' \succeq_{spm} \rho'$ but the agent is not more responsive with a higher mean under $\Sigma_{\rho''}$. In this*

[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 $(\theta, s) \in \Theta \times S$.

*sense, the class of preferences $\mathcal{U}^1$ is not only sufficient but also necessary for responsiveness with a higher mean. We present such an example in Section 7.2.1.*

# 4 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)).

## 4.1 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 $F_{\Theta_i}$ be the marginal distribution of $\tilde{\theta}_i$ induced by $\mu^0$. 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 $F_{\Theta_{-i}}(\cdot|\theta_i') \geq_{FOSD} F_{\Theta_{-i}}(\cdot|\theta_i)$,

which is a weaker assumption than affiliation (Milgrom and Weber, 1982).

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 $\theta$, 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 of Gossner (2000), 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 $\Sigma_{\rho_i} \triangleq (S_i, F(\cdot,\cdot;\rho_i))$.[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 $\rho_i$ is an index. For Bayesian rationality to hold, we assume that all information structures induce the same marginal on $\tilde{\theta}_i$ which corresponds to the prior CDF $F_{\Theta_i}$. Furthermore, we assume (WLOG) that all information structures induce the same marginal on $\tilde{s}_i$, denoted by $F_{S_i}$, with a positive and bounded density $f_{S_i}$.

Let $S \triangleq \times_{i\in N} S_i$. We denote the profile of information structures by $\Sigma_\rho \triangleq \left(\Sigma_{\rho_1}, \ldots, \Sigma_{\rho_n}\right)$. A profile $\Sigma_\rho$ 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 $(\theta, s) \in \Theta \times S$, $\boldsymbol{F}(s|\theta;\rho) = \prod_{i\in N} F(s_i|\theta_i;\rho_i)$.

(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 $i$ cannot directly learn about $(\tilde{\theta}_{-i}, \tilde{s}_{-i})$. Assumption (A.12) is an extension of (A.5) and implies that higher signal realizations lead to a first-order

[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).

stochastic shift 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 $\Sigma_{\rho_i}$. Then the players participate in the basic game $G$ by simultaneously choosing an action.

Momentarily ignoring existence issues, let $a^\star(\rho) = (a_1^\star(\rho), a_2^\star(\rho), \ldots, a_n^\star(\rho))$ 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

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

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 profile of information structures 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.

[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)).

## 4.2 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.

## 4.3 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 $i \in N$.*

## 4.4 Preferences and Main Result for Games

Let $\Gamma^I$ 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 $\Gamma^I$ are linked to responsiveness with a higher mean.[17]

Let $\Gamma^D$ 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 $\Gamma^D$ 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 $u^i \in \Gamma^I$ (resp., $u^i \in \Gamma^D$) 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 Appendix B (Section 8). 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.

   - Suppose $\rho_i'' \geq_{spm} \rho_i'$, 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.

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

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

   - Suppose $\rho_j'' \geq_{spm} \rho_j'$ 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 $\alpha_j''$ and $\alpha_j'$ such that $\alpha_j''$ dominates $\alpha_j'$ in the increasing convex order. Fix the monotone strategies of players $k \neq j, i$. Then player $i$'s best-reply to $\alpha_j''$ dominates her best-reply to $\alpha_j'$ 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.

## 4.5 Examples

While the conditions on payoffs in Theorem 2 might seem abstract, we illustrate with particular examples how they are satisfied naturally.

[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 $\tilde{\theta}_j$.

**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 $\beta_i \in (0,1)$. 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 $u^i \in \Gamma^I$ $(resp., u^i \in \Gamma^D)$ 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 $j \neq i$.

This 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 $h_i(a_{-i}) = \frac{1}{n-1}\sum_{j\neq i} a_j$.

A more general formulation encompassing all quadratic games (Angeletos and Pavan, 2007; Bergemann and Morris, 2013) is the following:

$$u^i(\theta, a) = -a_i^2 + 2a_i g_i(\theta, a_{-i}) + f_i(\theta, a_{-i}),$$

where $g$ satisfies the same conditions above and $f_i : \Theta \times A_{-i} \to \mathbb{R}$ is bounded and measurable but otherwise is left free. While $f_i(\theta, a_{-i})$ is not important from a positive perspective, it is important for welfare analysis.

The canonical quadratic games satisfy $u^i \in \Gamma^I \cap \Gamma^D$ for every agent. As a corollary of 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, and without the assumption of symmetric preferences. However, we do have to take into account the possibility of multiple equilibria (see Theorem 2).

**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).

This 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 $v_i(\theta) = v_i \prod_{j=1}^n \theta_j$. A player's productivity is given by $\theta_i a_i$ where $\theta_i$ 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 $v_i > 0$. 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 $\theta_i$ 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 above model in 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 about the state and her match partner from $x_{i1t} = \theta_{it} + \varepsilon_{i1t}$, $x_{i2t} = x_{j1t} + \varepsilon_{i2t}$, $x_{i3t} = x_{j2t} + \xi_t + \varepsilon_{i3t}$, 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 $\xi_t$, 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 $\xi_t$. 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 allow for higher average investment and capital accumulation.

# 5 Applications

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

## 5.1 Application: Pigouvian Subsidies and Monopoly Production

In the example from Section 2.1, 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 $\theta$ 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 $\theta \in \Theta$.[20]

Prior to making any production decisions, the monopolist can acquire information from $\mathcal{P}$, a set of information structures that satisfy (A.5). We assume that 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*

[20] A sufficient condition is that $c(\theta, q)$ is convex in $q$ and $P(q)$ is decreasing and concave in $q$.

*conditions does the social planner demand more information than the monopolist?*[21]

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

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

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

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

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 has 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.*

**Remark 4** *Sufficient conditions for $-qP''(q)/P'(q) \leq 1$ and $\pi \in \mathcal{U}^I$ are that the inverse demand $P(q)$ is linear and $-c(q,\theta) \in \mathcal{U}^I$.*

*In relation to Section 2.1, when $a \geq 0$ we have $-c(q,\theta) \in \mathcal{U}^I$.*[22] *Hence, given that $P(q)$ is linear, we have $\pi \in \mathcal{U}^I$ and Proposition 2 holds. When $a < 0$, we instead have $\pi \in \mathcal{U}^D \backslash \mathcal{U}^I$ and the ranking of the social value and the private value of information is ambiguous.*

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. In other words, the social planner is "more risk-loving" than the monopolist, i.e., consumers (and the planner) benefit when the

[21] 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.

[22] When $a = 0$, $-c(q,\theta) \in \mathcal{U}^I \cap \mathcal{U}^D$.

monopolist becomes more responsive with a higher mean as quality of information increases. From Theorem 1, we get the desired responsiveness behavior when $\pi \in \mathcal{U}^I$.

## 5.2 Application: Information Disclosure

In the information disclosure game of Rayo and Segal (2010) and 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 only depends on the posterior mean.[23]

We depart from that approach and restrict the preferences of the receiver to the class that allows unambiguous Bayesian comparative statics (Theorem 1), we then characterize the conditions on the preferences of the sender that give maximal or minimal disclosure in two cases: when the sender is restricted to a totally ordered set of signals, and when we allow for complete flexibility in disclosure policies.

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). For the next result we assume that the sender is restricted to $\mathcal{P}$, a set of information structures that satisfy (A.5) (recall that a signal satisfies (A.5) if it generates posteriors that are first order stochastically ordered).

[23]Examples of papers that use these simplifying conditions are Rayo and Segal (2010), Bergemann and Morris (2013), Bergemann et al. (2015), Gentzkow and Kamenica (2016), Kolotilin (2018), Kolotilin et al. (2017), Lipnowkski and Mathevet (2017), Taneva (2017), Li and Norman (2018), Dworczak and Martini (2018).

The sender's problem is given by

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

**Proposition 3** *Assume $v(\theta, a)$ satisfies increasing differences (resp., decreasing differences) in $(\theta; a)$, and 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 $\rho'' \succeq_{spm} \rho'$.*

Therefore, when the sender is constrained to disclosure policies in $\mathcal{P}$, the proposition gives conditions under which maximal or minimal information disclosure obtains.

Proposition 3 can be seen as characterizing minimal and maximal conflict between a sender and a receiver: if their desire to correlate actions and states goes in the same (opposite) direction and the sender likes (dislikes) dispersion of the actions there will be full (null) disclosure.[24]

**Remark 5** *The conditions in Proposition 3 are also related to Remark 1 in Kamenica and Gentzkow (2011) as they are sufficient conditions for $\hat{v}(\mu) = \int_\Theta v(\theta, a^*(\mu))\mu(d\theta)$ to be a "convex" or "concave" function over beliefs that are first order stochastically ranked —a weakening of convexity and concavity respectively.*

When full information revelation is a policy available to the sender, we provide a result that holds independently of (A.5) and the supermodular stochastic order. Under the corresponding

[24]Note that whenever the sender violates one of the conditions *i,ii,iii*, there is a receiver satisfying them for which there is interior disclosure. In that sense the conditions for the sender are also necessary.

conditions on payoffs from Proposition 3, full information revelation is the optimal persuasion policy over all unrestricted persuasion policies (all information structures).

**Theorem 3** *Assume $v(\theta, a)$ satisfies increasing 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 in* $a$*,*

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

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

*Then full information revelation is the optimal disclosure policy among all possible signals.*

Theorem 3 follows from a similar reasoning as the proofs of Proposition 1 and Corollary 1: the full information structure is Blackwell more informative than any other signal, and trivially induces posteriors that satisfy A.5 (because $\Theta$ is an ordered set). Thus, when the sender can use any information structure, Corollary 1 and the conditions in Theorem 3 imply that there is minimal conflict between the sender and the receiver, establishing the optimality of full disclosure.

Note also that when the sender has full flexibility, some of the feasible signals do not satisfy (A.5) and cannot be ordered by the supermodular order. However, in the special case when there are only two possible states of the world, Proposition 3 implies full or null disclosure. The reason is that (A.5) is always satisfied when there are only two possible states, which implies that any information structure is both dominated by the full information structure and dominates the null information structure in the supermodular stochastic order.

**Example 5 (Portfolio Agency Problem)**

To illustrate the value in Proposition 3, consider 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.

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 $G_\theta$ with density $g_\theta$. The state of the world $\theta$ captures the underlying distribution of returns of the risky asset. Suppose for $\theta'' > \theta'$,

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

for all $z \in [\underline{x}, \bar{x}]$, with equality when $z = \bar{x}$.[25]

Suppose the financial adviser gets a share $\pi \in (0, 1)$ of the return on the risky asset, where $\pi$ represents 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

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

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 $\theta$ 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 to use the concavification approach.

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

[25] Rothschild and Stiglitz (1971) 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 embed their agent into the portfolio management problem.

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[26]

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

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. Additionally, from Theorem 3, full information revelation is the optimal persuasion policy over all information structures.

## 5.3 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 $G \triangleq \left(\{A_i, u^i\}_{i=1,2}, \mu^0\right)$ 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 $\Sigma_\rho \triangleq \left(\Sigma_{\rho_1}, \Sigma_{\rho_2}\right) \in \mathcal{P}_1 \times \mathcal{P}_2$ 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.[27] Each choice of $\Sigma_{\rho_2}$ defines a Bayesian game $\mathcal{G}_\rho \triangleq (\Sigma_{\bar{\rho}_1}, \Sigma_{\rho_2}, G)$ as outlined in Figure 5.

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

[27] 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 5.2. The only difference here is that both the sender and the receiver take an action.

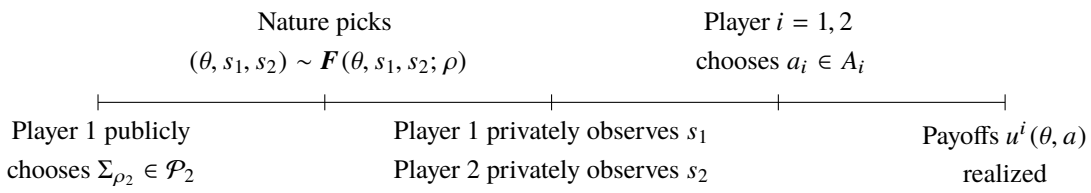



Figure 5: Timing of information sharing game

For each Bayesian game $\mathcal{G}_\rho$, we assume that the players can coordinate on the maximal monotone BNE $a^\star(\rho) = \left(a_1^\star(\rho), a_2^\star(\rho)\right)$ 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

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

*How much information, if any, would player 1 want to share with player 2?* The question of 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.[28] 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.[29] We establish that the full-information sharing result is robust to different specifications of information structures and payoffs. Moreover, we show that it is a dominant strategy and, therefore, the unique Nash outcome of the information sharing game.

## 5.4 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

[28] 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 $\beta_{ij} \geq 0 > \beta_{ii}$ $\forall j \neq i$.

[29] 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.

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 acquisition is no 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, and advertising (Bulow et al., 1985).

We focus our analysis on entry accommodation.[30] We 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.

### 5.4.1 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

[30] 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.

$i = 1, 2$ 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 $\Sigma_{\rho_2}$. 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.[31] 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 knowledge, i.e., both players know the Bayesian game is $\mathcal{G}_{\rho'} \triangleq (\Sigma_{\rho_1'}, \Sigma_{\rho_2}, G)$. Let $(a_1^\star(\rho'), a_2^\star(\rho'))$ 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 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 $(a_1^\star(\rho''), a_2^\star(\rho''))$.

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).

[31] 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.

Given actual first stage outcome $\rho$ 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

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

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

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

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 $\rho$ 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, $\Sigma_{\rho_1}$, 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 $\Sigma_{\rho_1}$. Instead, it captures the indirect effects of information stemming from a change in player 2's beliefs and, therefore, her strategic response.[32]

### 5.4.2 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

[32]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 $\Sigma_{\rho_1}$ over $\Sigma_{\hat{\rho}_1}$ when player 2 believes that player 1 will acquire $\Sigma_{\rho_1}$. It is straightforward to show that expectations conformity is equivalent to $VT(\rho;\hat{\rho}) + VT(\hat{\rho};\rho) \geq 0$.

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** *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 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.[33] 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).[34] 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 $U_1(\rho^o; \rho^o) - U_1(\rho^o; \rho^c) = VT(\rho^o; \rho^c) \geq 0 \Leftrightarrow \rho_1^o \succeq_{spm} \rho_1^c$. ■

[33] 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 $\kappa$ 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.

[34] 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.

### 5.4.3 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 4** *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 4. 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

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

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

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

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

The conditions in Theorem 4 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, $\text{sign}(u^1_{a_2a_2})$.

For example, suppose condition $i.$ of Theorem 4 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 4 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 $\Sigma_\rho$. 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 4 respectively. For instance, the fifth row of Figure 6 states that if a change from $\Sigma_{\hat{\rho}_1}$ to $\Sigma_{\rho_1}$ 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.

### 5.4.4 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.

| $a_2(\rho) - a_2(\hat{\rho})$ | | Externality | | Transparency |
|---|---|---|---|---|
| responsiveness | mean | $\text{sign}(u^1_{a_2})$ | $\text{sign}(u^1_{a_2a_2})$ | $VT(\rho;\hat{\rho})$ |
| $\nearrow$ | $\nearrow$ | $+$ | $+$ | $+$ |
| $\nearrow$ | $\nearrow$ | $-$ | $-$ | $-$ |
| $\nearrow$ | $\searrow$ | $-$ | $+$ | $+$ |
| $\nearrow$ | $\searrow$ | $+$ | $-$ | $-$ |
| $\nearrow$ | $cst$ | $\cdot$ | $+$ | $+$ |
| $\nearrow$ | $cst$ | $\cdot$ | $-$ | $-$ |
| $\searrow$ | $\nearrow$ | $+$ | $-$ | $+$ |
| $\searrow$ | $\nearrow$ | $-$ | $+$ | $-$ |
| $\searrow$ | $\searrow$ | $-$ | $-$ | $+$ |
| $\searrow$ | $\searrow$ | $+$ | $+$ | $-$ |
| $\searrow$ | $cst$ | $\cdot$ | $-$ | $+$ |
| $\searrow$ | $cst$ | $\cdot$ | $+$ | $-$ |

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

(1985).[35] 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 prices (strategic complements). Let $(a_1^\star(K_1), a_2^\star(K_1))$ 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 $U_1(K_1, a_1^\star(K_1), a_2^\star(K_1))$.

Fudenberg and Tirole (1984) show that the total marginal effect on the incumbent's payoff

[35] 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).

from increasing investment can be decomposed into

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

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 $\rho_1$. The total effect of overtly increasing investment in information from $\Sigma_{\rho_1}$ 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 player 1's payoff increases by her ability to make better informed decisions while holding 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 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 informa-

tion 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).

# 6 Conclusion

We provide a general framework to study changes in equilibria and welfare as the quality of private information increases. The theory has important implications in both 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 order of actions if, and only if, signal quality increases in the information order.

There are positive as well as normative implications. For example, the quality of private information affects price dispersion in industrial economics. In the macroeconomy, it might increase the cross-sectional volatility of investment or aggregate output and it can also induce a higher expected aggregate output.

In welfare analysis, we connect information, actions, and payoffs through the concept of externalities. A monopolist with more precise information about his costs increases both the volatility and the average level of production if he faces a sufficiently convex demand. This has positive externalities on the consumers.

In information disclosure (Bayesian persuasion) games, we characterize the minimal and maximal levels of conflict between a sender and a receiver, conditions under which extremal disclosure of information is optimal. We find that the conditions for full disclosure are also similar to the conditions that imply full information sharing is optimal for competing oligopolists, thereby extending the industrial organization literature to more general environ-

ments.

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. We characterize the value of transparency depending on the entrant's responsiveness to the incumbent's information and the sign of the externality imposed on the incumbent by the entrant's responsiveness.

The theory of Bayesian comparative statics will be useful to generalize many of the insights developed for quadratic economies to a class of payoffs with richer dynamics. One avenue for future research is to study the efficient and equilibrium use of information.[36] Another open question is how a central planner should intervene in markets with uncertain fundamentals and dispersed information in non-linear environments.[37]

More generally, the framework can be applied in information design.[38] Other extensions include studying the comparative statics of welfare and equilibrium outcomes with respect to the quality of public information, exogenous changes to the prior distribution of the market fundamentals, and changes in attitudes towards risk or temporal resolution of uncertainty.

# References

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

George-Marios Angeletos and Jennifer La'O. Sentiments. *Econometrica*, 81(2):739–779, 2013.

George-Marios Angeletos and Jennifer La'O. Optimal monetary policy with informational

[36] Angeletos and Pavan (2007) study the efficient and equilibrium use of information in quadratic economies.

[37] Angeletos and Pavan (2009); Lorenzoni (2010); Angeletos and La'O (2018) study policy making in quadratic economies with dispersed information.

[38] See Bergemann and Morris (2018) for a recent survey on information design.

frictions. *Working paper (MIT)*, 2018. URL https://economics.mit.edu/files/7025.

George-Marios Angeletos and Alessandro Pavan. Efficient use of information and social value of information. *Econometrica*, 75(4):1103–1142, 2007.

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

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

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

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

Coralio Ballester, Antoni Calvó-Armengol, and Yves Zenou. Who's who in networks. the key player. *Econometrica*, 74(5):1403–1417, 2006.

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

Dirk Bergemann and Stephen Morris. Information design: a unified perspective. *Journal of Economic Literature, forthcoming*, 2018.

Dirk Bergemann, Benjamin Brooks, and Stephen Morris. The limits of price discrimination. *American Economic Review*, 105(3):921–57, March 2015.

David Blackwell. 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.

David Blackwell. Equivalent comparisons of experiments. *The Annals of Mathematical Statistics*, 24(2):265–272, 06 1953.

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

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

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

Luca Colombo, Gianluca Femminis, and Alessandro Pavan. Information acquisition and welfare. *The Review of Economic Studies*, 81(4):1438–1483, 2014.

Piotr Dworczak and Giorgio Martini. The simple economics of optimal persuasion. *Journal of Political Economy, forthcoming*, 2018.

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

Esther Gal-Or. Information sharing in oligopoly. *Econometrica*, pages 329–343, 1985.

Matthew Gentzkow and Emir Kamenica. A Rothschild-Stiglitz approach to Bayesian persuasion. *American Economic Review P&P*, 106(5):597–601, May 2016.

Olivier Gossner. Comparison of information structures. *Games and Economic Behavior*, 30: 44–63, 2000.

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

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

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

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

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

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

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

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

Miles S Kimball. Precautionary saving in the small and in the large. *Econometrica*, pages 53–73, 1990.

Anton Kolotilin. Optimal information disclosure: A linear programming approach. *Theoretical Economics*, 13(2):607–635, 2018.

Anton Kolotilin, Tymofiy Mylovanov, Andriy Zapechelnyuk, and Ming Li. Persuasion of a privately informed receiver. *Econometrica*, 85(6):1949–1964, 2017.

René Leal-Vizcaíno and Teddy Mekonnen. Information sharing in supermodular games. *Working paper (Caltech)*, 2018. URL https://sites.google.com/view/teddymekonnen/research?authuser=0.

E. L. Lehmann. Comparing location experiments. *The Annals of Statistics*, 16(2):521–533, 06 1988.

Fei Li and Peter Norman. Sequential persuasion. *Working paper (UNC)*, 2018. URL https://lifei.weebly.com/research.html.

Elliot Lipnowkski and Laurent Mathevet. Simplifying Bayesian persuasion. *Working paper (NYU)*, 2017. URL http://www.laurentmathevet.com/wp-content/uploads/2017/03/BPCover.pdf.

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

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

Robert E Lucas Jr. Expectations and the neutrality of money. *Journal of Economic Theory*, 4(2):103–124, 1972.

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

Paul Milgrom and John Roberts. Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica*, pages 1255–1277, 1990.

Paul Milgrom and John Roberts. Comparing equilibria. *American Economic Review*, pages 441–459, 1994.

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

Paul R Milgrom and Robert J Weber. A theory of auctions and competitive bidding. *Econometrica*, 50(5):1089–1122, September 1982.

Stephen Morris and Hyun Song Shin. Social value of public information. *American Economic Review*, 92(5):1521–1534, 2002.

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

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

David P. Myatt and Chris Wallace. 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.

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

William Novshek and Hugo Sonnenschein. Fulfilled expectations Cournot duopoly with information acquisition and release. *The Bell Journal of Economics*, pages 214–218, 1982.

Alessandro Pavan and Xavier Vives. 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.

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

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

R. Radner. Team decision problems. *The Annals of Mathematical Statistics*, 33(3):857–881, 09 1962.

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

Luis Rayo and Ilya Segal. Optimal information disclosure. *Journal of Political Economy*, 118 (5):949–987, 2010.

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

René L Schilling. *Measures, integrals and martingales*. Cambridge University Press, 2005.

Moshe Shaked and J George Shanthikumar. *Stochastic orders*. Springer Science & Business Media, 2007.

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

Ina Taneva. Information design. *Working paper (University of Edinburgh)*, 2017. URL https://sites.google.com/site/itaneva13/research.

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

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

Jean Tirole. Cognitive games and cognitive traps. *Working paper (TSE)*, 2015. URL https://www.tse-fr.eu/sites/default/files/TSE/documents/doc/by/tirole/cognitive_games_210715.pdf.

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

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

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

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

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

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

Zhixiang Zhang. Comparison of information structures with infinite states of nature. *dissertation submitted to Johns Hopkins University (Department of Economics)*, 2008.

# 7 Appendix A

## 7.1 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 any increasing convex function $\varphi : \mathbb{R} \to \mathbb{R}$,*

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

*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 any decreasing convex function $\phi : \mathbb{R} \to \mathbb{R}$,*

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

*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* $\Sigma_{\rho'}$, $\rho'' \geq_{spm} \rho'$ *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* $\int_{x'}^{x''} g(t)h(t)dt \geq g(x'')\int_{x'}^{x''} h(t)dt$

## 7.2 Single-Agent

**Proof of Theorem 1**

*Proof.* ( $\Longrightarrow$ ) The payoff $u(\theta, a)$ satisfies ID in $(\theta; a)$ and the information structure $\Sigma_\rho$ has the property that $s > s'$ implies $\mu(\cdot|s;\rho) \geq_{FOSD} \mu(\cdot|s';\rho)$. 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.[39] For any two information structures $\rho'' \geq_{spm} \rho'$ 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 $u \in \mathcal{U}^1$, then $u_a(\theta, a)$ is convex in $a$ for all $\theta$. Thus,

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

and

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

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

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

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

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 $\theta$ 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 $t \in [0, 1]$,

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

where the last inequality follows from Lemma 2.

($\Longleftarrow$) By definition, if $\rho'' \not\succeq_{spm} \rho'$, 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, $u \in \mathcal{U}^I \bigcap \mathcal{U}^D$. For any given $\Sigma_\rho$,

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

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

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

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

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

which is independent of $\rho$. Thus,

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

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 $u \in \mathcal{U}^I$.

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

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 $\mu_2 \geq_{FOSD} \mu_1$ 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 $\theta$. The last inequality then follows from the definition of first-order stochastic dominance. Since the marginal value of $a_\lambda$ is non-positive at $\mu_\lambda$, we must have $a^*(\mu_\lambda) \leq a_\lambda$. A symmetric argument establishes that if $u \in \mathcal{U}^D$, then $a^*(\mu_\lambda) \geq a_\lambda$. ■

**Proof of Corollary 1**

*Proof.* We prove the case of increasing mean, the other case is analogous. We begin with a lemma.

**Lemma (A)** *Suppose $u \in \mathcal{U}^I$. For any finite sequence of beliefs $\{\mu_i\}_{i=1}^n$ with $\mu_n \succeq_{FOSD} \mu_{n-1} \succeq_{FOSD} \ldots \succeq_{FOSD} \mu_1$, and any sequence of weights $\{\lambda_i\}_{i=1}^n$ with $\lambda_i \in [0,1]$ and $\sum_{i=1}^n \lambda_i = 1$, we have*

$$a^*(\sum_{i=1}^n \lambda_i \mu_i) \leq \sum_{i=1}^n \lambda_i a^*(\mu_i)$$

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

*Proof.*

$$\begin{aligned} a^*\left(\lambda_1\mu_1 + \sum_{i=2}^n \lambda_i\mu_i\right) &\leq \lambda_1 a^*(\mu_1) + \left(\sum_{i=2}^n \lambda_i\right) a^*\left(\sum_{i=2}^n \frac{\lambda_i}{\sum_{k=2}^n \lambda_k}\mu_i\right) \\ &\leq \sum_{i=1}^n \lambda_i a^*\left(\mu_i\right) \end{aligned}$$

The first line follows from Proposition 1 using the property that first order stochastic dominance is preserved under convex combinations. The second line follows by induction.

■

Let $\Sigma_{\rho''}$ be an information structure that induces posteriors $\{\mu_i\}_{i=1}^n$ with corresponding probabilities $\{\tau_i^{\rho''}\}_{i=1}^n$ such that $\mu_i \succeq_{FOSD} \mu_j$ whenever $i > j$. Let $\Sigma_{\rho'}$ be another information structure that induces posteriors $\{\nu_k\}_{k=1}^m$ with corresponding probabilities $\{\tau_k^{\rho'}\}_{k=1}^m$.

Assume $\Sigma_{\rho'}$ be a garbling of $\Sigma_{\rho''}$ so that for each $k = 1, \ldots, m$, there exist weights $\{\lambda_i^k\}_{i=1}^n$ such that (i) $\lambda_i^k \in [0,1]$, (ii) $\sum_{i=1}^n \lambda_i^k = 1$, and (iii) $\nu_k = \sum_{i=1}^n \lambda_i^k \mu_i$. Furthermore, for each $i = 1, \ldots, n$, $\tau_i^{\rho''} = \sum_{k=1}^m \lambda_i^k \tau_k^{\rho'}$.

To show that the agent is more responsive with a higher mean under $\Sigma_{\rho''}$ than $\Sigma_{\rho'}$, take

any increasing and convex function $\varphi : \mathbb{R} \to \mathbb{R}$. Then

$$\begin{aligned}
\int \varphi(z) dH(z;\rho') &= \sum_{k=1}^{m} \varphi\left(a^*\left(\nu_k\right)\right) \tau_k^{\rho'} \\
&= \sum_{k=1}^{m} \varphi\left(a^*\left(\sum_{1=1}^{n} \lambda_i^k \mu_i\right)\right) \tau_k^{\rho'} \\
&\leq \sum_{k=1}^{m} \varphi\left(\sum_{1=1}^{n} \lambda_i^k a^*\left(\mu_i\right)\right) \tau_k^{\rho'} \\
&\leq \sum_{1=1}^{n} \sum_{k=1}^{m} \lambda_i^k \varphi\left(a^*\left(\mu_i\right)\right) \tau_k^{\rho'} \\
&= \sum_{1=1}^{n} \varphi\left(a^*\left(\mu_i\right)\right) \tau_i^{\rho''} = \int \varphi(z) dH(z;\rho'')
\end{aligned}$$

where the first inequality follows from Lemma A and monotonicity of $\varphi$, and the second inequality follows from the convexity of $\varphi$. The desired result follows by the characterization of responsiveness with a higher mean in Lemma 1.

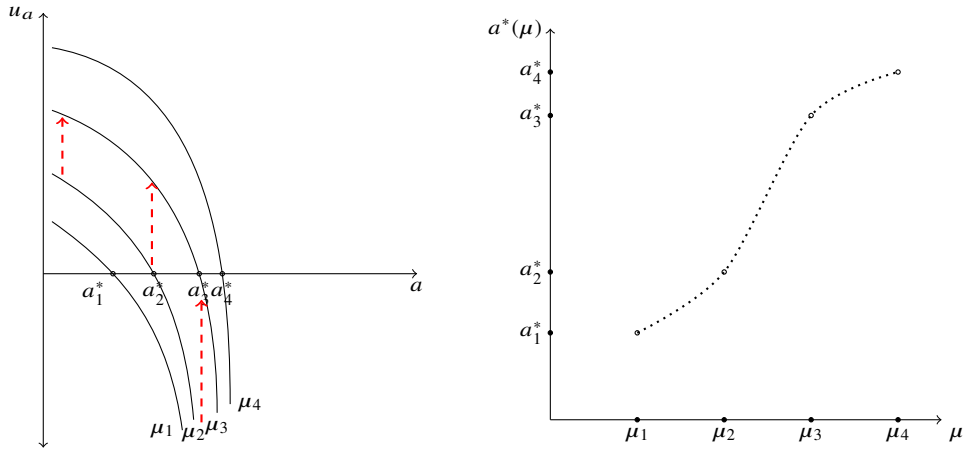
The above argument can be extended to the case of infinite posteriors following the methods in Zhang (2008). ■

### 7.2.1 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 $\mu_4 \geq_{FOSD} \mu_3 \geq_{FOSD} \mu_2 \geq_{FOSD} \mu_1$.

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 $\mu_{n+1}$ lies above the expected marginal utility of $\mu_n$. Thus, $a_{n+1} \geq a_n$. Furthermore, $u_a(\theta, a)$ also satisfies ID in $(\theta; a)$—the height of the dashed 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.



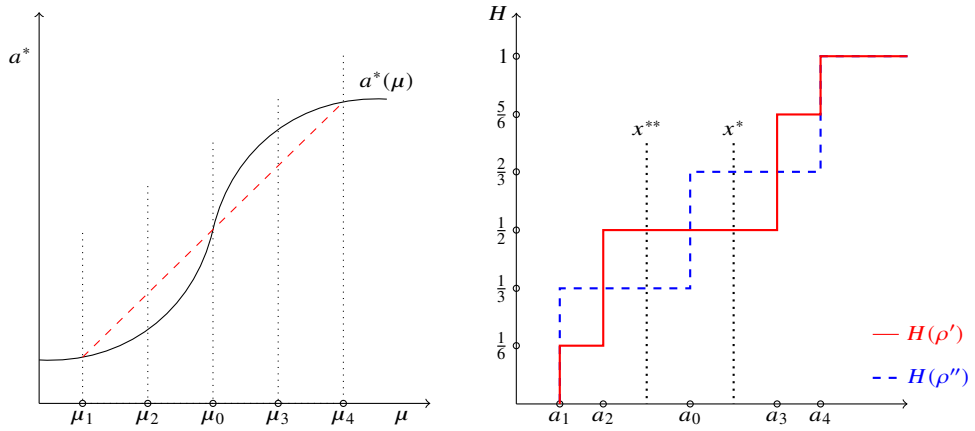

(a) Marginal utilities (b) Optimal action

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 $\{\mu_1, \mu_0, \mu_4\}$ with probabilities $\{1/3, 1/3, 1/3\}$ such that $\mu_4 \succeq_{FOSD} \mu_0 \succeq_{FOSD} \mu_1$. Let $\Sigma_{\rho'}$ induce posteriors $\{\mu_1, \mu_2, \mu_3, \mu_4\}$ with probability $\{1/6, 1/3, 1/3, 1/6\}$ where $\mu_2 = 0.5\mu_1 + 0.5\mu_0$ and $\mu_3 = 0.5\mu_4 + 0.5\mu_0$. 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, this corresponds to the point of intersection of the dashed line and the solid curved line at $\mu_0$. Figure 8b maps the distribution over optimal actions. $\Sigma_{\rho''}$ induces the dashed line while $\Sigma_{\rho'}$ induces the solid line.

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 $x^* \in (a^*(\mu_0), a^*(\mu_3))$. 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 $x^{**} \in (a^*(\mu_2), a(\mu_0))$. Thus, the agent is not more responsive with a lower mean under $\Sigma_{\rho''}$.




(a) Optimal action (b) Induced distribution

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

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.[40]

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., $a^*(\mu_0) = \bar{a}$. 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 $\rho'' \succeq_{spm} \rho'$.

# 8 Appendix B

## 8.1 Games

### Proof of Theorem 2

*Proof.*

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

To simplify exposition, let $n = 2$. Once again, we assume without loss of generality that for each player $i = 1, 2$, the marginal on signals, $F_{S_i}$, is the uniform distribution on the unit interval.

Fix a basic game $G$. For each player $i$, let $\alpha_i : S_i \to A_i$ be an arbitrary measurable and monotone strategy. Let $\mathcal{A}_i$ be the set of all such monotone and measurable strategies and let $\mathcal{A} \triangleq \mathcal{A}_1 \times \mathcal{A}_2$. Given an information structure $\Sigma_\rho$ and opponent's strategies $\alpha_{-i} \in \mathcal{A}_{-i}$, let $a_i^{BR}(\cdot; \alpha_{-i}, \rho) : S_i \to A_i$ be player $i$'s best response strategy. Specifically, for all $s_i \in [0, 1]$,

$$a_i^{BR}(s_i; \alpha_{-i}, \rho) = \arg\max_{a_i \in A_i} \int_{\Theta \times \mathcal{S}_{-i}} u^i\Big(\theta, \alpha_{-i}(s_{-i}), a_i\Big) d\boldsymbol{F}(\theta, s_{-i} | s_i; \rho).$$

By (A.6) and (A.10)-(A.13), $a_i^{BR}(\cdot; \alpha_{-i}, \rho) \in \mathcal{A}_i$ for $i = 1, 2$.[41]

For any given arbitrary monotone strategies $\alpha \triangleq (\alpha_1, \alpha_2) \in \mathcal{A}$, denote the profile of best-response strategies by $a^{BR}(\alpha, \rho) \triangleq \Big(a_1^{BR}(\cdot; \alpha_2, \rho), a_2^{BR}(\cdot; \alpha_1, \rho)\Big)$. Then, a BNE of $\mathcal{G}_\rho$, $a^\star(\rho)$, is given by the fixed point $a^{BR}(a^\star(\rho), \rho) = a^\star(\rho)$.

We only prove the case for $u^i \in \Gamma^\uparrow$. A symmetric argument establishes the result for the case of $u^i \in \Gamma^\downarrow$. The proof to Theorem 2 proceeds in four steps:

1. Player $i$'s best response strategy increases in the increasing convex order when player $i$'s information quality increases (Lemma 4)

2. Player $i$'s best response strategy increases in increasing convex order when player $-i$'s information quality increases (Lemma 5)

3. Player $i$'s best response strategy increases in increasing convex order when player $-i$'s strategy increases in increasing convex order (Lemma 6)

4. Given 1-3, apply comparative statics on fixed points to get desired result.

[41] By the monotonicity of the best response, $a_i^{BR}$ is equivalent to the quantile function almost everywhere. We can thus directly use $a_i^{BR}$ to characterize responsiveness by applying Lemma 1.

**Lemma 4** *Fix some arbitrary strategy $\alpha_{-i} \in \mathcal{A}_{-i}$. Take two structures $\Sigma_{\rho''} \triangleq (\Sigma_{\rho_i''}, \Sigma_{\rho_{-i}})$ and $\Sigma_{\rho'} \triangleq (\Sigma_{\rho_i'}, \Sigma_{\rho_{-i}})$ with $\rho_i'' \succeq_{spm} \rho_i'$. If $u^i \in \Gamma^{\uparrow}$, then $a_i^{BR}(\cdot; \alpha_{-i}, \rho'')$ dominates $a_i^{BR}(\cdot; \alpha_{-i}, \rho_i')$ in the increasing convex order.*

*Proof.* Given $\Sigma_{\rho_{-i}}$ and $\alpha_{-i} \in \mathcal{A}_{-i}$, let

$$\tilde{u}^i(\theta_i, a_i) = \int_{\Theta_{-i} \times S_{-i}} u^i\Big(\theta, \alpha_{-i}(s_{-i}), a_i\Big) d\boldsymbol{F}(\theta_{-i}, s_{-i} | \theta_i; \rho_{-i})$$

so that

$$a_i^{BR}(s_i; \alpha_{-i}, \rho) = \arg\max_{a_i \in A_i} \int_{\Theta_i} \tilde{u}^i(\theta_i, a_i) \mu(d\theta_i | s_i; \rho_i).$$

We have mapped this problem to the single-agent framework where the payoff is given by $\tilde{u}^i : \Theta_i \times A_i \to \mathbb{R}$. Thus, if $\tilde{u}^i \in \mathcal{U}^{\uparrow}$, then by Theorem 1, $a_i^{BR}(\cdot; \alpha_{-i}, \rho'')$ dominates $a_i^{BR}(\cdot; \alpha_{-i}, \rho')$ in the increasing convex order.

First, $\tilde{u}^i$ inherits the measurability, boundedness, and smoothness properties of $u^i$. Furthermore, concavity of $u^i$ in $a_i$ for all $(\theta, a_{-i}) \in \Theta \times A_{-i}$ implies concavity of $\tilde{u}^i$ in $a_i$ for all $\theta_i \in \Theta_i$. Similarly, convexity of $u^i_{a_i}$ in $a_i$ for all $(\theta, a_{-i}) \in \Theta \times A_{-i}$ implies convexity of $\tilde{u}^i_{a_i}$ in $a_i$ for all $\theta_i \in \Theta_i$.

To see that $\tilde{u}^i(\theta_i, a_i)$ has ID in $(\theta_i; a_i)$, let $\theta_i'' > \theta_i'$. Then,

$$\begin{aligned}
&\tilde{u}^i_{a_i}(\theta_i'', a_i) - \tilde{u}^i_{a_i}(\theta_i', a_i) \\
&= \int_{\Theta_{-i} \times S_{-i}} u^i_{a_i}\Big(\theta_i'', \theta_{-i}, \alpha_{-i}(s_{-i}), a_i\Big) d\boldsymbol{F}(\theta_{-i}, s_{-i} | \theta_i''; \rho_{-i}) \\
&- \int_{\Theta_{-i} \times S_{-i}} u^i_{a_i}\Big(\theta_i', \theta_{-i}, \alpha_{-i}(s_{-i}), a_i\Big) d\boldsymbol{F}(\theta_{-i}, s_{-i} | \theta_i'; \rho_{-i}) \\
&= \int_{\Theta_{-i} \times S_{-i}} \left( u^i_{a_i}\Big(\theta_i'', \theta_{-i}, \alpha_{-i}(s_{-i}), a_i\Big) - u^i_{a_i}\Big(\theta_i', \theta_{-i}, \alpha_{-i}(s_{-i}), a_i\Big) \right) d\boldsymbol{F}(\theta_{-i}, s_{-i} | \theta_i''; \rho_{-i}) \\
&+ \int_{\Theta_{-i} \times S_{-i}} u^i_{a_i}\Big(\theta_i', \theta_{-i}, \alpha_{-i}(s_{-i}), a_i\Big) \Big( d\boldsymbol{F}(\theta_{-i}, s_{-i} | \theta_i''; \rho_{-i}) - d\boldsymbol{F}(\theta_{-i}, s_{-i} | \theta_i'; \rho_{-i}) \Big)
\end{aligned}$$

Since $u^i(\theta_i, \theta_{-i}, a_{-i}, a_i)$ has increasing differences (ID) in $(\theta_i; a_i)$ for each $(\theta_{-i}, a_{-i}) \in$

$\Theta_{-i} \times A_{-i}$ and since ID is preserved under integration, the first term

$$\int_{\Theta_{-i}\times S_{-i}} \left( u^i_{a_i}\left(\theta_i'', \theta_{-i}, \alpha_{-i}(s_{-i}), a_i\right) - u^i_{a_i}\left(\theta_i', \theta_{-i}, \alpha_{-i}(s_{-i}), a_i\right)\right) d\boldsymbol{F}(\theta_{-i}, s_{-i}|\theta_i''; \rho_{-i}) \geq 0.$$

Furthermore, since $u^i(\theta_i, \theta_{-i}, a_{-i}, a_i)$ has ID in $(\theta_{-i}, a_{-i}; a_i)$ for each $\theta_i \in \Theta_i$, $u^i_{a_i}(\theta_i, \theta_{-i}, a_{-i}, a_i)$ is increasing in $(\theta_{-i}, a_{-i})$. As $\alpha_{-i}$ is a monotone strategy, by (A.13) and (A.6), the second term

$$\int_{\Theta_{-i}\times S_{-i}} u^i_{a_i}\left(\theta_i', \theta_{-i}, \alpha_{-i}(s_{-i}), a_i\right)\left( d\boldsymbol{F}(\theta_{-i}, s_{-i}|\theta_i''; \rho_{-i}) - d\boldsymbol{F}(\theta_{-i}, s_{-i}|\theta_i'; \rho_{-i})\right) \geq 0.$$

Hence, $\tilde{u}^i(\theta_i, a_i)$ has ID in $(\theta_i; a_i)$.

A similar argument establishes that $\tilde{u}^i_{a_i}(\theta_i, a_i)$ has ID in $(\theta_i; a_i)$. Thus, $\tilde{u}^i \in \mathcal{U}^{\uparrow}$. The desired result in the statement of the lemma follows from Theorem 1. ■

**Lemma 5** *Fix some arbitrary strategy $\alpha_{-i} \in \mathcal{A}_{-i}$. Take two structures $\Sigma_{\rho''} \triangleq (\Sigma_{\rho_i}, \Sigma_{\rho''_{-i}})$ and $\Sigma_{\rho'} \triangleq (\Sigma_{\rho_i}, \Sigma_{\rho'_{-i}})$ with $\rho''_{-i} \succeq_{spm} \rho'_{-i}$. If $u^i \in \Gamma^{\uparrow}$, then $a_i^{BR}(\cdot; \alpha_{-i}, \rho'')$ dominates $a_i^{BR}(\cdot; \alpha_{-i}, \rho')$ in the increasing convex order.*

*Proof.* Following the same first order condition argument we used in the proof of Theorem 1, for each $s_i \in [0, 1]$,

$$\left(a_i^{BR}(s_i; \alpha_{-i}, \rho') - a_i^{BR}(s_i; \alpha_{-i}, \rho'')\right) \underbrace{\int_{\Theta\times S_{-i}} -u^i_{a_i a_i}(\theta, \alpha_{-i}(s_{-i}), a_i^{BR}(s_i; \alpha_{-i}, \rho'))\, d\boldsymbol{F}(\theta, s_{-i}|s_i; \rho'')}_{\triangleq \hat{B}(s_i)^{-1}}$$

$$+ \int_{\Theta\times S_{-i}} u^i_{a_i}(\theta, \alpha_{-i}(s_{-i}), a_i^{BR}(s_i; \alpha_{-i}, \rho'))\left( d\boldsymbol{F}(\theta, s_{-i}|s_i; \rho'') - d\boldsymbol{F}(\theta, s_{-i}|s_i; \rho')\right) \leq 0.$$

Then, for each $t \in [0,1]$,

$$
\begin{aligned}
&\int_t^1 \Big(a_i^{BR}(s_i;\alpha_{-i},\rho') - a_i^{BR}(s_i;\alpha_{-i},\rho'')\Big)ds_i \\
&\leq \int_t^1 \hat{B}(s_i) \int_{\Theta\times S_{-i}} u^i_{a_i}(\theta,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\alpha_{-i},\rho'))\Big(d\boldsymbol{F}(\theta,s_{-i}|s_i;\rho') - d\boldsymbol{F}(\theta,s_{-i}|s_i;\rho'')\Big)ds_i \\
&= \int_{\Theta\times S} u^i_{a_i}(\theta_i,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\alpha_{-i},\rho'))\hat{B}(s_i)\mathbb{1}_{[s_i\geq t]}\Big(d\boldsymbol{F}(\theta,s;\rho') - d\boldsymbol{F}(\theta,s;\rho'')\Big) \\
&= \int_{\Theta_{-i}\times S_{-i}} \hat{\psi}(\theta_{-i},s_{-i};t)\Big(dF(\theta_{-i},s_{-i};\rho'_{-i}) - dF(\theta_{-i},s_{-i};\rho''_{-i})\Big),
\end{aligned}
$$

where

$$
\hat{\psi}(\theta_{-i},s_{-i};t) = \int_{\Theta_i\times S_i} u^i_{a_i}(\theta,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\alpha_{-i},\rho'))\hat{B}(s_i)\mathbb{1}_{[s_i\geq t]}dF(s_i|\theta_i;\rho_i)\mu(d\theta_i|\theta_{-i}).
$$

The equality follows from (A.11). Take $s''_{-i} > s'_{-i}$ which implies that $\alpha_{-i}(s''_{-i}) \geq \alpha_{-i}(s'_{-i})$. Then,

$$
\Big[u^i_{a_i}(\theta,\alpha_{-i}(s''_{-i}),a_i^{BR}(s_i;\alpha_{-i},\rho')) - u^i_{a_i}(\theta,\alpha_{-i}(s'_{-i}),a_i^{BR}(s_i;\alpha_{-i},\rho'))\Big]\mathbb{1}_{[s_i\geq t]} \geq 0,
$$

for all $(\theta,s_i) \in \Theta\times S_i$ because $u^i(\theta,a_{-i},a_i)$ has increasing differences in $(a_{-i};a_i)$. It is also increasing in both $\theta$ and $s_i$ because $u^i_{a_i}(\theta,a_{-i},a_i)$ has increasing differences in $(\theta,a_i;a_{-i})$. Similarly,

$$
\hat{B}(s_i) \geq 0
$$

by concavity of $u^i$ in $a_i$. It is increasing in $s_i$ because $u^i_{a_i}$ is convex in $a_i$, has ID in $(\theta,a_{-i};a_i)$, and because $\boldsymbol{F}(\theta,s_{-i}|s_i;\rho'')$ is increasing in FOSD as $s_i$ increases. Thus, along with (A.6)

and (A.13),

$$\begin{aligned}
&\hat{\psi}(\theta_{-i}, s''_{-i}; t) - \hat{\psi}(\theta_{-i}, s'_{-i}; t) \\
&= \int_{\Theta_i \times S_i} \Big\{ \Big[ u^i_{a_i}(\theta, \alpha_{-i}(s''_{-i}), a^{BR}_i(s_i; \alpha_{-i}, \rho')) - u^i_{a_i}(\theta, \alpha_{-i}(s'_{-i}), a^{BR}_i(s_i; \alpha_{-i}, \rho')) \Big] \\
&\qquad \times \hat{B}(s_i) \mathbb{1}_{[s_i \geq t]} \Big\} dF(s_i|\theta_i; \rho_i) \mu(d\theta_i|\theta_{-i})
\end{aligned}$$

is increasing in $\theta_{-i}$. In other words, $\hat{\psi}(\theta_{-i}, s_{-i}; t)$ has ID in $(\theta_{-i}; s_{-i})$. By Lemma 2, $\rho''_{-i} \succeq_{spm} \rho'_{-i}$ implies

$$\int_{\Theta_{-i} \times S_{-i}} \hat{\psi}(\theta_{-i}, s_{-i}; t) \Big( dF(\theta_{-i}, s_{-i}; \rho'_{-i}) - dF(\theta_{-i}, s_{-i}; \rho''_{-i}) \Big) \leq 0,$$

giving us the desired result. ■

**Lemma 6** *Fix $\Sigma_\rho$. Let $\alpha''_{-i}, \alpha'_{-i} \in \mathcal{A}_{-i}$ such that $\alpha''_{-i}$ dominates $\alpha'_{-i}$ in the increasing convex order. If $u^i \in \Gamma^\uparrow$, then, $a^{BR}_i(\cdot; \alpha''_{-i}, \rho)$ also dominates $a^{BR}_i(\cdot; \alpha'_{-i}, \rho)$ in the increasing convex order.*

*Proof.* Suppress the dependence on $\rho$ as it is held fixed. For any $t \in [0, 1]$, we use the first order conditions argument (similar to the proof of Lemma 5) to get the expression

$$\begin{aligned}
&\int_t^1 (a^{BR}_i(s_i; \alpha'_{-i}) - a^{BR}_i(s_i; \alpha''_{-i})) ds_i \\
&\leq \int_t^1 \Bigg\{ \underbrace{\Bigg( - \int_{\Theta \times S_{-i}} u^i_{a_i a_i}\Big(\theta, \alpha''_{-i}(s_{-i}), a^{BR}_i(s_i; \alpha'_{-i})\Big) d\boldsymbol{F}(\theta, s_{-i}|s_i) \Bigg)^{-1}}_{\triangleq \tilde{B}_i(s_i)} \\
&\times \int_{\Theta \times S_{-i}} \Big[ u^i_{a_i}\Big(\theta, \alpha'_{-i}(s_{-i}), a^{BR}_i(s_i; \alpha'_{-i})\Big) - u^i_{a_i}\Big(\theta, \alpha''_{-i}(s_{-i}), a^{BR}_i(s_i; \alpha'_{-i})\Big) \Big] d\boldsymbol{F}(\theta, s_{-i}|s_i) \Bigg\} ds_i.
\end{aligned}$$

By convexity of $u^i_{a_i}$ in $a_{-i}$,

$$u^i_{a_i}\left(\theta_i, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\right) - u^i_{a_i}\left(\theta_i, \alpha''_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\right)$$

$$\leq u^i_{a_i a_{-i}}\left(\theta_i, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\right)(\alpha'_{-i}(s_{-i}) - \alpha''_{-i}(s_{-i})).$$

Thus,

$$\int_t^1 (a_i^{BR}(s_i;\alpha'_{-i}) - a_i^{BR}(s_i;\alpha''_{-i}))ds_i$$
$$\leq \int_{S_{-i}} (\alpha'_{-i}(s_{-i}) - \alpha''_{-i}(s_{-i})) \int_{\Theta\times S_i} u^i_{a_i a_{-i}}\left(\theta, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\right)\tilde{B}(s_i)\mathbb{1}_{[s_i\geq t]}d\boldsymbol{F}(\theta, s_i|s_{-i})ds_{-i}.$$

From Lemma 1 and the equivalence of the monotone strategy $\alpha_{-i}$ with its quantile function, $\alpha''_{-i}$ dominates $\alpha'_{-i}$ in the increasing convex order if

$$\int_t^1 (\alpha'_{-i}(s_{-i}) - \alpha''_{-i}(s_{-i}))ds_{-i} \leq 0, \quad \forall t \in [0,1].$$

Furthermore, for each $s_{-i} \in [0,1]$

$$u^i_{a_i a_{-i}}\left(\theta, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\right)\mathbb{1}_{[s_i\geq t]} \geq 0, \ \forall(\theta, s_i) \in \Theta\times S_i$$

as $u^i$ has increasing differences in $(a_{-i}; a_i)$ for all $\theta \in \Theta$. It is also increasing in both $\theta$ and $s_i$ because $u^i_{a_i}$ has increasing differences in $(\theta, a_i; a_{-i})$. Similarly,

$$\tilde{B}(s_i) \geq 0$$

by concavity of $u^i$ in $a_i$. It is increasing in $s_i$ because $u^i_{a_i}$ is convex in $a_i$, has ID in $(\theta, a_{-i}; a_i)$, and because $\boldsymbol{F}(\theta, s_{-i}|s_i)$ is increasing in FOSD as $s_i$ increases. Thus, along with (A.6) and (A.13),

$$\int_{\Theta\times S_i} u^i_{a_i a_{-i}}\left(\theta, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\right)\tilde{B}(s_i)\mathbb{1}_{[s_i\geq t]}d\boldsymbol{F}(\theta, s_i|s_{-i})$$

is an increasing function of $s_{-i}$. Applying Lemma 3, we have

$$
\begin{aligned}
&\int_t^1 (a_i^{BR}(s_i;\alpha'_{-i}) - a_i^{BR}(s_i;\alpha''_{-i}))ds_i \\
&\leq \int_{S_{-i}} (\alpha'_{-i}(s_{-i}) - \alpha''_{-i}(s_{-i})) \int_{\Theta\times S_i} u^i_{a_i a_{-i}}\left(\theta, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\right)\tilde{B}(s_i)\mathbb{1}_{[s_i\geq t]}d\boldsymbol{F}(\theta, s_i|s_{-i})ds_{-i} \\
&\leq \underbrace{\int_{S_{-i}} (\alpha'_{-i}(s_{-i}) - \alpha''_{-i}(s_{-i}))ds_{-i}}_{\leq 0} \int_{\Theta\times S_i} \underbrace{u^i_{a_i a_{-i}}\left(\theta, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\right)\tilde{B}(s_i)\mathbb{1}_{[s_i\geq t]}}_{\geq 0} d\boldsymbol{F}(\theta, s_i|0) \\
&\leq 0
\end{aligned}
$$

for each $t \in [0,1]$. ■

We now tackle the last step in the proof: comparative statics of the BNEs. We apply the comparative statics of fixed points provided by Villas-Boas (1997). To do so, we will need the following definition.

**Definition 6 (Contractible Space)** *Let $X$ be a topological space. We say that $X$ is a contractible space if there exists a map $\Phi : X \times [0,1] \to X$ such that for all $x \in X$*

1. *$\Phi(\cdot, \lambda)$ is continuous in $\lambda$,*
2. *$\Phi(x, 0) = x$ and $\Phi(x, 1) = x^*$ for some $x^* \in X$*

Intuitively, $X$ is contractible if it can be continuously shrunk into a point inside itself.

**Villas-Boas (1997, Theorem 6)** *Let $X$ be a compact subset of a Banach space. Consider continuous mappings $T_1 : X \to X$ and $T_2 : X \to X$, and a transitive and reflexive order $\succeq$ on $X$. For all $x \in X$, let the upper-set $\{x' \in X : x' \succeq x\}$ be a compact and contractible subset. Let both $T_1$ and $T_2$ have a fixed point on $X$. Suppose $x' \succeq x \Rightarrow T_1(x') \succeq T_1(x)$, and suppose $T_1(x) \succeq T_2(x)$ for all $x \in X$. Then for every fixed point $x_2^\star$ of $T_2$, there is a fixed point $x_1^\star$ of $T_1$ such that $x_1^\star \succeq x_2^\star$.*

The remaining few steps prove that our setting satisfies the assumptions needed to apply the Villas-Boas result.[42]

Let $BV([0,1],\mathbb{R})$ be the space of functions of bounded variation from $[0,1]$ to $\mathbb{R}$. Given a function $g \in BV([0,1],\mathbb{R})$, let $V(g)$ be the total variation of $g$.[43] Define the bounded variation norm by $||g||_{BV} = \int_0^1 |g(s)|ds + V(g)$. The space $BV([0,1],\mathbb{R})$ equipped with the $||\cdot||_{BV}$ is a Banach space.

**Lemma 7** *For each $i = 1, 2$, $\mathcal{A}_i$ is a compact subset of the Banach space $(BV([0,1],\mathbb{R}), ||\cdot||_{BV})$.*

*Proof.* Any $\alpha_i \in \mathcal{A}_i$ is of bounded variation as it is an increasing function. Therefore, $\mathcal{A}_i$ is a subset of $BV([0,1],\mathbb{R})$.

To show that $\mathcal{A}_i$ is a compact subset $BV([0,1],\mathbb{R})$, take a sequence $\{\tilde{\alpha}_{i,k}\}_{k=1}^{\infty} \in \mathcal{A}_i$. The sequence is uniformly bounded as each function $\alpha_{i,k}$ maps into the compact interval $A_i$. By Helly's Selection Theorem, the sequence converges to an increasing function $\tilde{\alpha}_i \in BV([0,1],\mathbb{R})$.

Furthermore, as $\underline{a}_i \leq \tilde{\alpha}_{i,k}(0)$ for all $k$, the limit also satisfies $\underline{a}_i \leq \tilde{\alpha}_i(0)$. Similarly, as $\bar{a}_i \geq \tilde{\alpha}_{i,k}(1)$ for all $k$, the limit also satisfies $\bar{a}_i \geq \tilde{\alpha}_i(1)$. Finally, the point-wise limit of measurable functions is measurable (Corollary 8.9, Measure, Integrals, and Martingales, Schilling (2005)). As $\tilde{\alpha}_i$ is a monotone and measurable function that maps from $[0,1]$ to $A_i$, $\tilde{a}_i \in \mathcal{A}_i$. ■

Define a weak partial order over $\mathcal{A}_i$ by $\alpha_i'' \succeq_i \alpha_i'$ if, and only if, $\alpha_i''$ dominates $\alpha_i'$ in the increasing convex order. Using Lemma 1 (and the equivalence of $\alpha_i$ to its quantile function),

[42] For the case when $u^i \in \Gamma^{\downarrow}$ for all $i \in N$, we use Theorem 7 of Villas-Boas (1997) which uses the lower-sets generated by the decreasing convex order to get the desired comparative statics of fixed points.

[43] Specifically, $V(g) = \sup_{p \in P} \sum_{i=0}^{n_p - 1} |g(x_{i+1}) - g(x_i)|$ where $P$ is the set of all partitions $p = \{x_0, x_1, \ldots, x_{n_p}\}$ on $[0,1]$.

$\alpha_i'' \succeq_i \alpha_i'$ if, and only if,

$$\int_t^1 \alpha_i''(s_i)ds_i \geq \int_t^1 \alpha_i'(s_i)ds_i, \quad \forall t \in [0,1].$$

Denote the upper-set of $\alpha_i$ by $\mathcal{US}(\alpha_i) \triangleq \{\alpha_i' \in \mathcal{A}_i : \alpha_i' \succeq_i \alpha_i\} \subseteq \mathcal{A}_i$.

**Lemma 8** *For each $i = 1, 2$, and for any $\alpha_i \in \mathcal{A}_i$, $\mathcal{US}(\alpha_i)$ is a compact and contractible set.*

*Proof.* For a given $\alpha_i \in \mathcal{A}_i$, $\mathcal{US}(\alpha_i)$ is a closed subset of $\mathcal{A}_i$ (follows from the dominated convergence Theorem). Hence, it is compact. To show that $\mathcal{US}(\alpha_i)$ is contractible, let $\alpha_i^c : [0,1] \to A_i$ be the constant function with $\alpha_i^c(s) = \bar{a}_i$ for all $s \in [0,1]$. Note that $\alpha_i^c \in \mathcal{A}_i$. Furthermore, $\alpha_i^c(s) \geq \alpha_i(s)$, $\forall s \in [0,1]$ which implies $\alpha_i^c \succeq_i \alpha_i \Rightarrow \alpha_i^c \in \mathcal{US}(\alpha_i)$.

For each $\alpha_i \in \mathcal{A}_i$, defne the mapping $\Phi : \mathcal{US}(\alpha_i) \times [0,1] \to \mathcal{US}(\alpha_i)$ such that

$$\Phi(\alpha_i', \lambda) = (1-\lambda)\alpha_i' + \lambda\alpha_i^c.$$

$\Phi(\cdot, \lambda)$ is continuous in $\lambda$. As $\lambda$ increases from 0 to 1, $\Phi$ continuously deforms any strategy in $\mathcal{US}(\alpha_i)$ to the constant strategy $\alpha_i^c$, which is itself in $\mathcal{US}(\alpha_i)$. Therefore, $\mathcal{US}(\alpha_i)$ is contractible. ■

Thus far, we have an order, $\succeq_i$, on $\mathcal{A}_i$ that generates compact and contractible upper-sets. We extend these properties to $\mathcal{A} \triangleq \mathcal{A}_1 \times \mathcal{A}_2$ by the product order: given $\alpha'', \alpha' \in \mathcal{A}$, $\alpha'' \succeq \alpha'$ if, and only if, $\alpha_i'' \succeq_i \alpha_i'$ for each $i = 1, 2$. Along with the product topology, $\succeq$ is a partial order on $\mathcal{A}$ that generates compact and contractible upper-sets.[44]

For a Bayesian game $\mathcal{G}_\rho \triangleq (\Sigma_{\rho_1}, \Sigma_{\rho_2}, G)$, define an operator $T_\rho : \mathcal{A} \to \mathcal{A}$ with

$$T_\rho(\alpha) = \left(a_1^{BR}(\alpha_2, \rho), a_2^{BR}(\alpha_1, \rho)\right).$$

$T_\rho$ is continuous in $\alpha$ as utility functions are continuous in actions. A monotone BNE of $\mathcal{G}_\rho$,

[44] $\mathcal{A}$ is a subset of a Banach space equipped with the metric $d(\alpha', \alpha) = \sum_i ||\alpha_i' - \alpha_i||_{BV}$.

$a^\star(\rho)$, is a fixed point of $T_\rho$. We know such a fixed point exists (Van Zandt and Vives, 2007).

Consider two different games, $\mathcal{G}_{\rho''} \triangleq (\Sigma_{\rho''_1}, \Sigma_{\rho''_2}, G)$ and $\mathcal{G}_{\rho'} \triangleq (\Sigma_{\rho'_1}, \Sigma_{\rho'_2}, G)$, with $\rho''_i \succeq_{spm} \rho'_i$ for all $i = 1, 2$. For all $\alpha \in \mathcal{A}$,

$$\rho''_i \succeq_{spm} \rho'_i, \forall i \underbrace{\Rightarrow}_{\substack{\text{by Lemma 4} \\ \text{and Lemma 5}}} a_i^{BR}(\alpha_{-i}, \rho'') \succeq_i a_i^{BR}(\alpha_{-i}, \rho'), \forall i \Leftrightarrow T_{\rho''}(\alpha) \geq T_{\rho'}(\alpha).$$

Furthermore,

$$\alpha'' \geq \alpha' \Leftrightarrow \alpha''_i \succeq_i \alpha'_i, \forall i \underbrace{\Rightarrow}_{\text{by Lemma 6}} a_i^{BR}(\alpha''_{-i}, \rho'') \succeq_i a_i^{BR}(\alpha'_{-i}, \rho''), \forall i \Leftrightarrow T_{\rho''}(\alpha'') \geq T_{\rho''}(\alpha').$$

We can now directly apply Theorem 6 of Villas-Boas (1997) to conclude that, for every fixed point $a^\star(\rho')$ of $T_{\rho'}$, there is a fixed point $a^\star(\rho'')$ of $T_{\rho''}$ such that $a^\star(\rho'') \geq a^\star(\rho')$. ■

## 8.2 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 $V(\rho'') - V(\rho')$

$$\begin{aligned} &= \int_{\Theta \times S} v(\theta, a(s; \rho''))\,[dF(\theta, s; \rho'') - dF(\theta, s; \rho')] \\ &+ \int_{\Theta \times S} [v(\theta, a(s; \rho'')) - v(\theta, a(s; \rho'))]\,dF(\theta, s; \rho'). \end{aligned} \tag{1}$$

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[45] and convex in $a$ for all $\theta \in \Theta$, the second integral term satisfies

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

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

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

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

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

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 $a$.

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

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 $a$,

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

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

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

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 Theorem 3**

*Proof.*

First case: Assume $u \in \mathcal{U}^I$, and $v(\theta, a)$ is increasing and convex in $a$, and satisfies increasing differences in $(\theta; a)$:

Let $\mu_2 \succeq_{FOSD} \mu_1$ and let $\mu_\lambda = \lambda\mu_1 + (1-\lambda)\mu_2$. By Proposition 1 $a^*(\mu_\lambda) \leq \lambda a^*(\mu_1) + (1-\lambda)a^*(\mu_2)$. Let $a_1 = a^*(\mu_1)$, $a_2 = a^*(\mu_2)$, $a_\lambda = \lambda a_1 + (1-\lambda)a_2$ then

$$\begin{aligned}
\int v(a^*(\mu_\lambda), \theta) d\mu_\lambda &\leq \int v(a^*(\mu_\lambda), \theta) d\mu_\lambda \\
&\leq \lambda \int v(a_1, \theta) d\mu_\lambda + (1-\lambda) \int v(a_2, \theta) d\mu_\lambda \\
&= \lambda^2 \int v(a_1, \theta) d\mu_1 + (1-\lambda)^2 \int v(a_2, \theta) d\mu_2 \\
&\quad \lambda(1-\lambda) \left[ \int v(a_1, \theta) d\mu_2 + \int v(a_2, \theta) d\mu_1 \right] \\
&= \lambda \int v(a_1, \theta) d\mu_1 + (1-\lambda) \int v(a_2, \theta) d\mu_2 \\
&\quad \lambda(1-\lambda) \left[ \int v(a_1, \theta) d\mu_2 + \int v(a_2, \theta) d\mu_1 - \int v(a_1, \theta) d\mu_1 - \int v(a_2, \theta) d\mu_2 \right] \\
&\leq \lambda \int v(a_1, \theta) d\mu_1 + (1-\lambda) \int v(a_2, \theta) d\mu_2
\end{aligned}$$

By induction, if any signal $\Sigma_{\rho''}$ satisfies (A.5) and has finite posteriors, then for any garbling $\Sigma_{\rho'}$ with finite posteriors we have that $V(\rho'') > V(\rho')$ (see the proof of Corollary 1). The above argument can be extended to the case of infinite posteriors following the methods in Zhang (2008). The other two cases are analogous. ■

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

$$= \int_{\Theta \times S_2} u^1(\theta, a_1^\star(\theta; \rho''), a_2^\star(s_2; \rho'')) dF(\theta, s_2; \rho''_2) - \int_{\Theta \times S_2} u^1(\theta, a_1^\star(\theta; \rho'), a_2^\star(s_2; \rho')) dF(\theta, s_2; \rho'_2),$$

which can be written as

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

Thus, $\rho_2'' \geq_{spm} \rho_2'$ 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 4**

*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.[46]

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

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

$U_1(\rho;\hat{\rho})$ is given by

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

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

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

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

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

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

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

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

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

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

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

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

This concludes the proof. ■

## 8.3 Blackwell, Lehmann, and Supermodular Order

It is natural to ask why the supermodular order is the relevant order to consider instead of the more familiar Blackwell informativeness (Blackwell, 1951, 1953) or the Lehmann (accuracy) order (Lehmann, 1988). The answer is two-fold: The first reason for focusing on the supermodular order has to do with the value of information in the class of decision problems we consider. Blackwell (1951, 1953) shows that all decision makers value a higher quality of

information if, and only if, information quality is ranked by Blackwell informativeness. Athey and Levin (2017) show that if the class of decision problems is restricted to supermodular preferences, then a higher quality of information is valuable if, and only if, information quality is ranked by the more general supermodular order. Our results further solidify the link between the class of supermodular payoffs and the supermodular order by providing conditions on the marginal utilities of supermodular payoff functions such that, agents are more responsive when information quality increases if, and only if, information quality is ranked by the supermodular order.

Second, within the class of information structures that satisfy (A.5), the supermodular order is a more general ordering than Blackwell informativeness and the Lehmann ordering. In particular, if information structures satisfy the MLRP property (a stronger assumption than (A.5)), then Blackwell informativeness implies the Lehmann order which in turn implies the supermodular order. The converse however is not true, as shown by the example below. Figure 7 depicts the nesting of information orders and the associated class of decision problems.

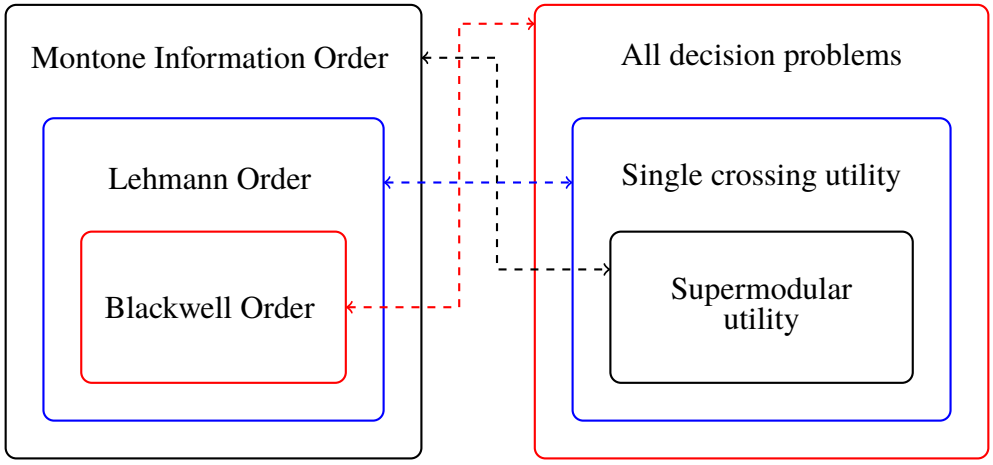



Figure 9: Information ordering and decision problems

The following is an example of information structures that can be ordered using the supermodular order but not the Lehmann order.[47] For this section only, we consider information structures $\Sigma_\rho \triangleq (S, \{F(\cdot|\theta;\rho)\}_{\theta\in\Theta})$ such that $\{F(\cdot|\theta;\rho)\}_{\theta\in\Theta}$ satisfies the MLRP property, i.e.,

[47]See Lehmann (1988), Persico (2000), and Jewitt (2007) for a more complete analysis of the Lehmann ordering.

for any $s < s'$, the likelihood function

$$\frac{f(s'|\theta;\rho)}{f(s|\theta;\rho)}$$

is non-decreasing in $\theta$.[48]

**Lehmann (Accuracy) Order:** $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the Lehmann order, denoted $\rho'' \geq_L \rho'$, if for all $s \in S$,

$$F^{-1}\Big(F(s|\theta;\rho')|\theta;\rho''\Big)$$

is non-decreasing in $\theta$.

**Example:** Let $\theta \in \{\theta_1, \theta_2, \theta_3\}$ with $\theta_1 < \theta_2 < \theta_3$. Let $\mu_i^o$ be the mass at $\theta_i$ with $\mu_1^o = \mu_2^o = \frac{2}{5}$ and $\mu_3^o = \frac{1}{5}$. Consider two information structure $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$ such that the signal space $S$ is the unit interval for both structures and $F(s|\theta_i;\rho')$ is given by

| | $0 \leq s < 1/2$ | $1/2 \leq s \leq 1$ |
|---|---|---|
| $\theta_1$ | $s\frac{3}{2}$ | $\frac{1+s}{2}$ |
| $\theta_2$ | $s$ | $s$ |
| $\theta_3$ | 0 | $2s-1$ |

while $F(s|\theta_i;\rho'')$ is given by For both information structures, the marginal on the signal is

| | $0 \leq s < 1/2$ | $1/2 \leq s \leq 1$ |
|---|---|---|
| $\theta_1$ | $2s$ | 1 |
| $\theta_2$ | $\frac{s}{2}$ | $\frac{3s-1}{2}$ |
| $\theta_3$ | 0 | $2s-1$ |

simply the uniform distribution on $S = [0,1]$, i.e., $F_S(s;\rho') = F_s(s;\rho'') = s$ for all $s \in [0,1]$. Furthermore, both structures satisfy the MLRP property: for any $s < s' < 1/2$ or $1/2 \leq s < s'$,

[48]This is a more restrictive assumption on signal structures than (A.5).

the likelihood functions satisfy

$$\frac{f(s'|\theta_i;\rho')}{f(s|\theta_i;\rho')} = \frac{f(s'|\theta_i;\rho'')}{f(s|\theta_i;\rho'')} = 1 \qquad \forall i = 1,2,3,$$

while for any $s < 1/2 \leq s'$, the likelihood ratios satisfy

$$\underbrace{\frac{f(s'|\theta_1;\rho')}{f(s|\theta_1;\rho')}}_{=1/3} < \underbrace{\frac{f(s'|\theta_2;\rho')}{f(s|\theta_2;\rho')}}_{=1} < \underbrace{\frac{f(s'|\theta_3;\rho')}{f(s|\theta_3;\rho')}}_{=\infty},$$

and

$$\underbrace{\frac{f(s'|\theta_1;\rho'')}{f(s|\theta_1;\rho'')}}_{=0} < \underbrace{\frac{f(s'|\theta_2;\rho'')}{f(s|\theta_2;\rho'')}}_{=3} < \underbrace{\frac{f(s'|\theta_3;\rho'')}{f(s|\theta_3;\rho'')}}_{=\infty}.$$

As a result, $s' > s$ implies $\mu(\cdot|s';\rho) \geq_{FOSD} \mu(\cdot|s;\rho)$, for $\rho = \rho', \rho''$ (Milgrom; 1981).

We first show that $\rho' \not\geq_L \rho''$ and $\rho'' \not\geq_L \rho'$. If $\rho' \geq_L \rho''$, then

$$F^{-1}\Big(F(s|\theta;\rho'')|\theta;\rho'\Big)$$

must be increasing in $\theta$ for every $s \in [0,1]$. However, for all $s \in [0,1]$

$$F^{-1}\Big(F(s|\theta_3;\rho'')|\theta_3;\rho'\Big) = s$$

whereas

$$F^{-1}\Big(F(s|\theta_1;\rho'')|\theta_1;\rho'\Big) \geq F^{-1}\Big(F(s|\theta_1;\rho')|\theta_1;\rho'\Big) = s$$

since $F(s|\theta_1;\rho'') \geq F(s|\theta_1;\rho')$. Similarly,

$$F^{-1}\Big(F(s|\theta_2;\rho'')|\theta_2;\rho'\Big) \leq F^{-1}\Big(F(s|\theta_2;\rho')|\theta_2;\rho'\Big) = s$$

because $F(s|\theta_2;\rho'') \leq F(s|\theta_2;\rho')$. Altogether, we have

$$F^{-1}\Big(F(\cdot|\theta_2;\rho'')|\theta_2;\rho'\Big) < F^{-1}\Big(F(\cdot|\theta_3;\rho'')|\theta_3;\rho'\Big) < F^{-1}\Big(F(\cdot|\theta_1;\rho'')|\theta_1;\rho'\Big)$$

violating the Lehmann monotonicity condition. Thus, $\rho' \not\succeq_L \rho''$.

Figure 8 depicts the conditional distributions of the signals. The solid black line is the conditional distribution of signals given $\theta_3$ under both $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$. The solid and dashed blue lines are the conditional distribution of signals given $\theta_1$ under $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$ respectively. Similarly, solid and dashed red lines are the conditional distribution of signals given $\theta_2$ under $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$ respectively. Starting from $s^* \in [0,1]$, the arrows show the transformation to $\tau_i = F^{-1}\Big(F(s^*|\theta_i;\rho'')|\theta_i;\rho'\Big)$ where the blue, red, and black arrows correspond to $\theta_1$, $\theta_2$, and $\theta_3$ respectively. Similarly, If $\rho'' \succeq_L \rho'$, then

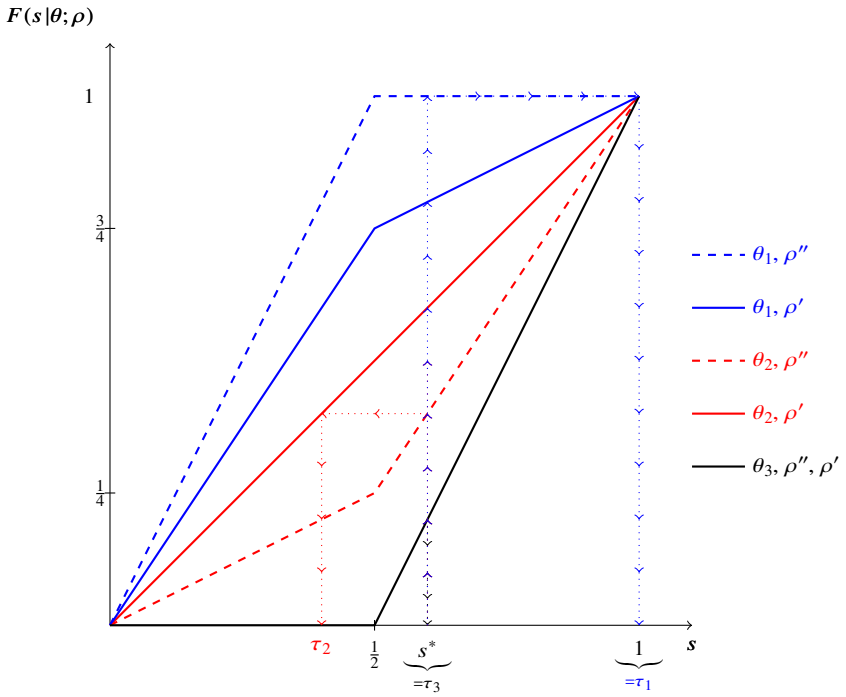



Figure 10: $\rho' \not\succeq_L \rho''$

$$F^{-1}\Big(F(s|\theta;\rho')|\theta;\rho''\Big)$$

must be increasing in $\theta$ for every $s \in [0,1]$. However, for all $s \in [0,1]$,

$$F^{-1}\Big(F(s|\theta_3;\rho')|\theta_3;\rho''\Big) = s$$

whereas

$$F^{-1}\Big(F(s|\theta_1;\rho')|\theta_1;\rho''\Big) \leq F^{-1}\Big(F(s|\theta_1;\rho'')|\theta_1;\rho''\Big) = s,$$

and

$$F^{-1}\Big(F(s|\theta_2;\rho')|\theta_2;\rho''\Big) \geq F^{-1}\Big(F(s|\theta_2;\rho'')|\theta_2;\rho''\Big) = s.$$

Altogether, we have

$$F^{-1}\Big(F(\cdot|\theta_1;\rho')|\theta_1;\rho''\Big) < F^{-1}\Big(F(\cdot|\theta_3;\rho')|\theta_3;\rho''\Big) < F^{-1}\Big(F(\cdot|\theta_2;\rho')|\theta_2;\rho''\Big)$$

violating the Lehmann monotonicity condition. Thus, $\rho'' \not\geq_L \rho'$. Furthermore, $\Sigma_{\rho''}$ and $\Sigma_{\rho'}$ are also not Blackwell ordered since Blackwell ordering implies Lehmann ordering (within the class of information structures with MLRP property).

Figure 9 depicts the conditional distributions of the signals. The solid black line is the conditional distribution of signals given $\theta_3$ under both $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$. The solid and dashed blue lines are the conditional distribution of signals given $\theta_1$ under $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$ respectively. Similarly, solid and dashed red lines are the conditional distribution of signals given $\theta_2$ under $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$ respectively. Starting from $s^* \in [0,1]$, the arrows show the transformation to $\tilde{\tau}_i = F^{-1}\Big(F(s^*|\theta_i;\rho')|\theta_i;\rho''\Big)$ where the blue, red, and black arrows correspond to $\theta_1$, $\theta_2$, and $\theta_3$ respectively.

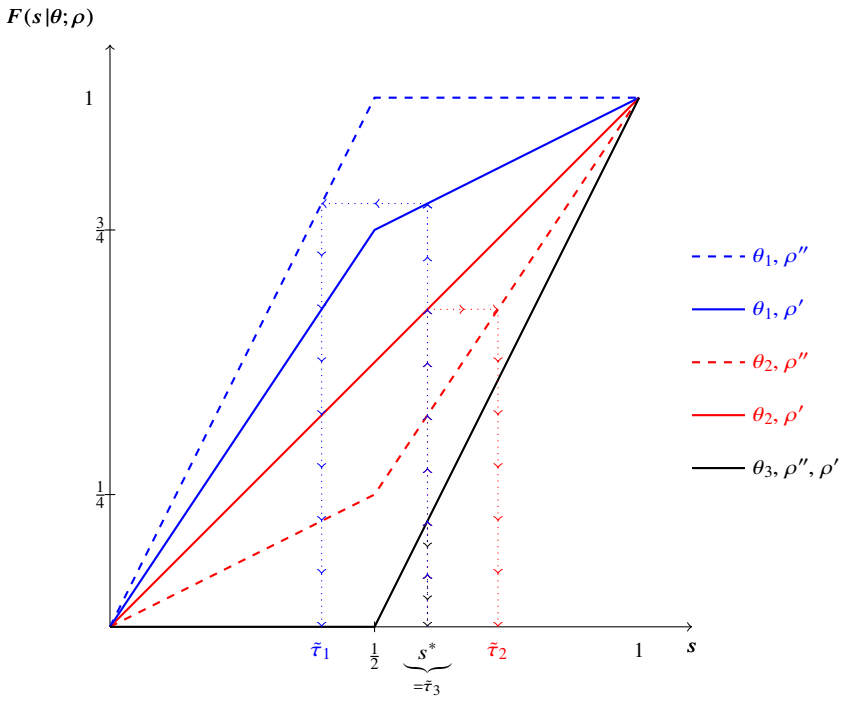



Figure 11: $\rho'' \not\succeq_L \rho'$

Next, we show that $\rho'' \succeq_{spm} \rho'$. From Lemma 2, $\rho'' \succeq_{spm} \rho'$ if $F(\theta_i, s; \rho') - F(\theta_i, s; \rho'') \leq 0$ for all $(\theta_i, s)$. Notice that for all $s \in [0, 1]$,

$$F(\theta_1, s; \rho') - F(\theta_1, s; \rho'') = \mu_1^o\Big(F(s|\theta_1; \rho') - F(s|\theta_1; \rho'')\Big) \leq 0.$$

Furthermore, for all $s \in [0, 1]$,

$$\begin{aligned} F(\theta_2, s; \rho') - F(\theta_2, s; \rho'') =& \mu_1^o\Big(F(s|\theta_1; \rho') - F(s|\theta_1; \rho'')\Big) + \mu_2^o\Big(F(s|\theta_2; \rho') - F(s|\theta_2; \rho'')\Big) \\ =& \frac{2}{5}\Big(F(s|\theta_1; \rho') - F(s|\theta_1; \rho'') + F(s|\theta_2; \rho') - F(s|\theta_2; \rho'')\Big) = 0. \end{aligned}$$

Finally, $F(\theta_3, s; \rho') - F(\theta_3, s; \rho'') = \sum_{i=1}^3 \mu_i^o\Big(F(s|\theta_i; \rho') - F(s|\theta_i; \rho'')\Big) = F_S(s; \rho') - F_S(s; \rho'') = 0$. Hence, $\rho'' \succeq_{spm} \rho'$.