Ex-ante Comparative Statics: Responsiveness to Information Quality*

Teddy Mekonnen † René Leal Vizcaíno ‡

November 19, 2017

Abstract

An agent chooses an action after acquiring information about an uncertain state. From an ex-ante perspective, the agent's optimal action is an endogenously determined random variable. We study how the quality of information affects the distribution of the optimal action. In particular, we study *responsiveness*, a comparative statics that captures mean-preserving spreads and second-order stochastic dominance shifts in the distribution of the optimal action. The higher the quality of information, the more closely the agent tailors her actions to the state, and consequently, under conditions we derive on payoffs, the more *responsive* the optimal action. We extend our results to Bayesian games with strategic complementarities in which different players have information of varying quality. We show that a player's equilibrium actions become more responsive as the quality of an opponent's information improves. We apply the comparative statics of responsiveness to compare the demand for information in covert and overt information acquisition games.

Keywords: Supermodularity, comparative statics, stochastic dominance, informativeness, decision-making under uncertainty, games of incomplete information

JEL Codes: C44, C61, D42, D81

*We are indebted to Asher Wolinsky, Eddie Dekel, Bruno Strulovici, and Jeff Ely for their unwavering guidance and encouragement. Our work has also greatly benefited from conversations with Laura Doval, Alessandro Pavan, Rob Porter, Phil Reny, Marciano Siniscalchi, Jean Tirole, and Zenon Zabinski.

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

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

1 Introduction

An agent faces a decision-making problem under uncertainty with learning: She first observes a signal from an information structure (also referred to as an experiment) that is informative about an unobserved state of the world. She then chooses an action. Since her action choice depends on the signals she receives, from an ex-ante perspective, the agent's optimal action is an endogenously determined random variable.

In this paper, we study how the quality of the agent's information structure affects the induced distribution of optimal action. Specifically, we consider a setting where the agent's action and the state are complements. We study how the mean and dispersion of the optimal action change when the quality of the information structure increases. We first define two orders over the distribution of optimal actions we collectively call *responsiveness* that captures different notions of dispersion such as mean-preserving spreads and second-order stochastic dominance relations. We then identify conditions on payoffs and experiments such that *more informative* experiments lead to *more responsive* optimal actions.

To concretely motivate our question, consider a monopolist facing a linear demand curve $P(q) = 1 - q$ and a quadratic cost function $c(\theta, q) = (1 - \theta)q + q^2/2$, where q is the number of units produced and $\theta \in [0, 1]$ is a cost parameter. Higher values of θ correspond to lower marginal costs. Consequently, the monopolist would like to produce more units as θ increases.

However, the monopolist does not observe the value of θ and only knows that it is uniformly distributed on the unit interval. Prior to any production decision, she observes a signal realization s from an information structure that fully reveals the state of the world ($s = \theta$) with probability $0 \leq \rho \leq 1$ and is completely uninformative ($\theta \perp s \sim U[0, 1]$) with probability $1 - \rho$. The quality of the information structure is increasing in ρ: When $\rho = 0$, signals are always uninformative; when $\rho = 1$, signals always fully reveal the state of the world.

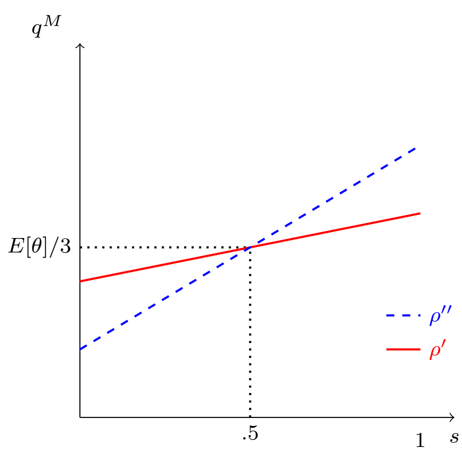
From an "interim" perspective, a monopolist that observes a signal realization s from an information structure of quality ρ optimally produces

$$q^M(s; \rho) = \underbrace{\frac{E[\theta]}{3}}_{\substack{\text{Monopoly quantity} \\ \text{based only on the prior}}} + \underbrace{\rho \left(\frac{s - E[\theta]}{3} \right)}_{\substack{\text{Adjustment based} \\ \text{on observed signal realization}}} .$$

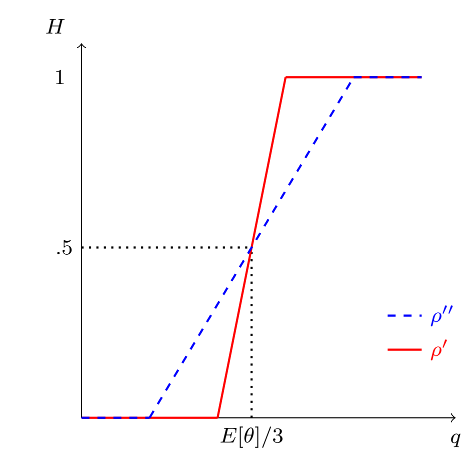
The first term on the right-hand-side is the quantity the monopolist produces absent an informative signal. The second term reflects how the monopolist adjusts her production decisions when she is able to learn something about her cost parameter. However, from an ex-ante perspective, the signal realization is a random variable yet to be observed by the monopolist. Thus, the induced optimal quantity $q^M(s;\rho)$ is also a realization of a random variable whose distribution is given by $H(z;\rho)$, the probability that the monopolist optimally produces at most z units given an information structure of quality ρ.[1]

Our goal in this paper is to characterize how $H(\cdot;\rho)$ changes when ρ increases. Will the optimal quantity produced increase or decrease on average when ρ increases? Will the quantities produced become more dispersed? In this example, we can answer these questions by using the closed-form solution of q^M.

Suppose the quality of the monopolist's information structure increases from ρ' to $\rho'' > \rho'$. "Good news" ($s > E[\theta]$) from ρ'' is a stronger evidence of high values of θ than "good news" from ρ'. As a result, the monopolist produces more when she observes "good news" from ρ'' than when she observes "good news" from ρ'.[2] Symmetrically, "bad news" ($s < E[\theta]$) from ρ'' is a stronger evidence of low values of θ than "bad news" from ρ'. As a result, the monopolist produces less when she observes "bad news" from ρ'' than when she observes "bad news" from ρ'. In either case, the monopolist makes more extreme decisions under the higher quality of information. In Figure 1(a), the rotation of the solid red line, $q^M(\cdot;\rho')$, to the dashed blue line, $q^M(\cdot;\rho'')$, captures the more extreme production decision due to an increase in the quality of information. This in turn induces a mean-preserving spread in the distribution H as shown in Figure 1(b).



(a) Quantity produced



(b) Distribution of optimal quantities

Figure 1

[1] $H(z;\rho) = \Pr(\{s : q^M(s;\rho) \leq z\})$.

[2] Recall that the higher the value of θ, the lower the marginal cost.

Now, suppose the monopolist chooses the quality of her information structure. The higher the quality, the more information the monopolist acquires. Consider a social planner who cannot regulate quantities or prices but can influence the monopolist's information acquisition.[3] Should the planner encourage or discourage information acquisition by the monopolist? Equivalently, do consumers prefer the distributions of quantities induced by more informative experiments?

The consumer surplus for a given quantity q is given by the function $CS(q) = q^2/2$. From the discussion above, increasing the quality of an information structure induces a mean-preserving spread in the distribution of optimal quantities produced by the monopolist. The convexity of the consumer surplus function implies that consumers benefit from such a mean-preserving spread.[4] Hence, the planner should encourage the monopolist's information acquisition, for example, by subsidizing information acquisition costs.

In the monopolist example, we made several assumptions, such as linear demand and quadratic cost functions, to simplify the analysis. This paper builds the tools to address the comparative statics of a higher quality of information in a more general model with supermodular payoffs, i.e., actions and the state of the world are complements. We present (i) an order over the distributions of optimal actions that captures changes in the mean and variability, (ii) an order over information structures that captures quality, and (iii) conditions on payoff functions that connect the two orders.

Consider two information structures, ρ' and ρ'', and compare the distributions of optimal actions they induce. We say the actions induced by ρ'' are more *responsiveness with higher mean* than those induced by ρ' if any risk-loving third party (e.g. the social planner in the above example) prefers the distribution of optimal actions induced by ρ''. Alternatively, we say the actions induced by ρ'' are more *responsiveness with lower mean* than those of ρ' if any risk-averse third party prefers the distribution of optimal actions induced by ρ'. Loosely, responsiveness with a higher mean corresponds to higher variability and higher actions on average (increasing convex stochastic order) while responsiveness with lower mean corresponds to higher variability but lower actions on average (second-order stochastic dominance).

To compare the quality of information structures, we first restrict attention to a class of experiments in which higher signal realizations lead to first-order stochastic shifts in posterior beliefs. The restriction is weaker than the common assumption that signals are ordered by the monotone likelihood ratio property (MLRP). Within this restricted class of experiments,

[3] Athey and Levin (2001) consider a similar problem. However, in their application, the planner can regulate prices/quantities as well as information.

[4] Specifically, the expected consumer surplus, $\frac{1}{2}\int z^2 dH(z;\rho)$, is increasing in ρ.

we then use the *monotone information order* to capture quality.[5] Information structure ρ'' dominates information structure ρ' in the monotone information order if, on average, high signals from ρ'' are a stronger evidence of high values of the state (than are high signals from ρ') and low signals from ρ'' are a stronger evidence of low values of the state (than are low signals from ρ'). Intuitively, the signals from ρ'' are more positively correlated with the state and thus, ρ'' is more informative than ρ'.

If ρ'' dominates ρ' in the monotone information order, we show that optimal actions induced by ρ'' are *more responsive with a higher mean* than the optimal actions induced by ρ' when payoffs exhibit increasing complementarities between actions and states. Furthermore, we show that the monotone information order is necessary to characterize responsiveness for all payoffs that exhibit increasing complementarities between actions and states. We also present corresponding results linking *responsiveness with lower mean* and decreasing complementarities between actions and states.

We then extend our comparative statics results to Bayesian games with strategic complementarities, which encompass such games as differentiated Bertrand competition and global games. We consider a setting where different players have information structures of varying quality. The key observation is that the distribution of a player's action in a Bayesian Nash equilibrium become more responsive not only to an increase in the quality of own information but also to an increase in the quality of other players' information.

As an application, we study endogenous information acquisition in Bayesian games with two players. The game is composed of two stages: only player 1 acquires information in the first stage followed by a second stage in which both players choose actions simultaneously. Whether or not player 2 observes player 1's choice of information corresponds to overt and covert information acquisition games respectively. We define the value of *transparency* as the difference in the marginal utility to player 1 between the overt and the covert games and we show how responsiveness is useful to characterize it. Specifically, we show that the value of *transparency* is positive or negative depending on (i) the responsiveness of player 2's action to player 1's information, and (ii) the sign of the externality on player 1 imposed by player 2's action. This in turn has implications on how much information player 1 acquires in the two games.

[5]The monotone information order is equivalent to the supermodular stochastic ordering, and the positive dependence ordering when the state is one-dimensional.

1.1 Related Literature

Our paper is closely related to Jensen (2017) who also studies the comparative statics of distributions. In his paper, the agent observes all relevant parameters when making a decision. His paper links changes in the distribution of parameters to the changes in the induced distribution over optimal actions. For instance, in the context of our motivating example, the monopolist observes the state θ and optimally produces quantity $q^M(\theta)$.[6] The interest is in how different distributions of θ change the subsequent distribution of $q^M(\theta)$. In our setting, the monopolist does not observe the state and the prior distribution of θ is held fixed. Instead, we characterize how different distributions over the posterior beliefs affect the monopolist's production decision.

Our work is also closely related to Lu (2016), who studies how informative signals affects choice from a menu. In particular, he shows that a decision-maker has a more dispersed willingness-to-pay for any given menu if the quality of information increases. We instead show that the choice from within a menu becomes more dispersed as the quality of information increases.[7]

Our paper is also related to the monotone comparative statics literature of single agent optimization problems: Topkis (1978), Milgrom and Shannon (1994), Athey (2002), and Quah and Strulovici (2009). Athey (2002) marks the first milestone on problems involving uncertainty and shows when optimal actions are increasing as a function of beliefs. We take the next step and ask how the distribution of optimal actions changes as a function of the distribution over beliefs.[8] Our key conceptual contribution is introducing an ex-ante comparative statics, responsiveness.

Our work is also related to the literature on the value of information which was studied by Blackwell (1951), Lehmann (1988), Persico(2000), and Athey and Levin (2001). Athey and Levin show that in the class of payoff functions that exhibit complementarities between actions and states, an agent values more information if, and only if, information quality is increasing in the monotone information order. Our results complement theirs in that we show in the subclass of payoff functions with increasing/decreasing complementarities, the agent's optimal actions are more responsive if, and only if, information quality is increasing in the monotone

[6] $q^M(\theta) = q^M(s;\rho)$ at $\rho = 1$: the quantity produced under complete information.

[7] For instance, there cannot be any meaningful dispersion in choice of action from within a singleton menu. However, the willingness-to-pay for the singleton menu can vary depending on the decision-maker's private information.

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

information order.

Our comparative statics results to games with strategic complementarities are also a natural extension of Vives (1990), Milgrom and Roberts (1994), Villas-Boas (1997), and Van Zandt and Vives (2007). At the same time, our work relates to the literature on information acquisition in games such as Persico (2000), Hellwig and Veldkamp (2009), Myatt and Wallace (2012), Colombo, Femminis and Pavan (2014), and more recently Yang (2015), Amir and Lazzati (2016), Denti (2016), Tirole (2015), and Pavan (2016). None of these papers, however, look at the dispersion of actions as a comparative static.

Moreover, our analysis of the value of *transparency* in Bayesian games is related to the characterization of strategic investment in sequential versus simultaneous games of complete information in Fudenberg and Tirole (1984) and Bulow, Geanakoplos and Klemperer (1985). We defer a detailed discussion of the relationship to Section 4.

The remainder of the paper is structured as follows: In section 2, we present the single agent framework, introduce *responsiveness*, and provide sufficient and necessary conditions for actions to become more responsive as information quality increases. Section 3 extends the analysis to Bayesian games with strategic complementarities. Section 4 applies *responsiveness* to analyze the value of *transparency* in Bayesian games. Section 5 concludes. Any proofs skipped in the text are in the Appendix.

1.2 Preliminary Definitions and Notation

Let X_i, $i = 1, 2, \ldots, m$, and Y be compact subsets of $\mathbb{R}$. Let $X \triangleq \times_{i=1}^{m} X_i$ be the Cartesian product endowed with the product order so that for $x', x \in X$, $x' \geq x$ if, and only if, $x'_i \geq x_i$ for $i = 1, 2, \ldots, m$. Let $x' \vee x$ denote the join of x' and x, the component-wise maximum, and let $x' \wedge x$ denote the meet of x' and x, the component-wise minimum.

A function $g : X \to \mathbb{R}$ is supermodular (submodular) if $g(x' \vee x) + g(x' \wedge x) \geq (\leq) g(x') + g(x)$ for all $x, x' \in X$. We say that g is modular if it is both supermodular and submodular. We use the terms 'increasing', and 'decreasing' in the weak sense, for example, we say a function $f : Y \to \mathbb{R}$ is increasing if $y' > y$ implies $f(y') \geq f(y)$. We will be explicit when we refer to strict monotonicity. A function $h : X \times Y \to \mathbb{R}$ has increasing (decreasing) differences in $(x; y)$ if for $x' \geq x$, $h(x', y) - h(x, y)$ is increasing (decreasing) in y.

For a differentiable function, $g : X \to \mathbb{R}$, we write $g_{x_i}(x)$ as a shorthand for $\frac{\partial}{\partial x_i} g(x)$ and $g_{x_i x_j}(x)$ for $\frac{\partial^2}{\partial x_i x_j} g(x)$. If g is differentiable and supermodular, then $g_{x_i x_j} \geq 0$ for all $i \neq j$.

2 Single Agent Problem

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

(A.1) $u(\theta, a)$ is uniformly bounded, measurable in θ, and twice differentiable in a,

(A.2) for all $\theta \in \Theta$, $u(\theta, \cdot)$ is strictly concave in a with $u_{aa}(\theta, \cdot) < 0$,

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

(A.4) $u(\theta, a)$ is supermodular.

Supermodularity implies that the agent's action and the unknown state of the world are complements. That is, the agent prefers a high action when the state is high. Assumptions (A.2)-(A.3) allow us to characterize the optimal actions by their first order conditions. Furthermore, the assumptions guarantee that the agent's problem has an interior solution.[9]

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

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

The compactness of A and the continuity of the utility function guarantee that the solution exists and is measurable. For any two beliefs $\mu_1, \mu_2 \in \Delta(\Theta)$, we say that μ_2 first-order stochastically dominates μ_1, denoted $\mu_2 \succeq_{FOSD} \mu_1$, if $\mu_1(\theta) \geq \mu_2(\theta)$ for all $\theta \in \Theta$. An implication of (A.4) is that $a^*(\mu_2) \geq a^*(\mu_1)$ whenever $\mu_2 \succeq_{FOSD} \mu_1$.

Prior to decision-making, the agent can observe an informative signal about the unknown state. Signals are generated by an information structure $\Sigma_\rho \triangleq (S, \{F(\cdot|\theta; \rho)\}_{\theta \in \Theta})$ where $S \subseteq \mathbb{R}$ is the signal space, $F(\cdot|\theta; \rho) : S \to [0, 1]$ is a probability measure over S conditional on a given state θ, and ρ is an index that is useful when comparing multiple signal structures. For each $s \in S$, we assume that $F(s|\theta; \rho)$ is measurable in θ. Let

$$F_S(s; \rho) = \int_{\Theta} F(s|\theta; \rho) \mu(d\theta)$$

[9] In Section 2.4, we discuss the difficulties associated with violations of these assumptions.

denote the the marginal of the signal. We assume that all information structures have the same marginal on the signal, i.e., $F_S(\cdot;\rho) = F_S(\cdot)$ for any Σ_ρ. Moreover, $F_S(\cdot)$ has a positive bounded density $f_S(\cdot)$. The assumption is without loss of generality: we can apply the integral probability transform to any signal with a continuous marginal distribution $F_S(\cdot;\rho)$ and create a new signal which is uniformly distributed on the unit interval. The transformed signal still conveys the same information as the original signal. If $F_S(\cdot;\rho)$ is discontinuous, then, as noted by Lehmann (1988), we can construct a new equally informative signal with a continuous marginal by appropriately distributing the mass at discontinuity points.[10]

2.1 Order over Distributions of Optimal Actions

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

From an ex-ante perspective, the signal realizations are yet to be observed. Therefore, the optimal actions induced by an information structure Σ_ρ are random variables. In particular, $a(\rho)$ is a random variable that is distributed according to $H(\cdot;\rho)$ defined as

$$H(z;\rho) \triangleq F_S\big(\{s : a(s;\rho) \le z\}\big)$$

for $z \in \mathbb{R}$. The quantile function is defined as

$$\hat{a}(q;\rho) = \inf\{z : q \le H(z;\rho)\}$$

for $q \in (0,1)$.

Our goal is to characterize how information quality affects the distribution of optimal actions. Thus, the first step is to identify an order over distributions of optimal actions that appropriately captures changes in the mean and dispersion of actions.

Responsiveness: Given two information structures $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$, we say that

i. $a(\rho'')$ is more responsive with a lower mean than $a(\rho')$ if, and only if, for any increasing concave function $\phi : \mathbb{R} \to \mathbb{R}$

[10] If $F_S(\cdot;\rho)$ is discontinuous at s^* with $F_S(s^*;\rho) = q$, then construct a new signal $\tilde{s}$ with $\tilde{s} = s$ if $s < s^*$, $\tilde{s} = s + qT$ if $s = s^*$, and $\tilde{s} = s + q$ if $s > s^*$, where $T \sim U(0,1)$.

$$E\Big[\phi \circ a(\rho')\Big] \geq E\Big[\phi \circ a(\rho'')\Big],$$

and

ii. $a(\rho'')$ is more responsive with a higher mean than $a(\rho')$ if, and only if, for any increasing convex function $\varphi : \mathbb{R} \to \mathbb{R}$

$$E\Big[\varphi \circ a(\rho'')\Big] \geq E\Big[\varphi \circ a(\rho')\Big]$$

In other words, $a(\rho'')$ is more responsive with a lower mean than $a(\rho')$ if $a(\rho')$ second-order stochastically dominates $a(\rho'')$. Similarly, $a(\rho'')$ is more responsive with a higher mean than $a(\rho')$ if $a(\rho'')$ dominates $a(\rho')$ in the increasing-convex stochastic order. If $a(\rho'')$ is both more responsive with a lower mean and more responsive with a higher mean than $a(\rho')$, then $a(\rho'')$ is a mean-preserving spread of $a(\rho')$. The definitions capture a notion of increased variability in the optimal actions along with changes in the expectation. As a short hand, we say $a(\rho'')$ is more responsive than $a(\rho')$ if it is either more responsive with a higher mean or more responsive with a lower mean.

Lemma 1 below provides equivalent characterizations of responsiveness. The first equivalence provides an alternate definition by comparing the distribution functions of the optimal actions. The second equivalence characterizes responsiveness as a comparison of the quantile functions. These alternative definitions are particularly useful when the optimal actions are monotone in the signal realization, a natural consequence when payoffs are supermodular and beliefs are ordered by first-order stochastic dominance.

Lemma 1 *The following are equivalent:*

i. $a(\rho'')$ *is more responsive with lower mean than* $a(\rho')$.

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

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

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

$$\int_{0}^{t} \hat{a}(q;\rho')dq \geq \int_{0}^{t} \hat{a}(q;\rho'')dq.$$

Similarly, the following are equivalent:

iv. $a(\rho'')$ *is more responsive with higher mean than* $a(\rho')$.

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

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

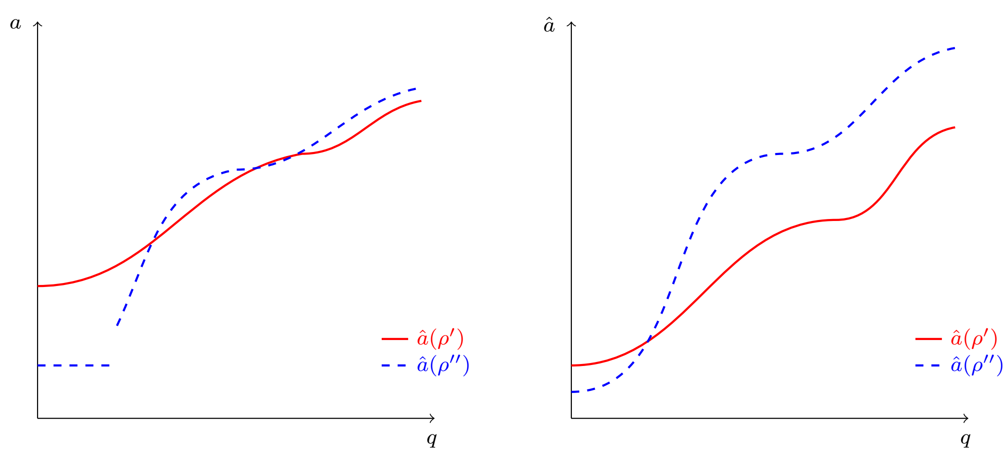
vi. For all $t \in [0,1]$*,*

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

Proof. The equivalence of $i.$ and $ii.$ $[iv.$ and $v.]$ follows from Shaked and Shanthikumar (Stochastic Orders, 2007. Theorem 4.A.2). The equivalence of $i.$ and $iii.$ $[iv.$ and $vi.]$follows from Shaked and Shanthikumar (Stochastic Orders, 2007. Theorem 4.A.3).

■

Figure 2 below plots the quantile functions induced by two information structures $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$. In Figure 2(a), the area under the quantile function $\hat{a}(\rho')$ (the solid red curve) is larger than that of $\hat{a}(\rho'')$ (the dashed blue curve) which implies the expectation of the optimal actions induced by $\Sigma_{\rho'}$ is higher than the optimal actions induced by $\Sigma_{\rho''}$. Furthermore, integrating $\hat{a}(q;\rho') - \hat{a}(q;\rho'')$ left to right always yields a non-negative value which, by Lemma 1. iii, implies Responsiveness with a lower mean. In contrast, in Figure 2(b), the area under $\hat{a}(\rho')$ is less than that of $\hat{a}(\rho'')$ which implies the expectation of the optimal actions induced by $\Sigma_{\rho'}$ is less than the optimal actions induced by $\Sigma_{\rho''}$. Furthermore, integrating $\hat{a}(q;\rho'') - \hat{a}(q;\rho')$ right to left always yields a non-negative value which, by Lemma 1. vi, implies Responsiveness with a higher mean.



(a) Responsiveness with a lower mean

(b) Responsiveness with a higher mean

Figure 2: Quantile Function and Responsiveness

2.2 The Monotone Information Order

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

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

Monotone Information Order: $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the monotone information order, denoted $\rho'' \succeq_{MIO} \rho'$, if for all $q \in [0,1]$

$$\mu\big(\cdot\,|F_S(s) \geq q;\rho''\big) \succeq_{FOSD} \mu\big(\cdot\,|F_S(s) \geq q;\rho'\big)$$

and

$$\mu\big(\cdot\,|F_S(s) \leq q;\rho'\big) \succeq_{FOSD} \mu\big(\cdot\,|F_S(s) \leq q;\rho''\big).$$

When $\rho'' \succeq_{MIO} \rho'$, the signal and the state are more positively correlated under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$. Intuitively, the agent has more faith in the signal realizations from $\Sigma_{\rho''}$ than the signal realizations from $\Sigma_{\rho'}$. By (A.5), high signal realizations are evidence of high states. The agent considers a signal realization above the q^{th} quantile from $\Sigma_{\rho''}$ as a stronger evidence that the state could be high (than a signal realization above the q^{th} quantile from $\Sigma_{\rho'}$). Consequently, the agent is more optimistic when she observes a signal realization above the q^{th} quantile from $\Sigma_{\rho''}$ than from $\Sigma_{\rho'}$. Similarly, a signal realization below the q^{th} quantile from $\Sigma_{\rho''}$ is a stronger evidence that the state could be low (than a signal realization below the q^{th} quantile from $\Sigma_{\rho'}$). Thus, the agent is more pessimistic when she observes a signal realization below the q^{th} quantile from $\Sigma_{\rho''}$ than from $\Sigma_{\rho'}$.

Example 1: Truth-or-Noise signals
Σ_ρ belongs to a class of information structures such that with probability $\rho \in [0,1]$, the signal reveals the state ($s = \theta$), and with probability $1-\rho$, the signal and the state are identically and independently distributed. Thus, with probability $1-\rho$, the signal is uninformative. Then, $\rho'' \succeq_{MIO} \rho'$ if $1 \geq \rho'' > \rho' \geq 0$.

Example 2: Normal additive noise
Σ_ρ belongs to a class of information structures such that conditional on the state θ, the signal is given by $s = \theta + \varepsilon$, where $\varepsilon \perp \theta$, $\varepsilon \sim \mathcal{N}(0, \rho^{-2})$, and $\rho > 0$. As ρ increases, the variance of the

additive noise shrinks (the precision of the signal increases). Then, $\rho'' \succeq_{MIO} \rho'$ if $\rho'' > \rho' > 0$.

Example 3: Uniform signals
Σ_ρ belongs to a class of information structures such that conditional on the state θ, the signal is distributed according to

$$s \sim U\big[\theta - \rho, \theta + \rho\big]$$

for $\rho > 0$. As ρ shrinks, the signal becomes more accurate. Then, $\rho'' \succeq_{MIO} \rho'$ if $\rho' > \rho'' > 0$.[11]

We provide the following lemma that gives alternative characterizations of monotone information order.

Lemma 2 *Given two information structures $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$, $\rho'' \succeq_{MIO} \rho'$ if, and only if,*

i. For all $\theta \in \Theta$ and all $x \in \mathbb{R}$,

$$\int_{-\infty}^{x} \mu(\theta|s;\rho'') f_S(s) ds \geq \int_{-\infty}^{x} \mu(\theta|s;\rho') f_S(s) ds,$$

and

$$\int_{x}^{\infty} \mu(\theta|s;\rho'') f_S(s) ds \leq \int_{x}^{\infty} \mu(\theta|s;\rho') f_S(s) ds.$$

ii. For all (integrable) supermodular functions $\psi : \Theta \times S \to \mathbb{R}$,

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

Proof. Fix any $\theta \in \Theta$ and any $x \in \mathbb{R}$. Let $F_S(x) = q$.

$$\begin{aligned}\int_{-\infty}^{x} \Big(\mu(\theta|s;\rho'') - \mu(\theta|s;\rho')\Big) dF_S(s) =& \int_{-\infty}^{x} F(\theta, ds; \rho'') - F(\theta, ds; \rho') \\ =& F(\theta, x; \rho'') - F(\theta, x; \rho') \\ =& F(\theta, F_S^{-1}(q); \rho'') - F(\theta, F_S^{-1}(q); \rho') \\ =& \Big(\mu(\theta|F_S(s) \leq q; \rho'') - \mu(\theta|F_S(s) \leq q; \rho')\Big) \underbrace{F_S(F_S^{-1}(q))}_{=q}\end{aligned}$$

[11]This is also a class of information structures that are ordered by MIO but not by Blackwell informativeness.

We then have

$$\rho'' \succeq_{MIO} \rho' \Leftrightarrow \mu(\theta|F_S(s) \leq q; \rho') \succeq_{FOSD} \mu(\theta|F_S(s) \leq q; \rho'')$$

$$\Leftrightarrow \Big(\mu(\theta|F_S(s) \leq q; \rho'') - \mu(\theta|F_S(s) \leq q; \rho')\Big) q \geq 0$$

giving us the desired result. By Bayes consistency,

$$\int_{-\infty}^{\infty} \mu(\theta|s;\rho'') f_S(s) ds = \int_{-\infty}^{\infty} \mu(\theta|s;\rho') f_S(s) ds = \mu(\theta),$$

which then proves the second expression of Lemma 2.i is also equivalent to MIO. Lemma 2.ii follows from Müller and Stoyan (2002), Theorem 3.9.5. ■

There is a close parallel between the characterization of responsiveness in Lemma 1 and the characterization of monotone information order in Lemma 2. This close connection between the two definitions provides us a way to link changes in posteriors to changes in the distribution of actions.

It is natural to ask why the monotone information order is the relevant order to consider. Athey and Levin (2001) answer this question by showing that all decision makers with supermodular payoffs prefer a "more informative" signal if, and only if, informativeness is ranked by the monotone information order. We complement their result by providing conditions on the marginal utilities of supermodular payoff functions such that, actions are more responsive to "more informative" signals if, and only if, informativeness is ranked by the monotone information order.

2.3 Monotone Information Order and Responsiveness

The final step is to find a class of payoff functions for which the optimal actions can be ranked by responsiveness when information quality increases according to the monotone information order. Let $\mathcal{U}^{\uparrow}$ 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 is

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

ii. supermodular in (θ, a).

Below, we show that payoffs in $\mathcal{U}^{\uparrow}$ are linked to optimal actions that become more responsive with a higher mean (hence the up arrow) as information quality increases in the monotone information order.

Let $\mathcal{U}^{\downarrow}$ 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 is

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

ii. submodular in (θ, a).

Payoffs in $\mathcal{U}^{\downarrow}$ are linked to optimal actions that become more responsive with a lower mean (hence the down arrow) as information quality increases in the monotone information order.

Theorem 1 *Consider two information structures $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$.*

i. If $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the monotone information order, then for all payoffs payoff $u \in \mathcal{U}^{\uparrow}$ $\left[u \in \mathcal{U}^{\downarrow}\right]$, $a(\rho'')$ is more responsive with a higher [lower] mean than $a(\rho')$.

ii. Suppose the prior μ^o is absolutely continuous on Θ. If $\Sigma_{\rho''}$ does not dominates $\Sigma_{\rho'}$ in the monotone information order, then there exists a payoff $u \in \mathcal{U}^{\uparrow} \bigcap \mathcal{U}^{\downarrow}$ such that $a(\rho'')$ is not more responsive than $a(\rho')$.

Before giving the intuition for Theorem 1, it is helpful to first explore why these class of payoff functions are interesting. We show that a payoff function $u \in \mathcal{U}^{\uparrow}$ $\left[u \in \mathcal{U}^{\downarrow}\right]$ leads to optimal actions that are "convex" ["concave"] in the agent's posterior belief. We then show how this convexity/concavity interacts with the informativeness of signals to result in responsiveness.

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

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

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

Henceforth, we focus on payoffs in $\mathcal{U}^{\uparrow}$ but the intuition and the arguments we provide can be symmetrically applied to payoffs in $\mathcal{U}^{\downarrow}$. Supermodularity of the agent's payoff function implies that the state and the action are complements, i.e., for two states $\theta > \theta'$, the difference in the marginal utility $u_a(\theta, a) - u_a(\theta', a)$ is non-negative. However, supermodularity does not tell us anything about the strength of the complementarities between the action and the state.

It could be that the complementarities are negligible when a is low but substantial when a is high. In such a case, the agent's payoff additionally satisfies *increasing supermodularity*, i.e., the marginal utility $u_a(\theta, a) - u_a(\theta', a)$ is increasing in a. Therefore, when a is high, the agent is more willing to increase her action if the state is high.

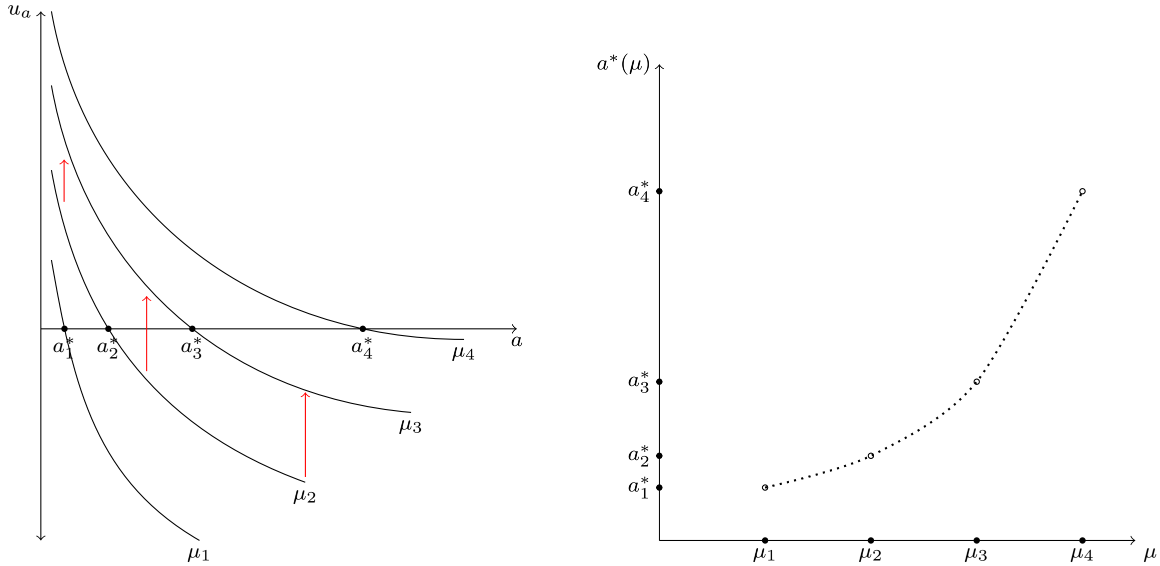
On the other hand, the concavity of the agent's payoff function implies that she has diminishing marginal utility. When a is high, she is less willing to increase her action regardless of the state. Thus, there are two opposing forces at work. When the agent's payoff u belongs to $\mathcal{U}^{\uparrow}$, the rate at which her marginal utility diminishes is less than the rate at which the complementarities between her action and the state increase.

Figure 3(a) below plots out the expected marginal utility of a payoff function $u \in \mathcal{U}^{\uparrow}$ for different beliefs $\mu_i \in \Delta(\Theta)$, $i = 1, 2, 3, 4$. Since the payoff is concave in a, the marginal utilities are downward sloping. The optimal action $a^*(\mu_i)$ is given by the action at which the expected marginal utility under belief μ_i intersects the x-axis. The beliefs are ordered by first-order stochastic dominance with $\mu_4 \succeq_{FOSD} \mu_3 \succeq_{FOSD} \mu_2 \succeq_{FOSD} \mu_1$. Supermodularity implies that the expected marginal utility of μ_i always lies below the expected marginal utility of μ_{i+1}. Thus, $a^*(\mu_4) \geq a^*(\mu_3) \geq a^*(\mu_2) \geq a^*(\mu_1)$.

Furthermore, *increasing supermodularity* implies that the gap between the expected marginal utilities of μ_{i+1} and μ_i is increasing as the action increases. We capture this by showing that the height of the red arrows increases left to right. Finally, the marginal utilities themselves are convex curves which implies that the marginal utility diminishes at a diminishing rate. All these properties combined give us the "convexity" property described in Proposition 1. Figure 3(b) depicts this for the case when beliefs can be plotted on a one-dimensional axis.

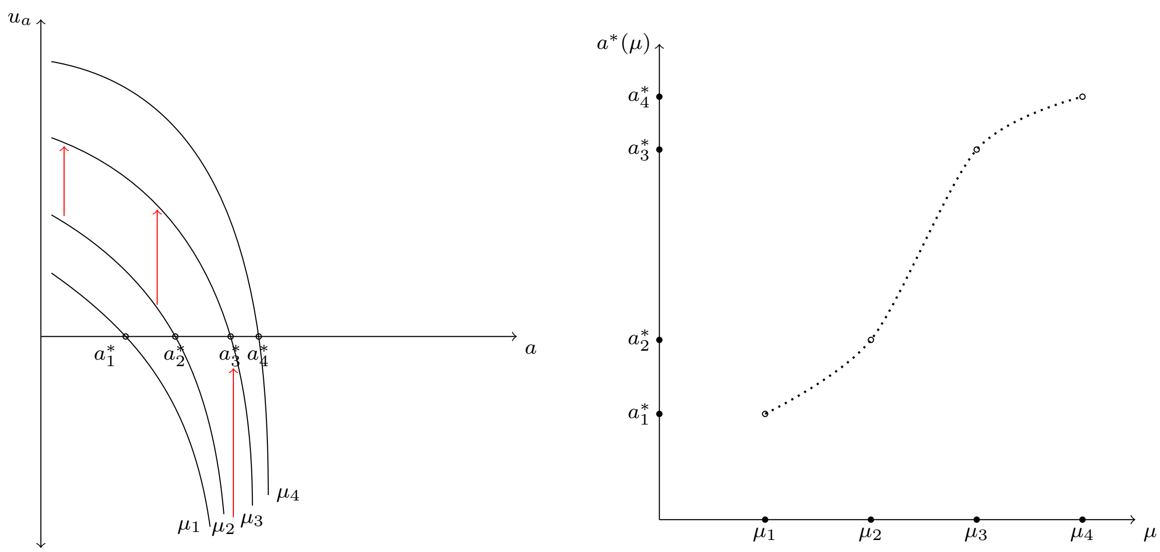
Figure 4 below shows why *increasing supermodularity* alone is not sufficient to get the convexity property from Proposition 1. In Figure 4(a), we plot the expected marginal utility of a payoff function $u \notin \mathcal{U}^{\uparrow}$ for different beliefs $\mu_i \in \Delta(\Theta)$, $i = 1, 2, 3, 4$. Once again, beliefs are ordered by first-order stochastic dominance with $\mu_4 \succeq_{FOSD} \mu_3 \succeq_{FOSD} \mu_2 \succeq_{FOSD} \mu_1$. Supermodularity still holds – the expected marginal utility of μ_i lies below the expected marginal utility of μ_{i+1}. Thus, $a^*(\mu_{i+1}) \geq a^*(\mu_i)$. Furthermore, *increasing supermodularity* still holds – the height of the red arrows increases left to right. However, the marginal utilities are now concave which implies that the marginal utility diminishes at an increasing rate. Figure 3(b) then depicts the resulting non-convex optimal action as a function of beliefs.

To see how the "convexity" of the optimal action is related to responsiveness, consider a setting where the prior puts mass only on two states $\{\theta_L, \theta_H\}$ with $\underline{\theta} \leq \theta_L < \theta_H \leq \bar{\theta}$. We abuse notation and use μ_o to denote the mass at θ_H, i.e., $\mu_o = \mu^o(\{\theta_H\})$. Let $\Sigma_{\rho'}$ be a completely



(a) Marginal utilities (b) Optimal action

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



(a) Marginal utilities (b) Optimal action

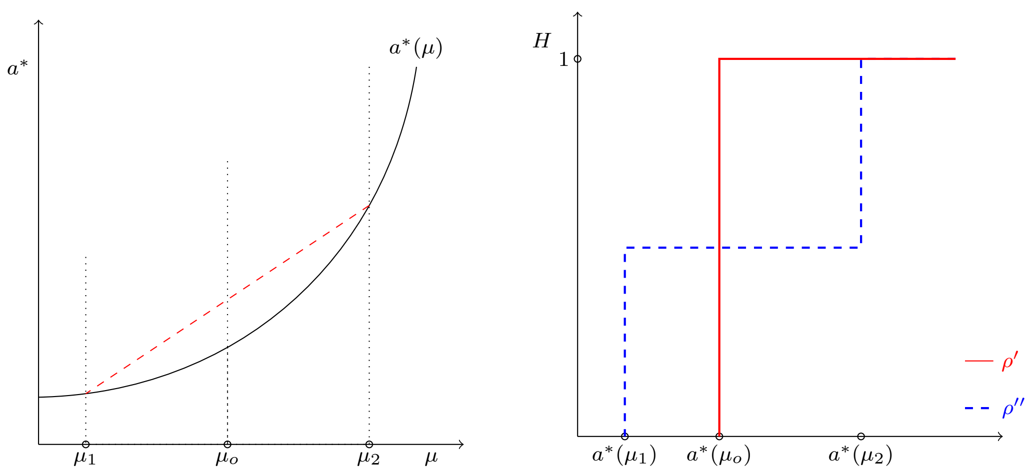
Figure 4: Non-convexity for $u \notin \mathcal{U}^{\uparrow}$

uninformative information structure. Then, $\Sigma_{\rho'}$ induces $a^*(\mu_o)$ with probability one.

Let $\Sigma_{\rho''}$ be an information structure that induces two posteriors $\{\mu_1, \mu_2\}$ with probability $\{\lambda, 1-\lambda\}$. Without loss of generality, assume $\mu_2 > \mu_1$ which implies $\mu_2 \succeq_{FOSD} \mu_1$. Consistency of Bayes-updating implies $\mu_o = \lambda\mu_1 + (1-\lambda)\mu_2$. $\Sigma_{\rho''}$ induces optimal actions $a^*(\mu_1)$ with probability λ and $a^*(\mu_2)$ with probability $1-\lambda$. Furthermore, given the supermodularity of $u(\theta, a)$ and $\mu_2 \succeq_{FOSD} \mu_1$, $a^*(\mu_2) \geq a^*(\mu_1)$.

From Proposition 1, if $u \in \mathcal{U}^{\uparrow}$, then $\lambda a^*(\mu_1)+(1-\lambda)a^*(\mu_2) \geq a^*(\lambda\mu_1+(1-\lambda)\mu_2) = a^*(\mu_o)$. In Figure 5(a) below, the average action from the more informative structure $\Sigma_{\rho''}$ is given by the point on the dashed red line directly above μ_o while the average action from the uninformative structure $\Sigma_{\rho'}$ is given by the point on the black line directly above μ_o.

Figure 5(b) maps the induced distribution over optimal actions. The dashed blue line, $H(\rho'')$, maps the distribution of actions under $\Sigma_{\rho''}$ with an atom of size λ at $a^*(\mu_1)$ and another atom of size $1-\lambda$ at $a^*(\mu_2)$. Similarly, the solid red line, $H(\rho')$, maps the distribution of actions under $\Sigma_{\rho'}$ which places all the mass at $a^*(\mu_o)$. Notice the integral $\int_x^\infty H(z;\rho'')-H(z;\rho')dz \leq 0$ for all $x \in \mathbb{R}$ which implies that $a(\rho'')$ is more responsive with a higher mean than $a(\rho')$.



(a) Optimal action

(b) Induced distribution

Figure 5: Convexity of a^* and responsiveness with higher mean

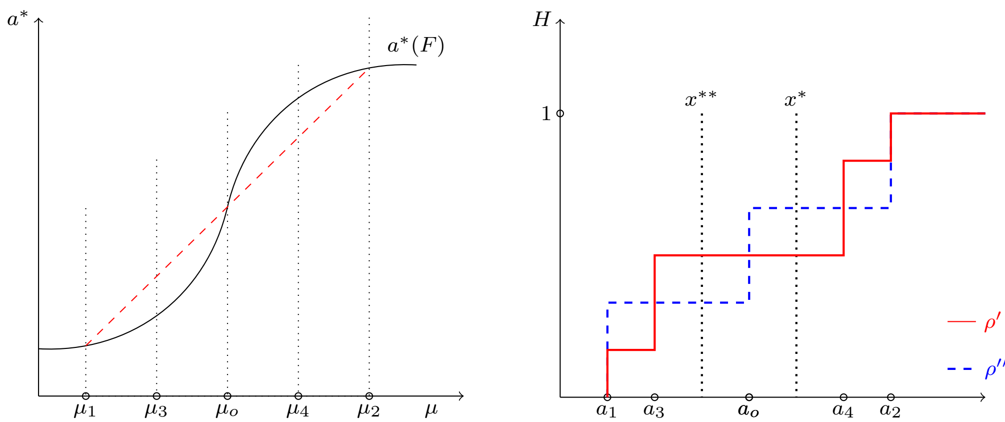
2.4 Non-responsive Optimal Actions

In this section, we explore why a higher quality of information may not lead to more responsive optimal actions when $u \notin \mathcal{U}^{\uparrow} \cup \mathcal{U}^{\downarrow}$. Figure 6 below illustrates why we may fail to rank actions by responsiveness when the optimal action is neither convex nor concave over posteriors that are first-order stochastically ordered.

Once again, consider a simplified setting in which the prior puts mass only on two states $\{\theta_L, \theta_H\}$. Let $\Sigma_{\rho''}$ be an information structure that induces three posteriors $\{\mu_1, \mu_o, \mu_2\}$ with equal probability such that $\mu_2 \succeq_{FOSD} \mu_o \succeq_{FOSD} \mu_1$. Let $\Sigma_{\rho'}$ induce posteriors $\{\mu_1, \mu_2, \mu_3, \mu_4\}$

with probability $\{1/6, 1/6, 1/3, 1/3\}$ such that $\mu_4 \succeq_{FOSD} \mu_3 \succeq_{FOSD} \mu_2 \succeq_{FOSD} \mu_1$. Notice that $\Sigma_{\rho'}$ is a garbling of $\Sigma_{\rho''}$ and thus, $\rho'' \succeq_{MIO} \rho'$.[12]

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 6(a) below, this corresponds to the point of intersection of the dashed red line and the solid black curve at μ_o. Figure 6(b) maps the distribution over optimal actions. $\Sigma_{\rho''}$ induces the dashed blue line while $\Sigma_{\rho'}$ induces the solid red line. If we start integrating from the right, then $\int_x^{\infty} H(z;\rho'') - H(z;\rho')dz \leq 0$ for all $x > a^*(\mu_4)$ but the sign changes at some point $x^* \in (a^*(\mu_o), a^*(\mu_4))$. If we integrate from the left, then $\int_{-\infty}^{x} H(z;\rho'') - H(z;\rho')dz \geq 0$ for all $x < a^*(\mu_3)$ but the sign changes at some point $x^{**} \in (a^*(\mu_3), a(\mu_o))$.



(b) Induced distribution

(a) Optimal action

Figure 6: Non-convexity/concavity and non-responsiveness

We can therefore conclude that $a(\rho'')$ and $a(\rho')$ cannot be ordered by 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.[13]

Figure 7 below illustrates another reason why a higher quality of information may not lead to more responsiveness; when the interior solution assumption, (A.3), is violated. Figure 7 depicts the optimal action as a function of beliefs. In the interior of the action space, the optimal action is convex in beliefs. Thus, as long as information structures induce beliefs that

[12] A garbling is a kernel $Q : S \times S \to [0,1]$ so that $F(s'|\theta;\rho') = \int_{s\in S} Q(s'|s)F(ds|\theta;\rho'')$

[13] Shaked and Shantikumar (2007) provide a thorough treatment of these orders.

lead to actions in $(\underline{a}, \bar{a})$, we can apply the same arguments we used in Figure 5 to conclude actions become more responsive with a higher mean as information quality increases in the monotone information order.

However, there are some beliefs for which the upper limit on the action space, $\bar{a}$, is a binding constraint. Suppose the prior is one such belief so that $a^*(\mu_o) = \bar{a}$. Let $\Sigma_{\rho'}$ be a completely uninformative information structure. Then, $\Sigma_{\rho'}$ induces $\bar{a}$ with probability one. Furthermore, any distribution over actions induced by any other information structure Σ_ρ will be first-order stochastically dominated, even if $\rho \succeq_{MIO} \rho'$. Thus, the optimal actions induced by the least informative information structure are actually more responsive with a higher mean than all other information structures.

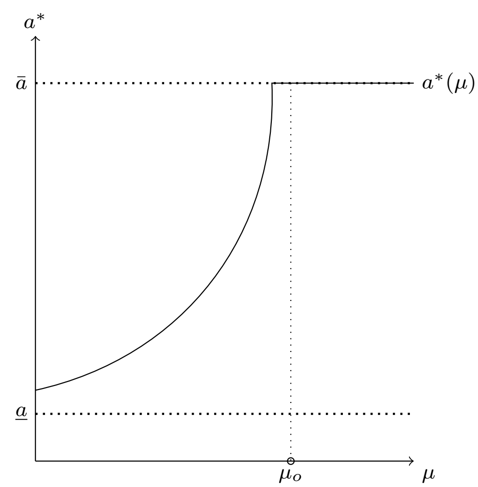


Figure 7: Boundary solution and non-responsiveness

2.5 Application: Pigouvian Subsidies and Monopoly Production

In the motivating example of the Introduction, we asked how a social planner, who cannot regulate prices or quantities, improves welfare by affecting the quality of information a monopolist uses to make production decisions. We used highly structured information structures and payoff functions to conclude the planner should encourage more information acquisition by the monopolist. In this subsection, we consider the example in a more general setting.

In this section, 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 and the planner share a common prior $\mu^o \in \Delta(\Theta)$. As θ increases, the marginal cost declines, i.e. $c(\theta, q)$ is submodular in (θ, 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 θ.

Prior to making any production decisions, the monopolist can acquire an information structure from a set of experiments $\{\Sigma_\rho\}_{\rho\in\mathbb{R}}$ at cost $k(\rho)$. We assume the experiments are totally ordered by the monotone information order so that $\rho'' > \rho'$ implies $\rho'' \succeq_{MIO} \rho'$ and $k(\rho'') \geq k(\rho')$.

Given a choice of information structure Σ_ρ and a signal realization $s \in S$, the monopolist updates her belief to the posterior $\mu(\cdot|s;\rho)$ and produces the monopolist optimal quantity $q^M(s;\rho)$. Thus, the monopolist's ex-ante problem is to choose an information structure that maximizes

$$\int_{\Theta\times S} \pi\big(\theta, q^M(s;\rho)\big)\mu(d\theta|s;\rho)dF_S(s) - k(\rho).$$

In contrast, the social planner takes the consumer surplus into account. Let $CS(q)$ be the consumer surplus when the monopolist produces q units. The planner's ex-ante choice of information structure maximizes

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

Thus, the planner has a higher demand for information than the monopolist if a higher quality of information increases the expected consumer surplus, i.e., information is a positive externality on the consumer even if the unknown parameter θ has no direct effect on consumer welfare.

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

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 $\pi \in \mathcal{U}^\uparrow$, an increase in ρ (higher quality of information by MIO) leads to an optimal action $q^M(\rho)$ that is more responsive with a higher mean. By definition of responsiveness with a higher mean, $E[CS(q^M(\rho)]$ is increasing in ρ. ■

Intuitively, the assumption that $-qP''(q)/P'(q) \leq 1$ implies that as production increases, the consumers capture more and more of the welfare gains than does the monopolist.[14] Therefore, the consumer surplus is a convex function of the quantity produced, which in turn implies that consumers benefit as the monopolist production becomes more responsive with higher

[14] The assumption implies that the gap between the inverse demand curve and the marginal revenue curve is widening as quantity increases. We can interpret the assumption as saying that the monopolist has diminishing market power.

mean. We can then use Theorem 1 to identify sufficient conditions that lead to the desired responsiveness behavior.

3 Supermodular Games

In this section, we extend our results from the single-agent framework to supermodular games with incomplete information. This class of games includes beauty contests, oligopolistic competition, games with network effects, search models, and investment games. It is useful to understand how information quality affects the equilibrium of these games in a general setting.

3.1 Setup

There are n players with $N \triangleq \{1, 2, \ldots, n\}$ denoting the set of players. Let $\Theta_i \triangleq [\underline{\theta}_i, \bar{\theta}_i]$ be the state space for player i. Let $\Theta = \times_{i \in N} \Theta_i$ and $\Theta_{-i} = \times_{j \neq i} \Theta_j$. The players hold a common prior $\mu^o \in \Delta(\Theta)$. Once again, we allow for beliefs to be discrete measures with finite support in Θ, absolutely continuous, or a mixture. Additionally, we assume that

(A.6) for all $i \in N$, $\theta_i' > \theta_i$ implies $\mu^o(\theta_{-i}|\theta_i') \succeq_{FOSD} \mu^o(\theta_{-i}|\theta_i)$

which is a weaker assumption than affiliation.

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

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

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

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

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

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 now strategic complementarities between the players' actions. Thus, when player j takes a higher action, player i wants to do the same.

Notice that our setup accommodates games of independent private values and common values. For the former, we simply need to impose an additional assumption that the prior is independent, i.e., $\mu^o(\theta) = \prod_{i \in N} \mu^o(\theta_i)$. The latter requires that the prior has support only on states with $\theta_i = \theta_j$ for all $i, j \in N$.

Following the terminology introduced by Bergemann and Morris (2016), we decompose the entire game of incomplete information into two components: the basic game and the information structure. The basic game $G \triangleq (N, \{A_i, u^i\}_{i \in N}, \mu^o)$ 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 $u^i : \Theta_i \times A \to \mathbb{R}$, and (iii) a common prior $\mu^o \in \Delta(\Theta)$.

The second component of the Bayesian game is the information structure $\Sigma_\rho = \times_{i \in N} \Sigma_{\rho_i}$ where for each player $i = 1, ..., n$, signals are generated by $\Sigma_{\rho_i} \triangleq \big(S_i, \{F(\cdot|\theta_i; \rho_i)\}_{\theta_i \in \Theta_i}\big)$. $S_i \subseteq \mathbb{R}$ is the signal space, $F(\cdot|\theta_i, \rho_i) : S_i \to [0, 1]$ is a probability measure over S_i conditional on a given state θ_i, and ρ_i is an index.[15] For each $s_i \in S_i$, we assume that $F(s_i|\theta_i; \rho_i)$ is measurable in θ_i. Let $F_{S_i}(\cdot; \rho_i)$ be the marginal of the signal. Once again, we assume without loss of generality that for any information structure Σ_{ρ_i}, $F_{S_i}(\cdot; \rho_i) = F_{S_i}(\cdot)$. Moreover, F_{S_i} has a positive and bounded density f_{S_i}.

Let $S = \times_{i \in N} S_i$. An information structure Σ_ρ induces a joint distribution over $S \times \Theta$ which we denote by $\boldsymbol{F}(s, \theta; \rho)$. The following are working assumptions for this section:

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

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

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

Assumption (A.11) implies that player i can only directly learn about θ_i. To see this, first notice that if players cannot directly learn about each other's signal realizations, then conditional on the state $\theta \in \Theta$, $\boldsymbol{F}(s|\theta; \rho) = \prod_{i \in N} \boldsymbol{F}(s_i|\theta; \rho)$ for all $s \in S$. Furthermore, if the signal s_i depends only on θ_i, then $\boldsymbol{F}(s_i|\theta; \rho) = F(s_i|\theta_i; \rho_i)$. Assumption (A.12) is an extension of (A.5) and implies that higher signal realizations lead to a first-order increase in a player's belief. Assumption (A.13) implies the converse: higher states are likely to lead to higher signal realizations. A distribution over the state and signal space that satisfies the monotone likelihood ratio property also satisfies (A.12)-(A.13).

[15] There is an implicit assumption in the setup that player i can directly learn only about θ_i. We make this assumption explicit in (A.11).

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

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

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

In 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 s_i and her opponents use strategies $a_{-i}^\star(\rho)$. Fixing the basic game G, we are interested in how a change in the information structure from $\Sigma_{\rho'}$ to $\Sigma_{\rho''}$ affects the BNEs of the full game $\mathcal{G}_{\rho'} = (\Sigma_{\rho'}, G)$ and $\mathcal{G}_{\rho''} = (\Sigma_{\rho''}, G)$.

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 beliefs when they participate in the basic game G, which is relevant in our setting as we do not impose any smoothness restriction on the joint distribution of signals and state.

3.2 Monotone Information Order and Equilibrium Responsiveness

We parallel the single-agent framework as closely as possible. We first extend the responsiveness order over optimal actions and the monotone information order over information structures into a multi-player setting using the product order. We then identify the class of payoff functions for which BNE are ordered by responsiveness when information quality changes according to the monotone information order.

Equilibrium Responsiveness: Given two games of incomplete information $\mathcal{G}_{\rho'}$ and $\mathcal{G}_{\rho''}$,

[16] By assumptions (A.6), (A.10), and (A.12), player i's best response is monotone 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), and Van Zandt and Vives (2007)).

$a^\star(\rho'')$ is more responsive with a higher [lower] mean than $a^\star(\rho')$ if, and only if, $a_i^\star(\rho'')$ is more responsive with a higher [lower] mean than $a_i^\star(\rho')$ for all $i \in N$.

Monotone Information Order: Given two information structures $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$, $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the monotone information order, denoted $\rho'' \succeq_{MIO} \rho'$ if, and only if, $\Sigma_{\rho''_i}$ dominates $\Sigma_{\rho'_i}$ in the monotone information order for all $i \in N$.

Let $\mathcal{P}^\uparrow$ be the class of payoff functions $u : \Theta_i \times A \to \mathbb{R}$ 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 $(\theta_i, a_{-j}) \in \Theta_i \times A_{-j}$,

ii. has increasing differences in $(\theta_i; a_j)$ for all $a_{-j} \in A_{-j}$ and,

iii. has increasing differences in $(a_i; a_j)$ for all $(\theta_i, a_{-j,-i}) \in \Theta_i \times A_{-j,-i}$.

Below, we show that payoffs in $\mathcal{P}^\uparrow$ are linked to BNE strategies that become more responsive with a higher mean (hence the up arrow) as information quality increases in the monotone information order. Notice that we now require the marginal utility of player i to be convex not only in player i's action but also in each opponent's action. Similarly, we now require the marginal utility of player i to be supermodular not only in (θ_i, a_i) but also in (θ_i, a_{-i}) and (a_i, a_{-i}).

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

i. is convcave in a_j for all $(\theta_i, a_{-j}) \in \Theta_i \times A_{-j}$,

ii. has decreasing differences in $(\theta_i; a_j)$ for all $a_{-j} \in A_{-j}$ and,

iii. has decreasing differences in $(a_i; a_j)$ for all $(\theta_i, a_{-j,-i}) \in \Theta_i \times A_{-j,-i}$.

Below, we show that payoffs in $\mathcal{P}^\downarrow$ are linked to BNE strategies that become more responsive with a lower mean (hence the down arrow) as information quality increases in the monotone information order.

Theorem 2 *Consider two Bayesian games $\mathcal{G}_{\rho'}$ and $\mathcal{G}_{\rho''}$.*

Suppose for each player $i \in N$, $u^i \in \mathcal{P}^\uparrow$ [$u^i \in \mathcal{P}^\downarrow$]. If $\Sigma_{\rho''}$ dominates $\Sigma_{\rho'}$ in the monotone information order, then for any monotone Bayesian Nash equilibrium $a^\star(\rho')$ of $\mathcal{G}_{\rho'}$, there exists *a monotone Bayesian Nash equilibrium $a^\star(\rho'')$ of $\mathcal{G}_{\rho''}$ such that $a^\star(\rho'')$ is more responsive with higher [lower] mean than $a^\star(\rho')$.*

Each player i faces n sources of uncertainty: the unknown state θ_i and the random actions of the remaining $n-1$ players which depend on the signal realizations they observe. The proof for Theorem 2 proceeds in four steps. The first step shows that, holding all else fixed, player i's best-reply strategy becomes more responsive when only the quality of information for player i increases in the monotone information order. As quality of information increases, player i is more informed about θ_i. Thus, an application of Theorem 1 from the single-agent setting gives the result.

The second step shows that, holding all else fixed, player i's best-reply strategy becomes more responsive when the quality of information for player $j \neq i$ increases in the monotone information order. As player j's information quality increases, player j's signals become more correlated to the state θ_j, which in turn is (weakly) correlated to θ_i.[17] Thus, by increasing the quality of information for player j, the signals for player i and j indirectly become more correlated. Hence, player i can better predict player j's random action and match it.

The third step shows that, holding all else fixed, player i's best-reply strategy becomes more responsive when player $j \neq i$ chooses a more responsive strategy due to the increasing differences of $u^i_{a_i}$ in $(a_j; a_i)$. It is of similar spirit to the result that increasing differences of u^i in $(a_j; a_i)$ imply that the best-reply to a monotone opponent's strategy is monotone. Finally, we conclude by applying the main result in Villas-Boas (1997) to get a comparative statics of responsiveness on fixed points.

4 Overt vs Covert Information Acquisition and the Value of Transparency

In this section, we illustrate the use of responsiveness to characterize the value of *transparency*, as defined below, in the context of Bayesian games with information acquisition. We then highlight why the value of *transparency* is of economic interest.

We consider a Bayesian game, $\mathcal{G}$, with two players, and composed of two stages: an information acquisition stage followed by the stage game $G = (\{A_i, u^i\}_{i=1,2}, F_\Theta)$, as defined in Section 3. In the information acquisition stage, only player 1 acquires information by choosing a signal structure from $\{\Sigma_\rho = (\mathcal{S}_1, \{F_{\mathcal{S}_1|\Theta}(\cdot|\theta;\rho)\}_{\theta\in\Theta})\}_{\rho\in\mathcal{P}}$ at cost $\kappa(\rho)$. We assume that $\rho' > \rho$

[17] By weakly correlated, we mean that we allow for θ_i to be independent of θ_j.

implies $\Sigma_{\rho'}$ dominates Σ_ρ in the monotone information order. On the other hand, player 2 has an exogenously given information structure $\Sigma_{\rho''} = (\mathcal{S}_2, \{F_{\mathcal{S}_2|\Theta}(\cdot|\theta;\rho'')\}_{\theta\in\Theta})$.

Whether or not player 2 observes player 1's choice of information structure corresponds to the overt and the covert game respectively. After the information acquisition stage, each player $i = 1, 2$ privately observes a signal s_i, updates beliefs, and plays the stage game by simultaneously choosing an action. Throughout this section, we assume that best-responses in the stage game are in pure strategies and characterized by their first order conditions. We also assume that there exists a unique pure strategy BNE.

Consider the following two scenarios as a thought experiment. Suppose both players start with exogenously given information structures, $\Sigma_{\hat{\rho}'}, \Sigma_{\rho''}$, and there is common knowledge of the information structure, i.e both players know the Bayesian game is $\mathcal{G}_{\hat{\rho}',\rho''} = (\Sigma_{\hat{\rho}'}, \Sigma_{\rho''}, G)$. Suppressing dependence on ρ'', let $a_1^\star(\hat{\rho}'), a_2^\star(\hat{\rho}')$ be the pure strategy BNE.

In the first scenario, player 1 is allowed to choose a different information structure. Player 2 (i) is made aware that player 1 can choose a different information structure, and (ii) observes player 1's choice. This scenario of the thought experiment mirrors the the overt game. Common knowledge of information structures still holds; if player 1 chooses $\Sigma_{\rho'}$, the resulting pure strategy BNE is $a_1^\star(\rho'), a_2^\star(\rho')$.

In the second scenario, player 1 is again allowed to choose a different information structure. However, player 2 (i) is not aware that player 1 can choose a different information structure, (ii) does not observe player 1's choice. This scenario of the thought experiment mirrors the covert game. Player 2 will ignorantly continue playing $a_2^\star(\hat{\rho})$, even when player 1 chooses $\Sigma_{\rho'}$. On the other hand, player 1 best-replies to $a_2^\star(\hat{\rho})$ by playing the pure strategy $a_1(\rho', \hat{\rho}')$. Henceforth, we refer to $\hat{\rho}$ as player 2's belief of player 1's choice, and ρ as player 1's actual choice. We say player 2 has correct beliefs when $\hat{\rho} = \rho$.

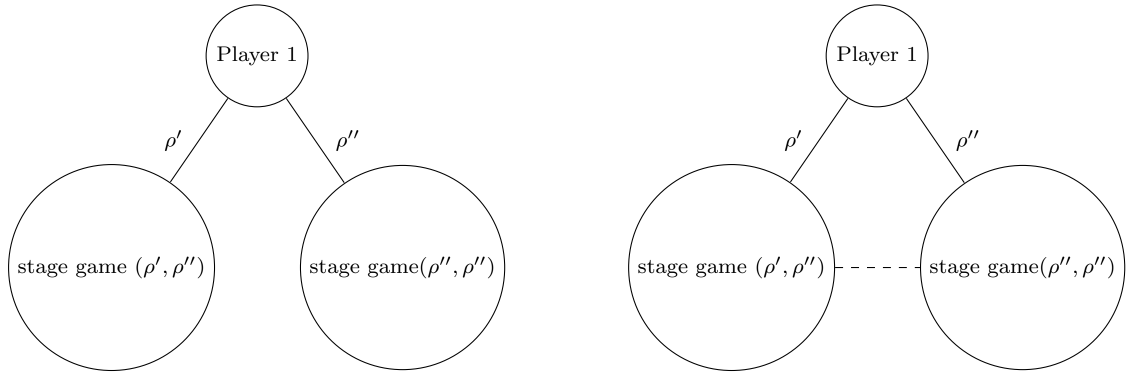
We define the value of *transparency* as the difference in ex-ante payoffs to player 1 between the two scenarios. Specifically, let player 1's ex-ante payoff in the covert game (second scenario), given actual signal choice ρ, and given player 2's belief $\hat{\rho}$ be

$$U_1(\rho, \hat{\rho}) = \int_{\Theta\times\mathcal{S}_1\times\mathcal{S}_2} u^1\big(\theta, a_1(s_1;\rho,\hat{\rho}), a_2^\star(s_2;\hat{\rho})\big) dF(\theta, s_1, s_2; \rho', \rho'').$$

Player 1's ex-ante payoff in the overt game (first scenario), given actual choice ρ, is $U_1(\rho, \rho)$.

Definition 1 *The Value of Transparency from choice of information structure ρ at belief $\hat{\rho}$ is given by:*

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



(a) Overt Information Acquisition (b) Covert Information Acquisition

Figure 8: The diagrams illustrate the information set of player 2 before signals are realized in the overt and covert games in a very simple version of the game where player 1 only has two available signals.

The Marginal Value of Transparency is given by:

$$MVT(\rho, \hat{\rho}) = \frac{d}{d\rho}\Big(U_1(\rho, \rho) - U_1(\rho, \hat{\rho})\Big)$$

We first show that the marginal value of *transparency* is key to compare player 1's demand for information across the two games. We then characterize when the marginal value of *transparency* is positive or negative depending on (i) player 2's responsiveness to a change in player 1's information structure, and (ii) the sign of the externality imposed on player 1 by player 2's responsiveness.

Before we discuss how to characterize the marginal value of transparency, we present why it is an interesting concept through the following propositions.

Proposition 3 *Suppose $\kappa(\rho) = 0$ for all $\rho \in \mathcal{P}$. If the Marginal Value of Transparency at correct beliefs ($\hat{\rho} = \rho$) is non-negative,*

$$MVT(\rho, \rho) \geq 0$$

for all $\rho \in \mathcal{P}$, then player 1's Value of Information in the Overt game, $U_1(\rho, \rho)$, is increasing in ρ.

In covert games, the more information player 1 acquires, the more knowledgeable she is about the unknown state and can make better decisions in the game. Therefore, if information is costless, the Value of Information in Covert games is increasing in the amount of information she acquires (Neyman, 1989, Amir and Lazzati, 2016).

While acquiring more costless information has the same positive effect in overt games, there are additional effects to account for. Player 2 can observe how much information player 1 acquires, and respond to it during the stage game. Player 2 may find it optimal to choose an unfavorable action (punish player 1) in the stage game whenever player 1 acquires more information in the first stage. If player 2's unfavorable action is strong enough on average, player 1's Value of Information in Overt games may decrease despite becoming more informed. Proposition 3 provides conditions, fully captured by the marginal value of *transparency*, for more information to be beneficial even in overt games.

Furthermore, when the Value of Information in Overt games is concave (for example when $\kappa(\cdot)$ is convex enough), Proposition 4 allows us to use the marginal value of *transparency* to make comparative statics predictions on the demand of player 1's information acquisition.

Proposition 4 *Let the Value of Information in the Overt game, $U_1(\rho,\rho)$, be concave in ρ. If the $MVT(\rho,\rho) \geq 0$ $[\leq 0]$ for all $\rho \in \mathcal{P}$, then player 1 acquires more [less] information in the Overt game than the Covert game.*

We now discuss how to characterize the marginal value of *transparency*. The value of transparency depends on how the choice of a signal by player 1 affects the actions of player 2. In Lemma 5, we identify two key components: player 2's responsiveness to changes in player 1's information structure, and the externality player 2's response imposes on player 1.

Proposition 5 *Let $a_2^\star(s_2;\rho)$ be differentiable in ρ, for all $s_2 \in \mathcal{S}_2$. The Marginal Value of Transparency at correct beliefs is*

$$MVT(\rho,\rho) = \int_{\mathcal{S}_2} \frac{\partial a_2^\star(s_2;\rho)}{\partial \rho} \left\{ \int_{\Theta\times\mathcal{S}_1} \frac{\partial u^1\Big(\theta, a_1^\star(s_1;\rho), a_2^\star(s_2;\rho)\Big)}{\partial a_2} dF_{\Theta\times\mathcal{S}_1}(\theta, s_1|s_2;\rho,\rho'') \right\} dF_{\mathcal{S}_2}(s_2;\rho'')$$

When player 2's actions, $a_2^\star(\rho')$ is monotone in s_2 and monotonically responsiveness (with a higher mean) in ρ', the term

$$\int_t^\infty \frac{\partial a_2^\star(s_2;\rho)}{\partial \rho} dF_{\mathcal{S}_2}(s_2;\rho'') \geq 0, \;\; \forall t \in \mathbb{R}$$

as it represents player 2's marginal responsiveness to small changes in player 1's information structure. The term in the brackets is the expected externality player 2 imposes on player 1, conditional on signal realization s_2. The characterization of the marginal value of *transparency* depends on how these two terms interact.

Theorem 3 (The Marginal Value of Transparency)

1. For increasing responsiveness:

- *If $a_2^\star(\rho)$ is increasing in responsiveness with increasing (decreasing) mean, and the externality $\frac{\partial u^1}{\partial a_2}$ is positive, supermodular (submodular) in (a_1, θ) and increasing (decreasing) in a_2.*

  Then the Marginal Value of Transparency is positive (negative) for all ρ.

- *If $a_2^\star(\rho)$ is increasing in responsiveness with increasing (decreasing) mean, and the externality $\frac{\partial u^1}{\partial a_2}$ is negative, submodular (supermodular) in (a_1, θ) and decreasing (increasing) in a_2.*

  Then the Marginal Value of Transparency is negative (positive) for all ρ.

2. For decreasing responsiveness:

- *If $a_2^\star(\rho)$ is decreasing in responsiveness with decreasing (increasing) mean, and the externality $\frac{\partial u^1}{\partial a_2}$ is positive, supermodular (submodular) in (a_1, θ) and increasing (decreasing) in a_2.*

  Then the Marginal Value of Transparency is negative (positive) for all ρ.

- *If $a_2^\star(\rho)$ is decreasing in responsiveness with decreasing (increasing) mean, and the externality $\frac{\partial u^1}{\partial a_2}$ is negative, submodular (supermodular) in (a_1, θ) and decreasing (increasing) in a_2.*

  Then the Marginal Value of Transparency is positive (negative) for all ρ.

To gain some intuition on this result it is useful to observe that the combined assumptions on the externality together with the assumptions on Σ_ρ and $\Sigma_{\rho''}$ imply that *the average externality conditional on the information of player 2*

$$\mathbb{E}\left(\frac{\partial u^1}{\partial a_2} \middle| s_2 \right)$$

is monotonic in player 2's belief. For example, if the externality $\frac{\partial u^1}{\partial a_2}$ is supermodular in (a_1, θ) and increasing in a_2, then $\mathbb{E}\left(\frac{\partial u^1}{\partial a_2} \middle| s_2 \right)$ is increasing as a function of s_2. Then the table in figure 9 summarizes theorem 3.

A Taxonomy of Transparency

$a_2^\star(\rho)$		externality		Marginal Value
responsiveness	mean	$\frac{\partial u^1}{\partial a_2}$	$\mathbb{E}\left(\frac{\partial u^1}{\partial a_2}\middle\vert s_2\right)$	of Transparency
$\nearrow$	$\nearrow$	$+$	$\nearrow$	$+$
$\nearrow$	$\searrow$	$+$	$\searrow$	$-$
$\nearrow$	$\nearrow$	$-$	$\searrow$	$-$
$\nearrow$	$\searrow$	$-$	$\nearrow$	$+$
$\searrow$	$\searrow$	$+$	$\nearrow$	$-$
$\searrow$	$\nearrow$	$+$	$\searrow$	$+$
$\searrow$	$\searrow$	$-$	$\searrow$	$+$
$\searrow$	$\nearrow$	$-$	$\nearrow$	$-$

Figure 9: The table summarizes theorem 3 by stating some assumptions and combining others into the monotonic average externality condition.

Perhaps the most direct way to interpret the results is through Figure 8. Looking at columns 2 and 3, one can say that if the expectation of $a_2^\star(\rho)$ is increasing in ρ and there is a positive externality the marginal value of transparency is positive. However, one has to be careful making this assertion since it is only a necessary condition. Looking at row 1 and 6 in the table, the conclusion also depends on columns 1 and 4.

More in general, whenever the marginal value of transparency is positive columns 2 and 3 have to 'match' (in the sense that they are either increasing and positive or decreasing and negative), and columns 1 and 4 have to match too (both increasing or both decreasing). Notice that of the possible $2^4 = 16$ combinations of assumptions one could possibly fit in the table the characterization only includes 8. For the other 8 cases the result is ambiguous and the marginal value of transparency can be positive or negative.

4.1 Relation to the Taxonomy of Strategic Behavior

The application of responsiveness to characterize the value of transparency is related to the taxonomy of strategic behavior in Fudenberg and Tirole (1984), and Bulow, Geanakoplos and Klemperer (1985). For a thorough treatment of different examples and applications we recommend Shapiro (1986). For a more recent treatment using the tools of supermodular games see Vives (2000). Here we follow the textbook treatment of Tirole (1988).

To make the comparison transparent, we look only at the case of entry accommodation in duopoly. There are 2 firms and 2 periods; in the first period firm 1 can make an investment K_1, in the second period both firms compete either in quantities (strategic substitutes) or prices (strategic complements). The term investment is used in a very broad sense and can represent for example an action that lowers firm 1's marginal costs or the action of capturing a share of the market. For the taxonomy, the relevant feature of the investment is the effect on the subgame perfect equilibrium payoff of firm 2: if the effect is positive, the investment makes firm 1 look *soft*. On the contrary, if the effect is negative, the investment makes firm 1 look *tough*.

Fudenberg and Tirole show

$$sign \frac{dU_1}{dK_1} = sign\left(\frac{da_2}{da_1}\right) * sign\left(\frac{\partial U_2}{\partial a_1}\frac{da_1}{dK_1}\right)$$

where on the right hand side the first term is positive in the case of strategic complements (price competition) and negative in the case of strategic substitutes (quantity competition). The second term is the perception of *toughness* or *softness* . They proceed to give a taxonomy based on the 4 possible combinations of the right hand side signs.

	strategic complements	strategic substitutes
tough	− (puppy dog)	+ (top dog)
soft	+ (fat cat)	− (lean and hungry)

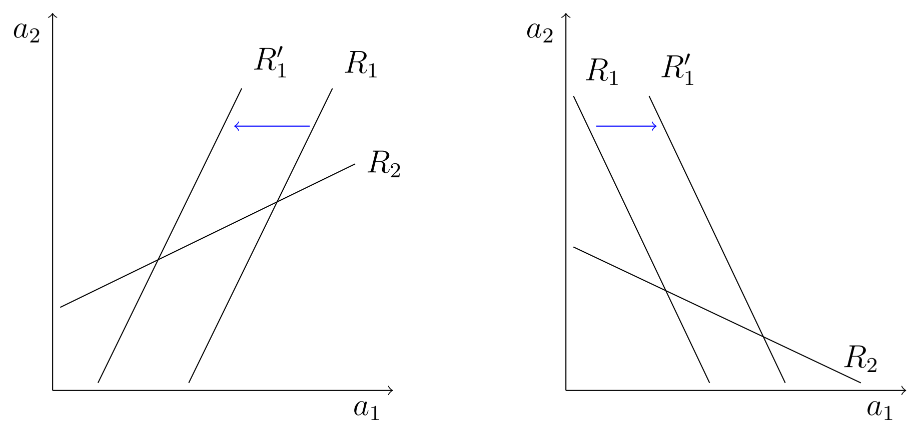
In our model, the investment K_1 corresponds to the player 1's information structure ρ'. Under the symmetric payoff assumption in Fudenberg and Tirole (1984) $sign\left(\frac{\partial U_2}{\partial a_1}\right) = sign\left(\frac{\partial U_1}{\partial a_2}\right)$. Therefore,

$$\begin{aligned} sign \frac{dU_1}{dK_1} &= sign\left(\frac{da_2}{da_1}\right) * sign\left(\frac{\partial U_2}{\partial a_1}\frac{da_1}{dK_1}\right) \\ &= sign\left(\frac{\partial U_1}{\partial a_2}\right) * sign\left(\frac{da_2}{dK_1}\right). \end{aligned}$$

The second line above is the deterministic version of the marginal value of *transparency* we derived in Lemma 5. It is therefore related to the characterization in Theorem 3, and corresponds to columns 2 and 3 in the Figure 8: the sign of the externality and the effect of K_1 on a_2.

Given the uncertainty in our model, we need to also account for responsiveness. The reason we do not use the Fudenberg and Tirole model as an analogy to derive our characterization is that responsiveness is not necessarily related to the strategic complementarities $sign\left(\frac{da_2}{da_1}\right)$. We showed in Theorem 3 that when there is full complementarities between a_1, a_2 and the state θ, increasing the quality of player 1's information will increase responsiveness of player 2, but this is not true in general.

Therefore, although there is a clear connection between this section and Fudenberg and Tirole (1984), the fact that now we have uncertainty shifts the focus from strategic complementarities to responsiveness. We have found that responsiveness is not characterized by strategic complementarities or substitutabilities but there is certainly a relationship that needs to be further explored. We leave the task for future work.



(a) Strategic Complements

(b) Strategic Substitutes

Figure 10: A decrease in the marginal cost for firm 1 shifts her reaction inwards in the case of price competition and shifts the reaction outwards in the case of quantity competition. In Fudenberg and Tirole's terminology investment (in lowering costs) makes firm 1 look tough which is beneficial in the case of strategic substitutes but firm 1 would rather look soft in the case of strategic complements.

5 Conclusion

We have advanced the Monotone Comparative Statics program in several directions. First of all we conceptualized how from an ex-ante perspective the optimal actions of an informed decision maker are endogenous random variables. A natural question then was comparing optimal actions as a function of the quality of the information the decision maker had, therefore we introduced the notion of "responsiveness" and characterized how it captures changes in the mean and dispersion of actions.

Furthermore, we gave sufficient conditions on payoffs so that actions are more responsive to higher quality of information. The conditions had an interpretation as risk sensitivity increasing with the state of the world. After that we extended the conditions to show that actions are more responsive to higher quality of information in games with complementarities.

In the final section we provided a taxonomy of the value of transparency in overt vs covert information acquisition games using the concepts and results of 'ex-ante comparative statics', showing in the way to the last result a variety of applications.

We expect that the methods of 'ex-ante comparative statics' will be specially fruitful when studying bayesian persuasion with restricted information structures, signal jamming games, and models of rational inattention and search. Moreover, methods based on monotone responses have been particularly useful in the intersection of industrial organization and econometrics.

REFERENCES

AMIR, R. and N. LAZZATI (2016): "Endogenous information acquisition in Bayesian games with strategic complementarities," *Journal of Economic Theory*, 163, 684-698.

ATHEY, S (2002): "Monotone Comparative Statics under Uncertainty," *Quarterly Journal of Economics*, CXVII(1), 187-223.

ATHEY, S. and J. LEVIN (2001): "The Value of Information in Monotone Decision Problems," *Stanford Working Paper* 01-003.

BLACKWELL, D. (1951): "Comparison of experiments," In *Second Berkeley Symposium on Mathematical Statistics and Probability*, Vol. 1, 93-102.

BLACKWELL, D. (1953): "Equivalent Comparisons of Experiments," *The Annals of Mathematical Statistics*, 24(2), 265-272.

BERGEMANN, D., and S. MORRIS (2016): "Bayes Correlated Equilibrium and the Comparison of Information Structures in Games," *Theoretical Economics*, 11, 487-522

BULOW, J. I., J. D. GEANAKOPLOS and P. D. KLEMPERER (1985): "Multimarket Oligopoly: Strategic Complements and Substitutes," *Journal of Political Economy*, 93(3), 488-511.

COLOMBO, L., G. FEMMINIS, and A. PAVAN (2014): "Information Acquisition and Welfare," *Review of Economic Studies*, 81, 1438-1483.

DENTI. T. (2016): "Unrestricted Information Acquisition" *Working Paper*.

FUDENBERG, D. and J. TIROLE (1984): "The Fat-Cat Effect, the Puppy-Dog Ploy, and the Lean and Hungry Look," *American Economic Review (Papers and Proceedings)*, 74(2),361-366.

GANUZA, J.J., and J. PENALVA (2010): "Signal Ordering Based on Dispersion and the Supply of Private Information in Auctions," *Econometrica*, 78 (3), 1007-1030.

HELLWIG, C. and L. VELDKAMP (2009): "Knowing What Others Know: Coordination Motives in Information Acquisition," *Review of Economic Studies*, 76, 223-251.

JENSEN, M.K. (2016): "Distributional Comparative Statics," *Working Paper*.

LEHMANN, E. L. (1988): "Comparing Location Experiments," *The Annals of Statistics*, 16 (2), 521-533.

LEVIN, J. (2001): "Information and the Market for Lemons," *Rand Journal of Economics*, 32(4), 657-666.

LU, J, (2016): "Random Choice and Private Information," *Econometrica*, 84 (6), 19832027

MEYER, M.A. (1990): "Interdependence in Multivariate Distributions: Stochastic Dominance Theorems and an Application to the Measurement of Ex Post Inequality under Uncertainty," *Nuffield College D. P.* No. 49.

MEYER, M. A., and D. MOOKHERJEE (1987): “Incentives, Compensation and Social Welfare,” *Review of Economic Studies*, 45, 209-26.

MEYER, M.A., and B. STRULOVICI (2012): “Increasing Interdependence of Multivariate Distributions,” *Journal of Economic Theory*, 147, 1460-89.

MILGROM, P. and J. ROBERTS (1990): “Rationalizability, Learning, and Equilibrium in Games with Strategic Complementarities,” *Econometrica*, 58(6), 1255-1277.

MILGROM, P. and J. ROBERTS (1994): “Comparing Equilibria,“ *American Economic Review*, 84(3), 441-459.

MILGROM, P. and C. SHANNON (1994): “Monotone Comparative Statics,” *Econometrica,* 62(1), 157-180.

MULLER, A., D. STOYAN (2002): “Comparison Methods for Stochastic Models and Risks,” *John Wiley & Sons, Ltd.*

MYATT, D. P. and C. WALLACE (2012): “Endogenous Information Acquisition in Coordination Games,” *Review of Economic Studies*, 79, 340-374.

NEYMAN, A. (1989): “The Positive Value of Information,” *Games and Economic Behavior*, 3, 350-355 (1991)

PAVAN, A. (2016): “Attention, Coordination and Bounded Recall,” *Working Paper.*

PERSICO, N. (1996): “Information Acquisition in Affiliated Decision Problems.” *Discussion Paper* No. 1149, Department of Economics, Northwestern University.

PERSICO, N. (2000): “Information Acquisition in Auctions,” *Econometrica*, 68(1), 135-148.

QUAH, J. and B. STRULOVICI (2009): “Comparative Statics, Informativeness, and the Interval Dominance Order”, *Econometrica*, 77(6), 1949-1992.

SCHILLING, R. (2005): “Measures, Integrals and Martingales,” *Cambridge University Press*, Cambridge CB2 8RU, UK.

SHAKED, M. and G. SHANTHIKUMAR (2007): “Stochastic Orders,” *Springer-Verlag* New York, XVI, 474pp.

SHAPIRO, C. (1986): “Theories of Oligopoly Behavior,” In *Handbook of Industrial Organization*, ed. R. Schmalensee and R. Willig (Amsterdam: North Holland).

TIROLE, J. (1988): “The Theory of Industrial Organization,” *MIT Press*, 479pp.

TIROLE, J. (2015): “Cognitive Games and Cognitive Traps,” *Working Paper.*

TOPKIS, D. M. (1978): “Minimizing a Submodular Function on a lattice,” *Operations Research*, 26, 305-321.

VAN ZANDT, T., and X. VIVES (2007): “Monotone Equilibria in Bayesian Games of Strategic Complementarities,” *Journal of Economic Theory*, 134, 339-360.

VILLAS-BOAS, J.M. (1997): "Comparative Statics of Fixed Points," *Journal of Economic Theory*, 73, 183-198.

VIVES, X. (1990): "Nash Equilibrium with Strategic Complementarities," *Journal of Mathematical Economics*, 19, 305-21.

VIVES, X. (2000): "Oligopoly Pricing: Old Ideas and New Tools," *MIT Press*, 441pp.

YANG, M., 2015: "Coordination with Flexible Information Acquisition," *Journal of Economic Theory*, 158(B), 721-738.

6 Appendix

6.1 Proofs from Section 2

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

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

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

Proof of Theorem 1.i.

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

Without loss of generality, we assume that the marginal on signals is uniformly distributed on the unit interval.[18] For any two information structures $\rho'' \succeq_{MIO} \rho''$ and any signal realization

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

$s \in [0,1]$, 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 $u \in \mathcal{U}^\uparrow$, then $u_a(\theta, a)$ is convex in a for all θ. Thus,

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

and

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

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

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

where $\mathbb{1}_{[s\geq t]}$ is the indicator function that equals 1 if $s \geq t$ and 0 otherwise.
Define $\psi(\theta, s; t) \triangleq u_a(\theta, a(s;\rho'))B(s)^{-1}\mathbb{1}_{[s\geq t]}$. For any $\theta > \theta'$, $\psi(\theta, s; t) - \psi(\theta', s; t) = 0$ for $s < t$ and

$$\psi(\theta, s; t) - \psi(\theta', s; t) = B(s)^{-1}\Big(u_a(\theta, a(s;\rho')) - u_a(\theta', a(s;\rho'))\Big) \geq 0$$

for $s \geq t$. The inequality follows from the supermodularity of u in (θ, a) and the strict concavity of u in a. Since $u \in \mathcal{U}^\uparrow$, u_a is also supermodular in (θ, 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, $u \in \mathcal{U}^\uparrow$ implies that $-u_a$ is submodular in (θ, a) and concave in a. Hence,

$-u_{aa}(\theta, a)$ is decreasing in both θ and a. Since higher signal realizations lead to higher actions and to first-order stochastic shifts in beliefs,

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

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

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

where the last inequality follows from the characterization of monotone information order in Lemma 2. ■

Proof of Theorem 1.*ii*.

Proof. From Lemma 2, if $\rho'' \not\geq_{MIO} \rho'$, there exists a $\theta^* \in \Theta$ and a $t^* \in [0, 1]$ such that

$$\int_0^{t^*} \mu(\theta^*|s; \rho')dF_S(s) > \int_0^{t^*} \mu(\theta^*|s; \rho'')dF_S(s)$$

By continuity of μ, there is a neighborhood $\Theta^* \subseteq \Theta$ around θ^* such that for all $\theta \in \Theta^*$

$$\int_0^{t^*} \mu(\theta|s; \rho')dF_S(s) > \int_0^{t^*} \mu(\theta|s; \rho'')dF_S(s).$$

Let $[\underline{\theta}^*, \bar{\theta}^*] \subseteq \Theta^*$ be a compact selection. We want to construct a payoff function $u : \Theta \times A \to \mathbb{R}$ such that $u \in \mathcal{U}^{\uparrow} \bigcap \mathcal{U}^{\downarrow}$ and admits a solution that is not ordered by responsiveness.

Define the function $v : \Theta \to \mathbb{R}$ as

$$v(\theta) = \begin{cases} \underline{a} + \mu(\underline{\theta}^*)(\bar{a} - \underline{a}) & \text{if} \quad \theta < \underline{\theta}^* \\ \underline{a} + \mu(\theta)(\bar{a} - \underline{a}) & \text{if} \quad \underline{\theta}^* \leq \theta \leq \bar{\theta}^* \\ \underline{a} + \mu(\bar{\theta}^*)(\bar{a} - \underline{a}) & \text{if} \quad \theta > \bar{\theta}^* \end{cases}.$$

Note that $v(\theta)$ is bounded, measurable, increasing in θ, and absolutely continuous. Define a payoff

$$u(\theta, a) = v(\theta)a - \frac{a^2}{2}.$$

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 is supermodular in (θ, a). For each $\theta \in \Theta$, the marginal utility $u_a(\theta, a) = 0$ if $a = v(\theta) \in [\underline{a}, \bar{a}]$. Furthermore, the marginal utility $u_a(\theta, a) = v(\theta) - a$ is

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

ii. modular in (θ, a).

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

For any given Σ_ρ, notice that $a(s;\rho) = E[v(\theta)|s;\rho]$. Then given $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$,

$$\begin{aligned}
&\int_0^{t^*} \big(a(s;p_2) - a(s;\rho')\big)dF_S(s) \\
=&\int_0^{t^*}\int_\Theta v(\theta)\big(\mu(d\theta|s;\rho'') - \mu(d\theta|s;\rho')\big)dF_S(s) \\
=&\int_0^{t^*}\int_\Theta v'(\theta)\big(\mu(\theta|s;\rho') - \mu(\theta|s;\rho'')\big)d\theta dF_S(s) \\
=&(\underline{a} - \bar{a})\int_{\theta^*}^{\theta^{**}} \underbrace{\int_0^{t^*} \big(\mu(\theta|s;\rho') - \mu(\theta|s;\rho'')\big)dF_S(s)}_{>0}\mu(d\theta) > 0
\end{aligned}$$

where the second equality follows from integration by parts. The third equality follows by changing the order of integration after noticing that $v'(\theta) = 0$ for all $\theta \notin [\underline{\theta}^*, \bar{\theta}^*]$. The inequality follows by the construction of the subset $[\underline{\theta}^*, \bar{\theta}^*]$ and the absolute continuity of μ. Therefore, $a(\rho'')$ is not more responsive with a lower mean than $a(\rho')$.

Notice that

$$E[a(\rho'')] = E\Big[E[v(\theta)|s;\rho'']\Big] = E[v(\theta)] = E\Big[E[v(\theta)|s;\rho']\Big] = E[a(\rho')].$$

Thus,

$$\int_{t^*}^{1} \Big(a(s;p_2) - a(s;\rho')\Big) dF_S(s)$$

$$= \underbrace{\int_0^1 \Big(a(s;p_2) - a(s;\rho')\Big) dF_S(s)}_{=0} - \left(\underbrace{\int_0^{t^*} \Big(a(s;p_2) - a(s;\rho')\Big) dF_S(s)}_{>0} \right) < 0$$

and thus, $a(\rho'')$ is not more responsive with a higher mean than $a(\rho')$. ■

6.2 Proofs from Section 3

Proof of Theorem 2

Proof. 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 Bayesian game $\mathcal{G}_\rho$. 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 $\mathcal{A} = \mathcal{A}_1 \times \mathcal{A}_2$. Given opponent strategies $\alpha_{-i} \in \mathcal{A}_{-i}$, let $a_i^{BR}(\alpha_{-i}, \rho) : S_i \to A_i$ be player i's best response strategy. Specifically,

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

By (A.6), (A.10), and (A.11)-(A.13), $a_i^{BR}(\alpha_{-i}, \rho) \in \mathcal{A}_i$ for $i = 1, 2$.[19]

For any given arbitrary monotone strategies $\alpha \in \mathcal{A}$, denote the profile of best-response strategies by $a^{BR}(\alpha, \rho) \triangleq \{a_i^{BR}(\alpha_{-i}, \rho)\}_{i=1,2}$. Then, a BNE of $\mathcal{G}_\rho$, $a^\star(\rho)$, is given by the fixed point $a^{BR}(a^\star(\rho), \rho) = a^\star(\rho)$.

The proof to Theorem 2 proceeds in four steps:

1. Player i's best response strategy increases in responsiveness when player i's information quality increases (Lemma 3)

2. Player i's best response strategy increases in responsiveness when player $-i$ information quality increases (Lemma 4)

[19] By the monotonicity of the best response, a_i^{BR} is equivalent to the quantile function almost everywhere. We can then apply Lemma 1 and directly use a_i^{BR} to characterize responsiveness.

3. Player i's best response strategy increases in responsiveness when player $-i$'s strategy increases in responsiveness (Lemma 5)

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

We only prove the case for $u^i \in \mathcal{P}^{\uparrow}$. A symmetric argument establishes the result for the case of $u^i \in \mathcal{P}^{\downarrow}$.

Lemma 3 *Fix some arbitrary strategy $\alpha_{-i} \in \mathcal{A}_{-i}$. Consider two information structures $(\Sigma_{\rho'_i}, \Sigma_{\rho_{-i}})$ and $(\Sigma_{\rho''_i}, \Sigma_{\rho_{-i}})$ with $\rho''_i \succeq_{MIO} \rho'_i$. If $u^i \in \mathcal{P}^{\uparrow}$, then $a_i^{BR}(\alpha_{-i}, \rho''_i, \rho_{-i})$ is more responsive with a higher mean than $a_i^{BR}(\alpha_{-i}, \rho'_i, \rho_{-i})$.*

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

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

and notice that

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

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 $a_i^{BR}(\alpha_{-i}, \rho''_i, \rho_{-i})$ is more responsive with a higher mean than $a_i^{BR}(\alpha_{-i}, \rho'_i, \rho_{-i})$ by Theorem 1.

First, notice that $\tilde{u}^i(\theta_i, a_i)$ inherits the measurability, boundedness, and smoothness properties of u^i. In particular $u^i_{a_i,a_i}(\theta_i, a_{-i}, a_i) < 0$ for all $(\theta_i, a_{-i}) \in \Theta_i \times A_{-i}$ implies that $\tilde{u}^i_{a_i,a_i}(\theta_i, a_i) < 0$ for all $\theta_i \in \Theta_i$. Similarly, $u^i_{a_i}(\theta_i, a_{-i}, a_i)$ is convex in a_i for all $(\theta_i, a_{-i}) \in \Theta_i \times A_{-i}$ implies that $\tilde{u}^i_{a_i}(\theta_i, a_i)$ is convex in a_i for all $\theta_i \in \Theta_i$.

To see supermodularity of $\tilde{u}^i$, let $\theta'_i > \theta_i$. Then,

$$\tilde{u}^i_{a_i}(\theta'_i, a_i) - \tilde{u}^i_{a_i}(\theta_i, a_i)$$

$$= \int_{\Theta_{-i}\times S_{-i}} u^i_{a_i}\Big(\theta'_i, \alpha_{-i}(s_{-i}), a_i\Big) F(ds_{-i}|\theta_{-i}; \rho_{-i})\mu^o(d\theta_{-i}|\theta'_i)$$

$$- \int_{\Theta_{-i}\times S_{-i}} u^i_{a_i}\Big(\theta_i, \alpha_{-i}(s_{-i}), a_i\Big) F(ds_{-i}|\theta_{-i}; \rho_{-i})\mu^o(d\theta_{-i}|\theta_i)$$

$$= \int_{\Theta_{-i}\times S_{-i}} \left(u^i_{a_i}\Big(\theta'_i, \alpha_{-i}(s_{-i}), a_i\Big) - u^i_{a_i}\Big(\theta_i, \alpha_{-i}(s_{-i}), a_i\Big)\right) F(ds_{-i}|\theta_{-i}; \rho_{-i})\mu^o(d\theta_{-i}|\theta'_i)$$

$$+ \int_{\Theta_{-i}\times S_{-i}} u^i_{a_i}\Big(\theta_i, \alpha_{-i}(s_{-i}), a_i\Big) F(ds_{-i}|\theta_{-i}; \rho_{-i})\left(\mu^o(d\theta_{-i}|\theta'_i) - \mu^o(d\theta_{-i}|\theta_i)\right)$$

Since $u^i(\theta_i, a_{-i}, a_i)$ is supermodular in (θ_i, a_i) for each $a_{-i} \in A_{-i}$ and since supermodularity is preserved under integration, the first term

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

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

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

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

Lemma 4 *Fix some arbitrary strategy $\alpha_{-i} \in \mathcal{A}_{-i}$. Consider two information structures $(\Sigma_{\rho_i}, \Sigma_{\rho'_{-i}})$ and $(\Sigma_{\rho_i}, \Sigma_{\rho''_{-i}})$ with $\rho''_{-i} \succeq_{MIO} \rho'_{-i}$. If $u^i \in \mathcal{P}^\uparrow$, then $a_i^{BR}(\alpha_{-i}, \rho_i, \rho''_{-i})$ is more responsive with a higher mean than $a_i^{BR}(\alpha_{-i}, \rho_i, \rho'_{-i})$.*

Proof. Following the same first order condition argument we used in Theorem 1.i, we get the expression

$$\Big(a_i^{BR}(s_i;\alpha_{-i},\rho') - a_i^{BR}(s_i;\alpha_{-i},\rho'')\Big)\underbrace{\int_{\Theta_i\times S_{-i}} -u^i_{a_ia_i}\big(\theta_i,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\alpha_{-i},\rho')\big)\boldsymbol{F}(d\theta_i,ds_{-i}|s_i;\rho'')}_{\triangleq \hat{B}(s_i)}$$

$$+\int_{\Theta_i\times S_{-i}} u^i_{a_i}\big(\theta_i,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\alpha_{-i},\rho')\big)\Big(\boldsymbol{F}(d\theta_i,ds_{-i}|s_i;\rho'') - \boldsymbol{F}(d\theta_i,ds_{-i}|s_i;\rho')\Big) \le 0.$$

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

$$\begin{aligned}
&\int_t^1 \Big(a_i^{BR}(s_i;\alpha_{-i},\rho') - a_i^{BR}(s_i;\alpha_{-i},\rho'')\Big)ds_i\\
&\le \int_t^1 \hat{B}(s_i)^{-1}\int_{\Theta_i\times S_{-i}} u^i_{a_i}\big(\theta_i,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\alpha_{-i},\rho')\big)\Big(\boldsymbol{F}(d\theta_i,ds_{-i}|s_i;\rho') - \boldsymbol{F}(d\theta_i,ds_{-i}|s_i;\rho'')\Big)ds_i\\
&= \int_{\Theta_i\times S} u^i_{a_i}\big(\theta_i,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\alpha_{-i},\rho')\big)\hat{B}(s_i)^{-1}\mathbb{1}_{[s_i\ge t]}\Big(\boldsymbol{F}(d\theta_i,ds;\rho') - \boldsymbol{F}(d\theta_i,ds;\rho'')\Big).
\end{aligned}$$

For a given information structure Σ_ρ,

$$\begin{aligned}
\boldsymbol{F}(d\theta_i,ds;\rho) =& \boldsymbol{F}(d\theta_i,ds_i|s_{-i};\rho)dF_{S_{-i}}(s_{-i})\\
=& \int_{\Theta_{-i}} \boldsymbol{F}(d\theta_i,ds_i|\theta_{-i},s_{-i};\rho)F(d\theta_{-i},ds_{-i};\rho_{-i})\\
=& \int_{\Theta_{-i}} \boldsymbol{F}(d\theta_i,ds_i|\theta_{-i};\rho)F(d\theta_{-i},ds_{-i};\rho_{-i})\\
=& \int_{\Theta_{-i}} F(ds_i|\theta_i;\rho_i)\mu^o(d\theta_i|\theta_{-i})F(d\theta_{-i},ds_{-i};\rho_{-i})
\end{aligned}$$

where the last two equalities follows from (A.11). Let

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

so that

$$\int_t^1 \Big(a_i^{BR}(s_i;\alpha_{-i},\rho') - a_i^{BR}(s_i;\alpha_{-i},\rho'')\Big)ds$$

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

Take $s'_{-i} > s_{-i}$. Then, $\hat{\psi}(\theta_{-i},s'_{-i};t) - \hat{\psi}(\theta_{-i},s_{-i};t)$ is increasing in θ_{-i} because

1. $u^i_{a_i}$ has increasing differences in (θ_i, a_i), (θ_i, a_{-i}) and (a_{-i}, a_i),
2. $\mathbb{1}_{[s_i \geq t]}$ is increasing in s_i,
3. $\boldsymbol{F}(d\theta_i, ds_i|d\theta_{-i};\rho)$ is increasing in FOSD as θ_{-i} increases, and
4. $\boldsymbol{F}(d\theta_i, ds_{-i}|ds_i)$ is increasing in FOSD as s_i increases.

Thus, $\hat{\psi}(\theta_{-i},s_{-i};t)$ is supermodular in (θ_{-i},s_{-i}). By Lemma 2, $\rho''_{-i} \succeq_{MIO} \rho'_{-i}$ implies

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

giving us the desired result. ∎

Lemma 5 *Fix Σ_ρ. Let $\alpha''_{-i}, \alpha'_{-i} \in \mathcal{A}_{-i}$ such that α''_{-i} is more responsive with higher mean than α'_{-i}. If $u^i \in \mathcal{P}^{\uparrow}$, then, $a_i^{BR}(\alpha''_{-i},\rho)$ is more responsive with a higher mean than $a_i^{BR}(\alpha'_{-i},\rho)$.*

Proof. Suppress the dependence on ρ as it is held fixed. For any $t \in [0,1]$, we use the first order conditions argument (similar to the proof of Lemma 4) to get the expression

$$\int_t^1 \big(a_i^{BR}(s_i;\alpha'_{-i}) - a_i^{BR}(s_i;\alpha''_{-i})\big)ds_i$$

$$\leq \int_t^1 \Bigg\{ \Bigg(\underbrace{-\int_{\Theta_i\times S_{-i}} u^i_{a_i a_i}\Big(\theta, \alpha''_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)\boldsymbol{F}(d\theta_i, ds_{-i}|s_i)}_{\triangleq \tilde{B}_i(s_i)}\Bigg)^{-1}$$

$$\times \int_{\Theta_i\times S_{-i}} \Bigg[u^i_{a_i}\Big(\theta_i, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big) - u^i_{a_i}\Big(\theta_i, \alpha''_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)\Bigg]\boldsymbol{F}(d\theta_i, ds_{-i}|s_i)\Bigg\}ds_i.$$

Since $u^i_{a_i}$ is continuous and increasing in a_{-i} (by supermodularity of u^i in (a_i, a_{-i})), it is differentiable in a_{-i} almost everywhere. By convexity of $u^i_{a_i}$ in a_{-i},

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

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

Thus,

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

We make use of the following result from Quah and Strulovici (2009)

Lemma 6 *Let $g : [x', x''] \to \mathbb{R}$ and $h : [x', x''] \to \mathbb{R}$ be integrable functions. If g is increasing and $\int_x^{x''} h(t)dt \geq 0$ for all $x \in [x', x'']$, then $\int_{x'}^{x''} g(t)h(t)dt \geq g(x') \int_{x'}^{x''} h(t)dt$.*

Proof. Quah and Strulovici (2009) Lemma 1 ■

By using the definition of responsiveness in Lemma 1 and the equivalence of the monotone strategy α_{-i} with its quantile function, α''_{-i} is more responsive with a higher mean than α'_{-i} if, and only if,

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

Furthermore,

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

is an increasing function of s_{-i} because

1. $u^i_{a_i}$ has increasing differences in (θ_i, a_i), (θ_i, a_{-i}) and (a_{-i}, a_i),
2. $u^i_{a_i}$ is convex in a_{-i} for all (θ_i, a_i),
3. $\mathbb{1}_{[s_i \geq t]}$ is increasing in s_i,
4. $\boldsymbol{F}(d\theta_i, ds_i | ds_{-i})$ is increasing in FOSD as s_{-i} increases, and

5. $\boldsymbol{F}(d\theta_i, ds_{-i}|ds_i)$ is increasing in FOSD as s_i increases.

Applying Lemma 6, we have

$$\int_t^1 \big(a_i^{BR}(s_i;\alpha'_{-i}) - a_i^{BR}(s_i;\alpha''_{-i})\big)ds_i$$

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

$$\leq \int_{S_{-i}} \big(\alpha'_{-i}(s_{-i}) - \alpha''_{-i}(s_{-i})\big)ds_{-i} \int_{\Theta_i\times S_i} u^i_{a_i a_{-i}}\Big(\theta_i, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)\tilde{B}(s_i)^{-1}\mathbb{1}_{[s_i\geq t]}\boldsymbol{F}(d\theta_i, ds_i|0)$$

$$\leq 0$$

for each $t \in [0,1]$ where the last inequality follows because

$$\int_{S_{-i}} \big(\alpha'_{-i}(s_{-i}) - \alpha''_{-i}(s_{-i})\big)ds_{-i} \leq 0$$

by responsiveness with a higher mean and because

$$u^i_{a_i a_{-i}}\Big(\theta_i, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)\tilde{B}(s)^{-1}\mathbb{1}_{[s_i\geq t]} \geq 0$$

by the supermodularity of u^i in (a_i, a_{-i}) and the concavity of u^i in a_i. ■ We will 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 2 (Contractible Space) *Let X be a topological space. Let $\Phi : X \to X$ be the identity map with $\Phi(x) = x, \forall x \in X$. We say that X is a contractible space if there exists a map $\Gamma : X \times [0,1] \to X$, and a function $\Psi : X \to X$ such that, for all $x \in X$,*

1. *$\Gamma(x, \lambda)$ is continuous in λ*
2. *$\Gamma(x, 0) = \Phi(x)$ and $\Gamma(x, 1) = \Psi(x)$*
3. *$\Psi(x)$ is a constant function*

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

Theorem 6, Villas-Boas (1997): Consider a subset of a Banach space X, 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 subspace. Let both T_1 and T_2 have a fixed point on X. Let $x' \succeq x \Rightarrow T_1(x') \succeq T_1(x)$, and let $T_1(x) \succeq T_2(x)$ 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. For each player $i \in N$, any $\alpha_i \in \mathcal{A}_i$ is both uniformly bounded and is of bounded variation as it is a monotone function from $[0, 1]$ to $[\underline{a}_i, \bar{a}_i] = A_i$. Therefore, $\mathcal{A}_i$ is a subset of $BV([0, 1], A_i)$, where $BV([0, 1], A_i)$ is the space of functions of bounded variation from $[0, 1]$ to A_i.

Consider any sequence $\{\tilde{\alpha}_{i,k}\}_{k=1}^{\infty} \in \mathcal{A}_i$ such that $\tilde{\alpha}_{i,k} \to \tilde{\alpha}_i$ point-wise. The point-wise limit of monotone functions is monotone. Similarly, the point-wise limit of measurable functions is measurable (Corollary 8.9, Measure, Integrals, and Martingales, Schilling, 2005). So, $\tilde{\alpha}_i \in \mathcal{A}_i$, which implies that $\mathcal{A}_i$ is a closed subset of $BV([0, 1], A_i)$. Therefore, $\mathcal{A}_i$ equipped with the sup-norm metric $|| \cdot ||_\infty$, is a Banach space.[20]

Define a partial order over $\mathcal{A}_i$ by $\alpha_i' \succeq_i \alpha_i$ if, and only if, α_i' is more responsive with a higher mean than α_i.

Lemma 7 *For all $\alpha_i \in \mathcal{A}_i$, the upper-set $\{\alpha_i' \in \mathcal{A}_i : \alpha_i' \succeq_i \alpha_i\}$ is a compact and contractible set.*

Proof. For some $\alpha_i \in \mathcal{A}_i$, the upper-set, $U(\alpha_i) = \{\alpha_i' \in \mathcal{A}_i : \alpha_i' \succeq_i \alpha_i\}$, is a closed subset of $\mathcal{A}_i$ (follows from the Dominated Convergence Theorem). Consider a sequence $\{\tilde{\alpha}_{i,k}\}_{k=1}^{\infty} \in U(\alpha_i)$. Using, Helly's Selection Theorem, there exists a subsequence $\{\tilde{\alpha}_{i,k_m}\}_{k_m} \in U(\alpha_i)$ that converges point-wise to a limit $\tilde{\alpha}_i \in \mathcal{A}_i$. As $U(\alpha_i)$ is closed, $\tilde{\alpha}_i \in U(\alpha_i)$. Therefore, $U(\alpha_i)$ is (sequentially) compact.

Next we show that $U(\alpha_i)$ 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 $s \in [0, 1]$. We have $\alpha_i^c \in \mathcal{A}_i$. For all $\alpha_i' \in \mathcal{A}_i$, $\alpha_i^c(s) \geq \alpha_i'(s)$ for all $s \in [0, 1]$, and thus $\alpha_i^c \succeq_i \alpha_i' \Rightarrow \alpha_i^c \in U(\alpha_i)$.

Define the mappings $\Psi : \mathcal{A}_i \to \mathcal{A}_i$ and $\Phi : \mathcal{A}_i \to \mathcal{A}_i$ such that $\Psi(\alpha_i) = \alpha_i^c$ and $\Phi(\alpha_i) = \alpha_i$ for all $\alpha_i \in \mathcal{A}_i$. For each $\alpha_i \in \mathcal{A}_i$, defne the mapping $\Gamma : U(\alpha_i) \times [0, 1] \to U(\alpha_i)$ such that

[20] For $\alpha_i, \alpha_i' \in \mathcal{A}_i$, $||\alpha_i - \alpha_i'||_\infty = \sup_{s \in [0,1]} |\alpha_i(s) - \alpha_i'(s)|$

$\Gamma(\alpha_i', \lambda) = (1-\lambda)\Phi(\alpha_i') + \lambda\Psi(\alpha_i')$. $\Gamma(\alpha_i', \lambda)$ is continuous in λ. As λ increases on $[0,1]$, Γ maps any point in $\mathcal{U}(\alpha_i)$ to α_i^c, which is itself in $\mathcal{U}(\alpha_i)$. Therefore, $\mathcal{U}(\alpha_i)$ is contractible. ■

We therefore have an order on a Banach space $\mathcal{A}_i$ that generates compact and contractible upper-sets. We extend these properties to $\mathcal{A} = \times_{i=1}^{I} \mathcal{A}_i$ by the product order: given $\alpha, \alpha' \in \mathcal{A}$, $\alpha' \succeq \alpha$ if, and only if, $\alpha_i' \succeq_i \alpha_i$ for all $i \in N$. Along with the product topology, $\succeq$ is an order on a Banach space $\mathcal{A}$ that generates compact and contractible upper-sets.[21]

Consider a Bayesian game $\mathcal{G}_\rho = (\Sigma_\rho, G)$. Define an operator $T_\rho : \mathcal{A} \to \mathcal{A}$ with

$$T_\rho(\alpha) = \big(a_1^{BR}(\alpha_{-1}, \rho), a_2^{BR}(\alpha_{-2}, \rho), \ldots, a_n^{BR}(\alpha_{-n}, \rho)\big).$$

T_ρ is continuous in α as utility functions are continuous in opponent's actions. A monotone BNE of $\mathcal{G}_\rho$ is a fixed point of T_ρ. We know such a fixed point exists (Van Zandt and Vives (2007).

Now consider two different games, $\mathcal{G}_{\rho''} = (\Sigma_{\rho''}, G)$ and $\mathcal{G}_{\rho'} = (\Sigma_{\rho'}, G)$, with $\rho'' \succeq_{MIO} \rho'$. From Lemma 5,

$$\alpha' \succeq \alpha \Rightarrow \alpha_i' \succeq_i \alpha_i, \forall i \in N \Rightarrow a_i^{BR}(\alpha_{-i}', \rho'') \succeq_i a_i^{BR}(\alpha_{-i}, \rho''), \forall i \in N \Rightarrow T_{\rho''}(\alpha') \succeq T_{\rho''}(\alpha).$$

From Lemma 3 and 4,

$$\rho'' \succeq_{MIO} \rho' \Rightarrow a_i^{BR}(\alpha_{-i}, \rho'') \succeq_i a_i^{BR}(\alpha_{-i}, \rho'), \forall i \in N \Rightarrow T_{\rho''}(\alpha) \succeq T_{\rho'}(\alpha)$$

for all $\alpha \in \mathcal{A}$. We can now directly apply Theorem 6: Villas-Boas 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 $a^\star(\rho'') \succeq a^\star(\rho')$. ■

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