Bayesian Comparative Statics*

Teddy Mekonnen[†] René Leal Vizcaíno[‡]

April 14, 2021

Abstract

We study how changes to the informativeness of signals in Bayesian games and single-agent decision problems affect the distribution of equilibrium actions. Focusing on supermodular environments, we provide conditions under which a more precise private signal for one agent leads to an increasing-mean spread or a decreasing-mean spread of equilibrium actions for all agents. We apply our comparative statics to information disclosure games between a sender and many receivers and derive sufficient conditions on the primitive payoffs that lead to extremal disclosure of information.

JEL Codes: C44, C61, D42, D81

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

[†]Brown University, Department of Economics. mekonnen@brown.edu

[‡]Banco de México. rlealv@banxico.org.mx

1. Introduction

Economists have long been interested in how equilibrium actions and welfare in Bayesian games vary with the quality of information. For example, what happens to the equilibrium distribution of prices under Cournot or Bertrand competition when firms observe a more precise signal about their cost parameters? Would more precise signals lead to more efficient outcomes? Such questions have generally been addressed in linear-quadratic settings (games with quadratic utility functions and normally distributed states and signals) in which it is well-known that a decrease in the noise of a player's private signal leads to a mean-preserving spread in the distribution of actions.

Beyond linear-quadratic settings however, characterizing the effect of information on the distribution of equilibrium actions has proven difficult. The standard monotone comparative static tools (Milgrom and Shannon, 1994) are not applicable because most stochastic orders, such as mean-preserving spreads or second-order stochastic dominance, over joint probability distributions of actions and states lack a lattice structure (Müller and Scarsini, 2006). Furthermore, Roux and Sobel (2015) show that comparative statics results about the dispersion of equilibrium actions cannot always be established even when players use monotone equilibrium strategies.

In this paper, we identify two classes of Bayesian games (and decision problems) for which a higher quality of private information unambiguously leads to more dispersed distribution of equilibrium actions along with monotone shifts to the mean. In particular, we consider Bayesian settings in which players have supermodular utility functions with either (i) supermodular and convex marginal utility, or (ii) submodular and concave marginal utility. Games with quadratic utility functions lie at the intersection of these two classes. Supermodular payoffs imply that there are complementarities between a player's action and the state as well as strategic complementarities between any two players. Supermodular and convex (resp., submodular and concave) marginal utilities further imply that these complementarities are getting stronger (resp., weaker) as a player's action increases.

To formally describe the comparative static results, we must first discuss an order over information structures that captures quality, and an order over distributions of actions that captures changes in the mean and dispersion.

To compare the quality of information, we first restrict attention to monotone information structures, i.e., higher signal realizations lead to first-order stochastic shifts in posterior beliefs. We then use the *supermodular stochastic order* (Tchen, 1980; Meyer and Strulovici, 2012), a more complete order than the Blackwell order (Blackwell, 1951, 1953) and the Lehmann order (Lehmann, 1988) within the class of monotone information structures. Loosely, an information structure ρ dominates another information structure ρ' in the supermodular stochastic order if ρ exhibits more interdependence between signals and states than ρ'.

Each information structure ρ induces a distribution $H_i(\rho)$ over player i's actions in equilibrium. Given two information structures ρ and ρ', we say the players are more *responsive with a higher mean* under ρ than ρ' if, for each player i, $H_i(\rho)$ dominates $H_i(\rho')$ in the increasing convex order. Loosely speaking, players become more responsive with a higher mean when each player's distribution over actions undergoes an "increasing-mean spread." Alternatively, we say that players are more *responsive with a lower mean* under ρ than ρ' if, for each player i, $H_i(\rho')$ second-order stochastically dominates $H_i(\rho)$.

Our main result shows that for a subclass of supermodular games, players are more responsive if, and only if, the quality of information increases in the supermodular stochastic order. In particular, we show that for supermodular games with supermodular and convex marginal utilities, players are more responsive with a higher mean under ρ than under ρ' whenever ρ dominates ρ' in the supermodular stochastic order. Furthermore, if players are more responsive with a higher mean under ρ than ρ' for all supermodular games with supermodular and convex marginal utilities, then ρ necessarily dominates ρ' in the supermodular stochastic order. We also present symmetric results linking responsiveness with a lower mean to supermodular games with submodular and concave marginal utilities.

As an application of our main result, we consider a Bayesian persuasion game between a sender and possibly many receivers. The literature commonly takes the sender's interim value function as a primitive of the persuasion framework. However, the sender's interim value function is endogenous and difficult to derive without assuming binary actions and states or assuming that the receivers' equilibrium strategies depend only on the first and second moments of their posterior beliefs. For example, these assumptions are used in Rayo and Segal (2010), Bergemann and Morris (2013), Gentzkow and Kamenica (2016), Taneva (2017), and Dworczak and Martini (2019). Instead, we restrict attention to the two classes of games we consider, and provide conditions for when there is minimal and maximal levels of conflict between the sender and the receivers based solely on their utility functions. Under our conditions, we show that the extremal disclosure of information is optimal.

Related Literature

The closest paper to ours is Jensen (2018) which also studies how the distribution of individual decisions and equilibrium outcomes vary with changes in some economic parameters. While Jensen's methodology would prove useful to answer the comparative statics we are interested in, it requires imposing a *quasiconvex differences* condition on the players' interim utility function. Since quasicovexity is not preserved under integration, we are left with the open question of what conditions on the primitives lead to interim utility functions with quasiconvex differences. We show that the class of games we consider do in fact lead to interim utility functions that satisfy quasiconvex differences with the added benefit that all our assumptions are only on the primitives. We provide a more detailed discussion in Section 3.1 following our main results.

Methodologically, this paper contributes to the literature on monotone comparative statics. Specifically focusing on Bayesian single-agent decision problems, Athey (2002) and Quah and Strulovici (2009) show that optimal actions are a monotone function of beliefs (for beliefs ordered by stochastic dominance). Similarly, in Bayesian games, Athey (2001) and Van Zandt and Vives (2007) show that a player's Bayesian Nash equilibrium action is a monotone function of the player's beliefs. We add to this literature by showing that the *distribution* of equilibrium actions is monotone (in the increasing/decreasing convex order) as a function of the *distribution* of beliefs (for distributions of beliefs ordered consistently with the supermodular stochastic order).

The supermodular stochastic order has been previously studied by Athey and Levin (2017) who show that for all single-agent decision makers with supermodular preferences, the value of information increases if, and only if, the quality of information increases in the supermodular stochastic order. Our main result for the single-agent case shows that for a subset of supermodular preferences, optimal actions become more dispersed if, and only if, the quality of information increases in the supermodular stochastic order. Amir and Lazzati (2016) extend Athey and Levin's result to supermodular Bayesian games and show that the value of information is increasing and convex in the supermodular stochastic order. Additionally, Amir and Lazzati show that the gap between a player's highest and lowest equilibrium actions increases as information quality increases in the supermodular stochastic order. Our main result shows that such "dispersion" in the players' equilibrium behavior extends beyond the gap between the highest and lowest actions; it holds for the entire distribution of equilibrium actions.

The remainder of the paper is structured as follows: Section 2 presents our model and introduces the orders over distributions of actions and information structures. Our main result is then presented in Section 3, followed by examples and an application in Section 4. All proofs are in the Appendix.

2. Model

2.1. Preliminary Definitions and Notation

Let X_i be a compact subset of $\mathbb{R}$ for $i = 1, \ldots, m$, and define $X = \times_{i=1}^m X_i$ and $X_{-i} = \times_{j \neq i} X_j$. We equip X with the coordinate-wise order $\geq$, i.e., for $x'', x' \in X$, $x'' \geq x'$ if $x''_i \geq x'_i$ for all $i = 1, 2, \ldots, m$. We also equip X_{-i} with the same coordinate-wise order.

We say a function $g : X \to \mathbb{R}$ has increasing (resp., decreasing, or constant) differences in $(x_{-i}; x_i)$ if $g(x_i, x''_{-i}) - g(x_i, x'_{-i})$ is increasing (resp., decreasing, or constant) in x_i for all $x''_{-i}, x'_{-i} \in X_{-i}$ with $x''_{-i} \geq x'_{-i}$.

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

All references to "increasing" or "decreasing," "increasing differences" or "decreasing differences," and "convex" or "concave" are in the weak sense.

2.2. Basic Game Setup

There are n players with $N = \{1, 2, \ldots, n\}$ denoting the set of players. While our exposition highlights games with $n > 1$, we emphasize that our setup and results also apply to single-agent decision problems with $n = 1$.

Each player $i \in N$ has a random state variable (or type) $\tilde{\theta}_i$ with support contained in $\Theta_i = [\underline{\theta}_i, \bar{\theta}_i]$. Define $\Theta = \times_{i \in N} \Theta_i$ and $\Theta_{-i} = \times_{j \neq i} \Theta_j$. To distinguish random variables from their realizations, we denote the random state variables by $\tilde{\theta} = (\tilde{\theta}_i, \tilde{\theta}_{-i})$ and the realized states by $\theta = (\theta_i, \theta_{-i})$.

The players hold a common prior given by the joint cumulative distribution function $F : \Theta \to [0, 1]$. Let F_{Θ_i} be the marginal distribution of $\tilde{\theta}_i$ and let $F_{\Theta_{-i}}(\cdot|\theta_i)$ be the joint distribution of $\tilde{\theta}_{-i}$ conditional on $\tilde{\theta}_i = \theta_i$ induced by F.

Assumption 1 *For all $i \in N$, $F_{\Theta_{-i}}(\cdot|\theta_i)$ first-order stochastically dominates $F_{\Theta_{-i}}(\cdot|\theta'_i)$ whenever $\theta_i > \theta'_i$. We adopt the notation $F_{\Theta_{-i}}(\cdot|\theta_i) \succeq_{FOSD} F_{\Theta_{-i}}(\cdot|\theta'_i)$.*

Let $A_i = [\underline{a}_i, \bar{a}_i]$ be the player i's action space, $A = \times_{i \in N} A_i$, and $A_{-i} = \times_{j \neq i} A_j$. Each player $i \in N$ has a utility function given by $u^i : \Theta \times A \to \mathbb{R}$.

Assumption 2 *For each $i \in N$,*

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

b. for all $(\theta, a_{-i}) \in \Theta \times A_{-i}$, $u^i(\theta, a_{-i}, \cdot)$ is strictly concave in a_i, *c. for all* $(\theta, a_{-i}) \in \Theta \times A_{-i}$*, there exists an action* $a_i \in A_i$ *such that* $u^i_{a_i}(\theta, a_{-i}, a_i) = 0$*, and*

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

Assumption 2.a-c imply that players have unique best responses that are characterized by first-order conditions. Assumption 2.d implies that there are complementarities between the state of the world and a player's action. Additionally, there are strategic complementarities between the players' actions. Thus, player i wants to take a high action when the state θ is high or when player j takes a high action.

2.3. Information Structures

Each player $i \in N$ observes a signal $\tilde{s}_i$ from an information structure $\Sigma_{\rho_i} = \langle S_i, G(\cdot, \cdot; \rho_i)\rangle$ where $S_i \subseteq \mathbb{R}$ is the signal space, $G(\cdot, \cdot; \rho_i) : \Theta_i \times S_i \to [0, 1]$ is a joint cumulative distribution of $(\tilde{\theta}_i, \tilde{s}_i)$, and ρ_i is an index. Let $G_{\Theta_i}(\cdot; \rho_i)$ and $G_{S_i}(\cdot; \rho_i)$ be the marginal distribution of $\tilde{\theta}_i$ and $\tilde{s}_i$ respectively. Let $G_{\Theta_i}(\cdot|s_i; \rho_i)$ be player i's posterior distribution conditional on $\tilde{s}_i = s_i$, and let $G_{S_i}(\cdot|\theta_i; \rho_i)$ be the distribution of signals conditional on $\tilde{\theta}_i = \theta_i$.

Assumption 3 *For all* $i \in N$*,*

a. $G_{\Theta_i}(\cdot; \rho_i) = F_{\Theta_i}(\cdot)$*,*

b. $G_{S_i}(\cdot; \rho_i) = G_{S_i}(\cdot)$*,*

c. $G_{\Theta_i}(\cdot|s_i; \rho_i) \succeq_{FOSD} G_{\Theta_i}(\cdot|s_i'; \rho_i)$ *whenever* $s_i > s_i'$*, and*

d. $G_{S_i}(\cdot|\theta_i; \rho_i) \succeq_{FOSD} G_{S_i}(\cdot|\theta_i'; \rho_i)$ *whenever* $\theta_i > \theta_i'$*.*

Assumption 3.a implies that posterior beliefs satisfy Bayes plausibility (Kamenica and Gentzkow, 2011). Assumption 3.b, which holds without loss of generality, states that all information structures induce the same marginal distribution on $\tilde{s}_i$.[1] Assumption 3.c implies that higher states are more likely when the signal realization is high while Assumption 3.d implies that higher signal realizations are more likely when the state is high.

Let $S = \times_{i \in N} S_i$. We denote the profile of information structures by $\Sigma_\rho = (\Sigma_{\rho_1}, \ldots, \Sigma_{\rho_n})$. A profile Σ_ρ induces a joint distribution $\boldsymbol{G}(\cdot, \cdot; \rho) : \Theta \times S \to [0, 1]$ over $(\tilde{\theta}, \tilde{s})$. We assume that player i cannot directly learn about $(\tilde{\theta}_{-i}, \tilde{s}_{-i})$. Formally,

Assumption 4 $\boldsymbol{G}(s|\theta; \rho) = \prod_{i \in N} G_{S_i}(s_i|\theta_i; \rho_i)$ *for all* $(\theta, s) \in \Theta \times S$*.*

[1] The assumption is without loss of generality because we can apply the integral probability transform to any random signal $\tilde{s}_i$ and create a new, equally informative signal which is uniformly distributed on the unit interval. See Lehmann (1988).

2.4. Equilibrium Outcomes

Following the terminology of Gossner (2000), we decompose a Bayesian game into a basic game and a profile of information structures. The basic game $\Gamma = \langle N, \{A_i, u^i\}_{i\in N}, F\rangle$ is comprised of (i) a set of players N, (ii) for each player $i \in N$, an action space A_i along with a payoff function $u^i : \Theta \times A \to \mathbb{R}$ that satisfies Assumption 2, and (iii) a common prior F that satisfies Assumption 1. The profile of information structures Σ_ρ satisfies Assumption 3 and Assumption 4. Both Γ and Σ_ρ are common knowledge. The full Bayesian game is given by $\mathcal{G}_\rho = (\Sigma_\rho, \Gamma)$. The setting is general enough to accommodate both finite and continuous types, as well as private, interdependent, or common values. For example, the IPV case is given by $\tilde{\theta}_i \perp \tilde{\theta}_j$ for all $j \neq i$ and $u^i(\theta, a) = u^i(\theta_i, a)$ for all $(\theta, a) \in \Theta \times A$. The pure common values case is given by $\tilde{\theta}_i = \tilde{\theta}_j$ for all $j \neq i$.

Each player $i \in N$ first privately observes a signal realization $s_i \in S_i$ generated from Σ_{ρ_i}. Then the players participate in the basic game Γ by simultaneously choosing an action. A pure strategy for player $i \in N$ is given by the measurable function $\alpha_i : S_i \to A_i$. Let $\alpha = (\alpha_i, \alpha_{-i})$ be a pure strategy profile. In a Bayesian game $\mathcal{G}_\rho$, player i's interim utility when taking action $a_i \in A_i$, given a signal realization s_i and a profile of opponent's strategies α_{-i}, is

$$U^i(a_i, \alpha_{-i}; s_i, \rho) = \int\limits_{\Theta\times S_{-i}} u^i\big(\theta, \alpha_{-i}(s_{-i}), a_i\big) d\boldsymbol{G}(\theta, s_{-i}|s_i; \rho).$$

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 $\mathcal{G}_\rho$. Thus, for each player $i \in N$ and each $s_i \in S_i$, $a_i^\star(s_i; \rho) = \arg\max_{a_i\in A_i} U^i(a_i, a_{-i}^\star(\rho); s_i, \rho)$.

We restrict our attention to monotone BNEs, i.e., each player's equilibrium strategy, $a_i^\star(s_i; \rho)$ is increasing in the signal realization s_i. The existence of monotone pure strategy BNE in supermodular games has long been established. In particular, the existence result of Van Zandt and Vives (2007) is noteworthy in our setting because their existence result does not require players to have atomless posterior beliefs when they participate in the basic game. While restricting attention to monotone BNEs may be with loss of generality, extremal equilibria are nonetheless monotone. Specifically, the least and the greatest BNEs of a supermodular Bayesian game are in monotone pure strategies.

Our goal in this paper is to characterize a comparative statics of $a^\star(\rho)$ as the information structure Σ_ρ changes while holding the underlying basic game Γ fixed. To do so, we will first introduce the relevant orders over actions and information structures on which our comparative statics is based.

2.5. Order over Distributions of Actions

From an interim perspective, each player $i \in N$ first observes a signal realization $s_i \in S_i$ and then takes some action $\alpha_i(s_i) \in A_i$. From an ex-ante perspective, the signal realizations are yet to be observed. Therefore, $\alpha_i(\tilde{s}_i)$ is a random variable that is distributed according to the CDF $H(\cdot; \alpha_i) : \mathbb{R} \to [0, 1]$ given by

$$H(z; \alpha_i) = \int_{S_i} \mathbf{1}_{[\alpha_i(s_i) \le z]} dG_{S_i}(s_i),$$

where $\mathbf{1}_{[\cdot]}$ is the indicator function.

Given two pure strategies α_i and α_i', we say that α_i dominates α_i' in the *increasing convex order* (resp., *decreasing convex order*) if for any measurable, convex, and increasing (resp., decreasing) function $\psi : \mathbb{R} \to \mathbb{R}$,

$$\int \psi(z) dH(z; \alpha_i) \ge \int \psi(z) dH(z; \alpha_i').$$

Loosely, α_i dominates α_i' in the increasing convex order if α_i is more dispersed and has a higher mean than α_i'. We write $\alpha_i \succeq_H \alpha_i'$ ("H" for higher mean) when α_i dominates α_i' in the increasing convex order. Similarly, α_i dominates α_i' in the decreasing convex order if α_i is more dispersed and has a lower mean than α_i'. We write $\alpha_i \succeq_L \alpha_i'$ ("L" for lower mean) when α_i dominates α_i' in the decreasing convex order. Given a profile of pure strategies $\alpha = (\alpha_1, \ldots, \alpha_n)$ and $\alpha' = (\alpha_1', \ldots, \alpha_n')$, we say that $\alpha \succeq_H \alpha'$ (resp., $\alpha \succeq_L \alpha'$) if and only if $\alpha_i \succeq_H \alpha_i'$ (resp., $\alpha_i \succeq_L \alpha_i'$) for all $i \in N$.

Definition 1 (Responsiveness)
Given a basic game Γ, we say players are ***more responsive with a higher mean*** *under Σ_ρ than $\Sigma_{\rho'}$ if and only if*

- *for each monotone BNE $a^\star(\rho')$ of $\mathcal{G}_{\rho'} = (\Sigma_{\rho'}, \Gamma)$, there exists a monotone BNE $a^\star(\rho)$ of $\mathcal{G}_\rho = (\Sigma_\rho, \Gamma)$ such that $a^\star(\rho) \succeq_H a^\star(\rho')$, and*
- *for each monotone BNE $a^\star(\rho)$ of $\mathcal{G}_\rho$, there exists a monotone BNE $a^\star(\rho')$ of $\mathcal{G}_{\rho'}$ such that $a^\star(\rho) \succeq_H a^\star(\rho')$.*

If $\succeq_H$ is replaced by $\succeq_L$, we say players are ***more responsive with a lower mean****.*

Responsiveness compares the set of BNE outcomes in the weak-set order. There is no analogous comparative statics such that *every* monotone BNE of $\mathcal{G}_\rho$ dominates *every* monotone BNE of $\mathcal{G}_{\rho'}$ in the increasing or decreasing convex order. In fact, there are no general comparative statics results for Nash equilibria (and fixed points) in the strong-set order.[2] Thus, while the definition for responsiveness takes into account the possibility of multiple BNE outcomes, we have to use equilibrium selection rules in applications. For example, if players are more responsive with a higher mean under Σ_ρ than $\Sigma_{\rho'}$, then the extremal (greatest and least) BNE outcomes of $\mathcal{G}_\rho$ dominate the respective extremal outcomes of $\mathcal{G}_{\rho'}$ in the increasing convex order. Of course, when $n = 1$, there is no multiplicity because of Assumption 2.

2.6. Order over Information Structures

The next step is to determine an appropriate way to compare different information structures. Recall that we focus, without loss of generality, on information structures that induces the same marginals on $\tilde{\theta}_i$ and $\tilde{s}_i$ (Assumption 3a-b).

Definition 2 (Supermodular Stochastic Order)
We say Σ_{ρ_i} dominates $\Sigma_{\rho_i'}$ in the ***supermodular stochastic order*** *if and only if for all $(\theta_i, s_i) \in \Theta_i \times S_i$, $G(\theta_i, s_i; \rho_i) \geq G(\theta_i, s_i; \rho_i')$.*

Intuitively, Σ_{ρ_i} dominates $\Sigma_{\rho_i'}$ in the supermodular stochastic order if the state and the signal are more correlated under Σ_{ρ_i}. Recall that Assumption 3.c implies low signal realizations are more likely for low states. When Σ_{ρ_i} dominates $\Sigma_{\rho_i'}$ in the supermodular stochastic order, a signal $\tilde{s}_i \leq s_i$ from Σ_{ρ_i} presents a stronger evidence of a low state than the same signal from $\Sigma_{\rho_i'}$, i.e.,

$$\mathbb{P}(\tilde{\theta}_i \leq \theta_i | \tilde{s}_i \leq s_i; \rho_i) \geq \mathbb{P}(\tilde{\theta}_i \leq \theta_i | \tilde{s}_i \leq s_i; \rho')$$
$$\Leftrightarrow \underbrace{\mathbb{P}(\tilde{\theta}_i \leq \theta_i | \tilde{s}_i \leq s_i; \rho_i) G_{S_i}(s_i)}_{=G(\theta_i, s_i; \rho_i)} \geq \underbrace{\mathbb{P}(\tilde{\theta}_i \leq \theta_i | \tilde{s}_i \leq s_i; \rho_i') G_{S_i}(s_i)}_{=G(\theta, s; \rho_i')}.$$

We use the notation $\rho_i \succeq_{spm} \rho_i'$ whenever Σ_{ρ_i} dominates $\Sigma_{\rho_i'}$ in the supermodular stochastic order. Given two profiles of information structures $\Sigma_\rho = (\Sigma_{\rho_1}, \ldots, \Sigma_{\rho_n})$ and $\Sigma_{\rho'} = (\Sigma_{\rho_1'}, \ldots, \Sigma_{\rho_n'})$, we write $\rho \succeq_{spm} \rho'$ if and only if $\rho_i \succeq_{spm} \rho_i'$ for all $i \in N$.

[2]Note that Vives (1990) and Milgrom and Roberts (1994) establish monotone comparative statics in strong-set order only for extremal Nash outcomes. Che et al. (2019) provides an example in which Nash equilibria before and after a policy change can be compared by the weak-set order but not the strong-set order.

3. Preferences and Main Result

The main contribution of this paper is to identify a class of games for which players become more responsive when information quality increases according to the supermodular stochastic order. We provide examples of games that fall into these classes in Section 4.

Let $\boldsymbol{\Gamma_H}$ be the class of basic games $\Gamma = \langle N, \{A_i, u^i\}_{i\in N}, F\rangle$ such that F satisfies Assumption 1 and the payoff function $u^i : \Theta \times A \to \mathbb{R}$ for each $i \in N$ satisfies Assumption 2 and has a marginal utility $u^i_{a_i}(\theta, a)$ that

(i) is convex in a_j for all $j \in N$, and

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

Below, we show $\boldsymbol{\Gamma_H}$ is linked to responsiveness with a higher mean.

Similarly, let $\boldsymbol{\Gamma_L}$ be the class of basic games $\Gamma = \langle N, \{A_i, u^i\}_{i\in N}, F\rangle$ such that F satisfies Assumption 1 and the payoff function $u^i : \Theta \times A \to \mathbb{R}$ for each $i \in N$ satisfies Assumption 2 and has a marginal utility $u^i_{a_i}(\theta, a)$ that

(i) is concave in a_j for all $j \in N$, and

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

Below, we show $\boldsymbol{\Gamma_L}$ is linked to responsiveness with a lower mean.

A basic game $\Gamma \in \boldsymbol{\Gamma_H} \cap \boldsymbol{\Gamma_L}$ has a prior F that satisfies Assumption 1 and a payoff function for each $i \in N$ that satisfies Assumption 2 with marginal utility $u^i_{a_i}(\theta, a)$ that

(i) is linear in a_j for all $j \in N$, and

(ii) has constant differences in $(\theta, a_{-j}; a_j)$ for all $j \in N$.

Below, we show that $\boldsymbol{\Gamma_H} \cap \boldsymbol{\Gamma_L}$ is linked to mean-preserving spreads of equilibrium actions.

Theorem 1 states that when the basic game Γ belongs to the class of games in $\boldsymbol{\Gamma_H}$, players are more responsive with a higher mean under $\mathcal{G}_\rho = (\Sigma_\rho, \Gamma)$ than under $\mathcal{G}_{\rho'} = (\Sigma_{\rho'}, \Gamma)$ whenever Σ_ρ dominates $\Sigma_{\rho'}$ in the supermodular stochastic order. Moreover, if Σ_ρ does NOT dominate $\Sigma_{\rho'}$ in the supermodular stochastic order, then there is a game $\Gamma \in \boldsymbol{\Gamma_H}$ in which the players are NOT more responsive with a higher mean under Σ_ρ.

The theorem also establishes a similar result relating $\boldsymbol{\Gamma_L}$ and responsiveness with a lower mean. Thus, for games in $\boldsymbol{\Gamma_H} \cap \boldsymbol{\Gamma_L}$, players are more responsive both with a higher and lower mean when information quality increases in the supermodular stochastic order; in other words, an increase in the quality of information leads to a mean-preserving spread in the distribution of equilibrium actions.

Theorem 1 *Consider any two profiles of information structures Σ_ρ and $\Sigma_{\rho'}$ that satisfy Assumption 3 and Assumption 4. Players are more responsive with a higher (resp., lower) mean under Σ_ρ than $\Sigma_{\rho'}$ for any basic game $\Gamma \in \boldsymbol{\Gamma_H}$ (resp., $\Gamma \in \boldsymbol{\Gamma_L}$) if, and only if, Σ_ρ dominates $\Sigma_{\rho'}$ in the supermodular stochastic order.*

For the case of a single-agent, we also provide a companion result in Theorem 2 relating responsiveness to the Blackwell order: the agent is more responsive under Σ_ρ than under $\Sigma_{\rho'}$ when the latter is a garbling of the former. While the result appears to be a corollary of Theorem 1, there is a difference—the garbling $\Sigma_{\rho'}$ does NOT have to be a monotone information structure, i.e., does not have to satisfy Assumption 3. The difference in these two results is made evident in Section 4 when we consider an application of Bayesian persuasion.

Theorem 2 *Let $n = 1$. Consider any information structure Σ_ρ that satisfies Assumption 3 and let $\Sigma_{\rho'}$ be any* garbling *of Σ_ρ. If $\Gamma \in \boldsymbol{\Gamma_H}$ $\big($resp., $\Gamma \in \boldsymbol{\Gamma_L}\big)$, then the agent is more responsive with a higher (resp., lower) mean under Σ_ρ than under $\Sigma_{\rho'}$.*

We defer the proof of Theorem 1 and Theorem 2 until Appendix B. Here, we provide some intuition. We start with the case of a single-agent ($n = 1$) and drop the player-index "i" for now. Since $u(\theta, a)$ has complementarities between θ and a, we know that the agent's optimal action is monotone, i.e., it increases as the agent's posterior belief increases in first-order stochastic dominance. However, monotonicity alone is not sufficient to obtain our desired comparative statics as shown in Roux and Sobel (2015); the mapping from posterior beliefs to actions would also have to be appropriately convex or concave. When Σ_ρ is Blackwell more informative than $\Sigma_{\rho'}$, the distribution over posteriors induced by Σ_ρ is a mean-preserving spread of the distribution over posteriors induced by $\Sigma_{\rho'}$. A monotone and convex (resp., monotone and concave) mapping from beliefs to actions transforms the mean-preserving spread in the distribution over posterior beliefs to a shift in the increasing convex order (resp., decreasing convex order) in the distribution over actions.

We show that if $\Gamma \in \boldsymbol{\Gamma_H}$, the mapping from (FOSD ordered) posterior beliefs to actions is indeed monotone and convex, thereby obtaining the *responsiveness with a higher mean* result in Theorem 2. For games in $\boldsymbol{\Gamma_H}$, the complementarities between actions and states are increasing as actions increase, i.e., $u_a(\theta, a)$ has increasing differences in $(\theta; a)$. Thus, for a fixed action a, the agent's marginal utility is increasing and also getting *flatter* as θ increases. Also, recall from Assumption 2 that the agent's utility is concave in her action. Additionally, for games in $\boldsymbol{\Gamma_H}$, preferences also feature convex marginal utilities, i.e., $u_a(\theta, a)$ is convex in a. Thus, for a fixed θ, the agent's marginal utility is diminishing but at an increasingly *slower rate* as actions increase. In other words, the higher the state or the higher the action, the agent's marginal value from further increasing her action diminishes at a *decelerating* rate. Thus, her incentives to take higher actions get amplified as her beliefs put more weight on higher states, leading to the desired convexity property.

For games in $\mathbf{\Gamma_L}$, the agent's preferences feature complementarities that are decreasing as actions increase. Thus, for a fixed action a, the agent's marginal utility is increasing and also getting *steeper* as θ increases. The preferences also feature concave marginal utilities, so that for a fixed θ, the agent's marginal utility is diminishing at an increasingly *faster rate* as actions increase. In other words, the higher the state or the higher the action, the agent's marginal value from further increasing her action diminishes at an *accelerating* rate. Thus, her incentives to take higher actions get dampened as her beliefs put more weight on higher states, leading to the desired concavity property.

The proof for Theorem 1 when $n = 1$ generalizes this intuition from the Blackwell order to the more general supermodular stochastic order.

Remark 1 *The convexity/concavity of the mapping from beliefs to actions is not only sufficient but also necessary for responsiveness. For example, if an agent's optimal action mapping is non-convex and non-concave, which necessarily implies $\Gamma \notin \mathbf{\Gamma_H} \cup \mathbf{\Gamma_L}$, we can find a prior and two information structures Σ_ρ and $\Sigma_{\rho'}$ such that $\rho \succeq_{spm} \rho'$ but the agent is NOT more responsive with a higher or lower mean under Σ_ρ than $\Sigma_{\rho'}$. We present such an example in Section Appendix C.*

When $n > 1$, payoffs with increasing differences in $(\theta, a_{-i}; a_i)$ imply that the players have monotone BNE strategies. However, once again, monotonicity alone is not sufficient. With multiple players, player i's relevant "state" is not just $\tilde{\theta}$ but also the random equilibrium actions of the other players $a^\star_{-i}(\tilde{s}_{-i})$. Thus, we extend the marginal utility conditions in the single-agent case to the enlarged "state space" in order to obtain the multi-player responsiveness result in Theorem 1.

It is worthwhile to highlight two effects that only exist in the case for $n > 1$. First, when the quality of information for player $j \neq i$ increases, the signals $\tilde{s}_i$ and $\tilde{s}_j$ become unconditionally more correlated. This implies that, holding player j's strategy fixed, player i can better predict player j's random action. Second, when the distribution of player j's actions shift in the increasing convex order, player i's best response also shifts in the increasing convex order due to the strategic complementarities in the payoffs and the marginal payoffs. The combination of these effects is that each player's distribution of BNE outcomes becomes more dispersed if at least one player gets a higher quality of information.

3.1. Connection to Jensen

As mentioned in the literature review, the closest paper to ours is Jensen (2018), which studies how the distribution of individual decisions and equilibrium outcomes vary with changes in the distribution of some economic parameter. As the connection is most clear in the single-agent setting, we will focus our discussion to the case when $n = 1$.

For some belief $\mu \in \Delta(\Theta)$ and action $a \in A$, define the interim payoff function as

$$U(\mu, a) = \int\limits_{\Theta} u(\theta, a)\mu(d\theta)$$

and let $a^*(\mu) = \arg\max_{a\in A} U(\mu, a)$ be the agent's optimal choice. Jensen shows that if $U(\mu, a) - U(\mu, a - \delta)$ is quasiconvex (resp., quasiconcave) for all $\delta > 0$ small enough, then $a^*(\mu)$ is a convex (resp., concave) function of μ. Consequently, the agent's optimal actions become more dispersed with a higher (resp., lower) mean as the information structure becomes Blackwell more informative.

While the quasiconvexity condition is useful to answer the questions we are interested in, it is not known what conditions on $u(\theta, a)$ would yield the quasiconvexity differences conditions on the interim utility $U(\mu, a)$. In particular, quasiconvexity of $u(\theta, a) - u(\theta, a - \delta)$ in (θ, a) does not imply quasiconvexity of $U(\mu, a) - U(\mu, a - \delta)$ in (μ, a), as quasiconvexity is not closed under integration.

Below, we show that our class of games are sufficient to establish Jensen's quasiconvexity conditions on the interim utility for posteriors that are ranked by first-order stochastic dominance. Since we are considering differentiable functions, Jensen's conditions are equivalent to quasiconvexity/quasiconcavity of $U_a(\mu, a)$.

Proposition 1 *Let $n = 1$. For any two beliefs $\mu_1, \mu_2 \in \Delta(\Theta)$ with $\mu_2 \succeq_{FOSD} \mu_1$, any $a_1, a_2 \in A$, and any $\lambda \in [0, 1]$, $\Gamma \in \boldsymbol{\Gamma_H}$ implies*

$$U_a\big(\lambda\mu_1 + (1-\lambda)\mu_2, \lambda a_1 + (1-\lambda)a_2\big) \leq \max\big\{U_a(\mu_1, a_1), U_a(\mu_2, a_2)\big\}$$

and $\Gamma \in \boldsymbol{\Gamma_L}$ implies

$$U_a\big(\lambda\mu_1 + (1-\lambda)\mu_2, \lambda a_1 + (1-\lambda)a_2\big) \geq \min\big\{U_a(\mu_1, a_1), U_a(\mu_2, a_2)\big\}.$$

4. Examples and Applications

In this section, we present several examples of games that fit under $\boldsymbol{\Gamma_H}$ or $\boldsymbol{\Gamma_L}$ as well as an application to a game of information disclosure between a sender and multiple receivers.

4.1. Examples

Example 1 (Beauty contests)
Consider a beauty contest game in which each player $i \in N$ has a payoff

$$u^i(\theta, a) = -(1-\beta_i)\Big(\theta_i - a_i\Big)^2 - \beta_i \left(\sum_{j \neq i} \frac{a_j}{n-1} - a_i\right)^2$$

with $\beta_i \in (0,1)$. Each player i wants to match her action to her payoff-relevant state $\tilde{\theta}_i$ as well as to the average action of the other players. The parameter β_i measures the strength of strategic complementarity. The formulation allows for payoff asymmetry as well as independent or correlated states, in contrast to the classical formulation with symmetric payoffs and a pure common value setting (Keynes, 1936; Morris and Shin, 2002). Beauty contests belong to $\mathbf{\Gamma_H} \cap \mathbf{\Gamma_L}$.

Example 2 (Joint projects)
Consider N players participating in a joint project. Each player $i \in N$ has a payoff

$$u^i(\theta, a) = \nu_i(\theta_i) \prod_{j=1}^{n} a_j - c_i a_i^2,$$

where $a_i \in [0,1]$, $c_i > 0$, and $\nu_i : \mathbb{R} \to \mathbb{R}$ is an increasing function with $\nu_i(0) = 0$. Each player chooses how much costly effort to exert on a group project. The project succeeds with probability $\prod_{j=1}^{n} a_j$ and yields the player a random monetary payoff of $\tilde{\theta}_i$ or fails and yields zero monetary payoff. A player's preferences over monetary outcomes is captured by $\nu_i(\cdot)$. This game belongs to $\mathbf{\Gamma_H}$.

Example 3 (Network games)
Consider a network of N players represented by a (possibly asymmetric) $n \times n$ matrix $\boldsymbol{\sigma}$ with σ_{ij} capturing how much player i interacts with player j. Each player $i \in N$ has a payoff

$$u^i(\theta, a) = \Big(Y_i(\theta) + \sum_{j \neq i} \sigma_{i,j} X_i(a_j)\Big) a_i - c_i a_i^2,$$

where $c_i > 0$, and $Y_i : \Theta \to \mathbb{R}_+$ and $X_i : A \to \mathbb{R}_+$ are non-negative and increasing functions. Each player i chooses how much effort to exert. Effort is costly but yields a direct marginal benefit of $Y_i(\theta)$ that is increasing in θ. Additionally, there are positive spillovers in the form of peer effects captured by $\sigma_{ij} X_i(a_j)$: the more i interacts with j and the more effort j exerts, the higher the peer effect. This game belongs to $\mathbf{\Gamma_H}$ (resp., $\mathbf{\Gamma_L}$) if X_i is convex (resp., concave) for all $i \in N$.

Example 4 (Portfolio choice by a prudent and risk-averse agent)
Consider a single risk-averse agent with a Bernoulli utility $\vartheta : \mathbb{R} \to \mathbb{R}$ which is continuous, strictly increasing, and strictly concave. She can place a fraction $a \in [0, 1]$ of her wealth $W > 0$ in stocks, which yield a random rate of return $\tilde{x}$ that is distributed according to P_θ on $[\underline{x}, \bar{x}]$ with $\underline{x} < 0 < \bar{x}$. The remaining $1 - a$ of her wealth is in cash which yields a zero rate of return. Her ex-post payoff is given by

$$u(\theta, a) = \int_{\underline{x}}^{\bar{x}} \vartheta\Big(W\big(1 + ax\big)\Big) dP_\theta(x).$$

The parameter θ captures the volatility in stocks. In particular, for $\theta'' > \theta'$,

$$\int_0^z x\big[dP_{\theta''}(x) - dP_{\theta'}(x)\big] \geq 0, \tag{RS}$$

for all $z \in [\underline{x}, \bar{x}]$, with equality when $z = \bar{x}$. Hence, θ does not affect the average rate of return on stocks but a higher θ is associated with a lower volatility. Rothschild and Stiglitz (1971) show that all risk-averse agents invest more in a risky asset distributed according to $P_{\theta''}$ than $P_{\theta'}$ if, and only if, (RS) holds. Additionally, we assume that $\int_{\underline{x}}^{\bar{x}} x dP_\theta(x) = E[\tilde{x}] > 0$; otherwise, the agent will never invest in stocks.

When the investor is prudent, i.e., $\vartheta'''(x) > 0$, the absolute prudence coefficient $-\frac{\vartheta'''(x)}{\vartheta''(x)}$ is non-increasing, and the relative prudence coefficient satisfies $-\frac{\vartheta'''(x)}{\vartheta''(x)}x \leq 1$, then $\Gamma \in \boldsymbol{\Gamma_H}$.

4.2. Application: Information Disclosure

We consider a sender who chooses an information structure for a number of receivers with the intention of affecting the actions the receivers play in a Bayesian game. In the case of a single receiver, this problem corresponds to the information disclosure game of Rayo and Segal (2010) as well as the seminal Bayesian persuasion problem of Kamenica and Gentzkow (2011). The latter show that the persuasion problem is equivalent to maximizing the sender's interim expected payoff over Bayes-plausible distributions of the receiver's posterior beliefs. Mathevet et al. (2020) extend this belief-based approach to many receivers and show that the information disclosure problem is equivalent to maximizing the sender's interim expected payoff over consistent and Bayes-plausible distributions of the receivers' hierarchies of beliefs.

These results are powerful and general as they allow the sender full flexibility in what information to disclose to the receiver(s). However, they are also difficult to use in applications. For example, deriving the sender's interim value function from the primitives of a persuasion problem is a non-trivial task which requires a closed form solution to the receiver's optimization strategy. The literature has mostly focused on tractable binary environments (binary action and state spaces) or when the optimal strategy of the receiver depends only on the posterior mean. The problem often becomes more intractable when considering applications with many receivers or a continuum of actions and states.

We thus depart from the generality in the literature and restrict the sender to choose monotone information structures such that the receivers' signals are conditionally independent. We then apply Theorem 1 to characterize conditions on the preferences of the sender and the receivers that give maximal or minimal information disclosure.

Formally, there are n receivers who play a basic game $\Gamma = \langle N, \{A_i, u^i\}_{i\in N}, F\rangle$. The payoff $u^i : \Theta \times A \to \mathbb{R}$ satisfies Assumption 2 for all $i \in N$ and the common prior F satisfies Assumption 1. The sender has preferences over the state and the receivers' actions given by $v : \Theta \times A \to \mathbb{R}$ which is continuous in a for all $\theta \in \Theta$.

The information disclosure game is composed of two stages: First, without observing the state, the sender publicly chooses a profile of information structures Σ_ρ. Then, each receiver/player $i \in N$ privately observes a signal realization $s_i \in S_i$ generated by $\Sigma_{\rho_i} = \langle S_i, G(\cdot,\cdot;\rho_i)\rangle$ and participates in the Bayesian game (Σ_ρ, Γ). We assume that all players coordinate on either the greatest or least BNE, which are necessarily monotone pure-strategies. The sender's ex-ante payoff is given by

$$V(\rho) = \int_{\Theta\times S} v(\theta, a^\star(s;\rho))d\boldsymbol{G}(\theta, s;\rho).$$

Let $\mathcal{P}$ be the set of information structures that satisfy Assumption 3 and Assumption 4. The no-information structure trivially belongs to $\mathcal{P}$. Similarly, player i's full-information structure that reveals $\tilde{\theta}_i$ is in $\mathcal{P}$.[3] We refer to these two information structures as the minimal and maximal disclosure policies. We restrict the sender to choose from information structures in $\mathcal{P}$. Thus, the sender's problem is $\sup_{\Sigma_\rho\in\mathcal{P}} V(\rho)$.

Proposition 2 *Suppose $v(\theta, a)$ is component-wise convex (resp., component-wise concave) in a, has increasing differences (resp., decreasing differences) in $(\theta, a_{-i}; a_i)$ for each $i \in N$, and one of the following holds:*

i) $\Gamma \in \boldsymbol{\Gamma_H}$ *and $v(\theta, a)$ is increasing (resp., decreasing) in a,*

ii) $\Gamma \in \boldsymbol{\Gamma_L}$ *and $v(\theta, a)$ is decreasing (resp., increasing) in a, or*

iii) $\Gamma \in \boldsymbol{\Gamma_H} \cap \boldsymbol{\Gamma_L}$.

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

[3]Note the difference from an information structure that reveals $\tilde{\theta} = (\tilde{\theta}_1, \ldots, \tilde{\theta}_n)$ to player i.

Proposition 2 provides sufficient conditions under which there is minimal and maximal conflict between the sender and the receiver(s): if their desire to correlate actions and states goes in the same (opposite) direction and the sender likes (dislikes) dispersion of the actions, there will be maximal (minimal) disclosure.

We are not the first to study conditions under which there is maximal or minimal information disclosure. For the case with multiple receivers, Taneva (2017) exploits the equivalence between Bayes-correlated equilibria and Bayesian Nash equilibria with information disclosure (Bergemann and Morris, 2016) to characterize conditions for maximal and minimal information disclosure in binary environments with symmetric payoffs. Bergemann and Morris (2013) similarly use Bayes-correlated equilibria to characterize conditions for maximal and minimal information sharing between firms in a linear-quadratic setting. While these papers allows the sender to choose any information structure, our result allows for richer state and action spaces and richer preferences while restricting the set of information structures available to the sender.

For the case of a single receiver, Kamenica and Gentzkow (2011) show that the sender will disclose all (resp., no) information if the sender's interim value function is convex (resp., concave) in the receiver's posterior beliefs. Kolotilin (2018) and Dworczak and Martini (2019) use duality theory to derive conditions for maximal and minimal information disclosure when the sender's interim utility depends only on the posterior mean. Mensch (2019) derives novel single-crossing conditions on what he calls the sender's "virtual utility" to characterize maximal and minimal information disclosure in environments with complementarities. However, the sender's interim value function or the sender's virtual utility are often complicated endogenous functions. Beyond binary environments or settings where the posterior mean is sufficient, it is unclear what conditions on the primitives of a persuasion problem would imply the desired conditions on the sender's interim/virtual utility. In contrast, our assumptions are directly on the primitives of the persuasion problem and do not necessitate computing the sender's interim value function.

When there is a single receiver, the conditional independence assumption (Assumption 4) is no longer relevant. In the special case with binary states, all posteriors are first-order ranked. This implies it is without loss of generality to focus on monotone information structures, i.e., Assumption 3 holds trivially. Any information structure is both dominated by the full-information structure and dominates the no-information structure in the supermodular stochastic order. Hence, Proposition 2 along with Theorem 1 imply maximal (resp., minimal) disclosure is the sender's optimal policy out of ALL possible information structures. The next result extends a part of this result to non-binary environments using Theorem 2, which allows for non-monotone information structures.

Theorem 3 *Let $n = 1$. Full-information revelation is the optimal disclosure policy among all possible information structures if $v(\theta, a)$ is convex in a, has increasing differences in $(\theta; a)$ and one of the following holds:*

i) $\Gamma \in \boldsymbol{\Gamma_H}$ *and $v(\theta, a)$ is increasing in a,*

ii) $\Gamma \in \boldsymbol{\Gamma_L}$ *and $v(\theta, a)$ is decreasing in a, or*

iii) $\Gamma \in \boldsymbol{\Gamma_H} \cap \boldsymbol{\Gamma_L}$.

Intuitively, the full-information structure is Blackwell more informative than any other signal and is trivially a monotone information structure because Θ is a totally ordered set when $n = 1$. Any other information structure, monotone or not, is necessarily a garbling of the full-information structure. Thus, when the sender can use any information structure, Theorem 2 and the conditions in Theorem 3 imply that there is minimal conflict between the sender and the receiver, establishing the optimality of full disclosure.

Once again, notice that the conditions in Theorem 3 are on the primitive preferences of the sender and the receiver. We do not need to assume further conditions, such as convexity of the sender's interim value function as in Kamenica and Gentzkow (2011), or quasiconvexity of the receiver's interim utility as in Jensen (2018).

Appendix A.

This section contains results that we use to prove our main theorem. These are not our results, and are provided for the ease of the reader.

Given a pure strategy $\alpha_i : S_i \to A_i$ for player $i \in N$ and ex-ante distribution of actions $H(\cdot; \alpha_i) : \mathbb{R} \to [0, 1]$, define the quantile function $\hat{a}(\cdot; \alpha_i) : [0, 1] \to \mathbb{R}$ by

$$\hat{a}(q; \alpha_i) = \inf\{z : q \leq H(z; \alpha_i)\}$$

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

Lemma A.1 *[Theorem 4.A.2-A.3 of Shaked and Shanthikumar (2007)]*
Given two pure strategies α_i and α_i', the following are equivalent:

(i) $\alpha_i \succeq_H \alpha_i'$.

(ii) For all $x \in \mathbb{R}$, $\int_x^\infty H(z; \alpha_i) dz \leq \int_x^\infty H(z; \alpha_i') dz$.

(iii) For all $t \in [0, 1]$, $\int_t^1 \hat{a}(q; \alpha_i) dq \geq \int_t^1 \hat{a}(q; \alpha_i') dq$.

Similarly, the following are equivalent:

(iv) $\alpha_i \succeq_L \alpha_i'$.

(v) *For all* $x \in \mathbb{R}$, $\int_{-\infty}^{x} H(z;\alpha_i)dz \geq \int_{-\infty}^{x} H(z;\alpha_i')dz$.

(vi) *For all* $t \in [0,1]$, $\int_0^t \hat{a}(q;\alpha_i)dq \leq \int_0^t \hat{a}(q;\alpha_i')dq$.

Lemma A.2 *[Theorem 3.8.2 of Müller and Stoyan (2002) or Tchen (1980)]*
Given two information structures Σ_{ρ_i} *and* $\Sigma_{\rho_i'}$, $\rho_i \succeq_{spm} \rho_i'$ *if, and only if, for all integrable functions* $\psi : \Theta_i \times S_i \to \mathbb{R}$ *that satisfy increasing differences (ID) in* $(\theta_i; s_i)$,

$$\int_{\Theta_i \times S_i} \psi(\theta_i, s_i) dG(\theta_i, s_i; \rho_i) \geq \int_{\Theta_i \times S_i} \psi(\theta_i, s_i) dG(\theta_i, s_i; \rho_i').$$

Lemma A.3 *[Lemma 1 of Quah and Strulovici (2009)]*
Let $g : [\underline{x}, \bar{x}] \to \mathbb{R}$ *and* $h : [\underline{x}, \bar{x}] \to \mathbb{R}$ *be integrable functions.*

1. *If* g *is increasing and* $\int_z^{\bar{x}} h(x)dx \geq 0$ *for all* $z \in [\underline{x}, \bar{x}]$, *then* $\int_{\underline{x}}^{\bar{x}} g(x)h(x)dx \geq g(\underline{x}) \int_{\underline{x}}^{\bar{x}} h(x)dx$.
2. *If* g *is decreasing and* $\int_{\underline{x}}^{z} h(x)dx \geq 0$ *for all* $z \in [\underline{x}, \bar{x}]$, *then* $\int_{\underline{x}}^{\bar{x}} g(x)h(x)dx \geq g(\bar{x}) \int_{\underline{x}}^{\bar{x}} h(x)dx$.

Appendix B.

Proof of Theorem 1. ($\Longrightarrow$) We only prove the case for $\boldsymbol{\Gamma_H}$. A symmetric argument establishes the result for the case for $\boldsymbol{\Gamma_L}$. Without loss of generality, let the marginal G_{S_i} be the uniform distribution on the unit interval for each $i \in N$ (see Footnote 1).

Fix a basic game $\Gamma \in \boldsymbol{\Gamma_H}$. For each player $i \in N$, let $\alpha_i : S_i \to A_i$ be an arbitrary measurable and monotone strategy. Since α_i is monotone, it is almost everywhere equal to its quantile function, i.e., $\alpha_i(s_i) = \hat{a}(s_i;\alpha_i)$ for almost all $s_i \in [0,1] = S_i$. Thus, given two monotone strategies α_i and α_i', from Lemma A.1, $\alpha_i \succeq_H \alpha_i'$ if, and only if,

$$\int_t^1 \alpha_i(s)ds_i \geq \int_t^1 \alpha_i'(s_i)ds_i$$

for all $t \in [0,1]$.

Let $\mathcal{A}_i$ be the set of all monotone and measurable strategies and let $\mathcal{A} = \times_{i\in N}\mathcal{A}_i$. Given a profile of information structures $\Sigma_\rho = (\Sigma_{\rho_1}, \ldots, \Sigma_{\rho_n})$ and opponents' strategies $\alpha_{-i} \in \mathcal{A}_{-i}$, let $a_i^{BR}(\cdot;\alpha_{-i},\rho) : S_i \to A_i$ be player i's best response strategy. Specifically, for all $s_i \in S_i$,

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

Using Assumption 1-Assumption 4 and monotone comparative statics of Bayesian supermodular games (Van Zandt and Vives, 2007), $a_i^{BR}(\cdot;\alpha_{-i},\rho) \in \mathcal{A}_i$ for all $i \in N$.

For any given profile of monotone strategies $\alpha = (\alpha_1,\ldots,\alpha_n) \in \mathcal{A}$, denote the profile of best-response strategies by $a^{BR}(\alpha,\rho) = \big(a_1^{BR}(\cdot;\alpha_{-1},\rho),\ldots,a_n^{BR}(\cdot;\alpha_{-n},\rho)\big)$. Then, a monotone BNE $a^\star(\rho)$ of a Bayesian game $\mathcal{G}_\rho = (\Sigma_\rho,\Gamma)$ is given by the fixed point $a^{BR}(a^\star(\rho),\rho) = a^\star(\rho)$.

The proof proceeds in three steps:

1. Player i's best response strategy increases in the increasing convex order when any player j's information quality increases in the supermodular stochastic order (Lemma B.1). This concludes the proof if $n=1$.

2. For all $j \in N\backslash\{i\}$, player i's best response strategy increases in the increasing convex order when player j's strategy increases in the increasing convex order (Lemma B.2).

3. Given 1-2, apply comparative statics on fixed points to get desired result.

Lemma B.1 *Fix some $i \in N$ and some monotone strategy $\alpha_{-i} \in \mathcal{A}_{-i}$. Take two profiles of information structures $\Sigma_{\rho''} = (\Sigma_{\rho''_j},\Sigma_{\rho_{-j}})$ and $\Sigma_{\rho'} = (\Sigma_{\rho'_j},\Sigma_{\rho_{-j}})$ for some $j \in N$. If $\rho''_j \succeq_{spm} \rho'_j$, then $a_i^{BR}(\cdot;\alpha_{-i},\rho'') \succeq_H a_i^{BR}(\cdot;\alpha_{-i},\rho')$.*

Proof. To economize on notation, we suppress the dependence of $a_i^{BR}(\cdot;\alpha_{-i},\rho'_i)$ on α_{-i}. For any signal realization $s_i \in S_i$, the first order conditions imply that

$$\int_{\Theta\times S_{-i}} \Big[u^i_{a_i}\big(\theta,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\rho'')\big) - u^i_{a_i}\big(\theta,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\rho')\big)\Big]\, d\boldsymbol{G}(\theta,s_{-i}|s_i;\rho'')$$
$$+\int_{\Theta\times S_{-i}} u^i_{a_i}\big(\theta,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\rho')\big)\Big[d\boldsymbol{G}(\theta,s_{-i}|s_i;\rho'') - d\boldsymbol{G}(\theta,s_{-i}|s_i;\rho')\Big] = 0$$

As $\Gamma \in \boldsymbol{\Gamma_H}$, $u^i_{a_i}(\theta,a)$ is convex in a_i for all $(\theta,a_{-i}) \in \Theta\times A_{-i}$. Thus,

$$\begin{aligned}&u^i_{a_i}\big(\theta,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\rho'')\big) - u^i_{a_i}\big(\theta,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\rho')\big)\\ \geq\, &u^i_{a_ia_i}\big(\theta,\alpha_{-i}(s_{-i}),a_i^{BR}(s_i;\rho')\big)\Big(a_i^{BR}(s_i;\rho'') - a_i^{BR}(s_i;\rho')\Big),\end{aligned}$$

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

$$\int_t^1 \Big(a_i^{BR}(s_i;\rho') - a_i^{BR}(s_i;\rho'')\Big) ds_i$$

$$\leq \int_t^1 B(s_i) \int\limits_{\Theta \times S_{-i}} u^i_{a_i}\big(\theta, \alpha_{-i}(s_{-i}), a_i^{BR}(s_i;\rho')\big)\Big[d\boldsymbol{G}(\theta, s_{-i}|s_i;\rho') - d\boldsymbol{G}(\theta, s_{-i}|s_i;\rho'')\Big] ds_i$$

$$= \int\limits_{\Theta_j \times S_j} \underbrace{\int\limits_{\Theta_{-j}\times S_{-j}} \mathbf{1}_{[s_i \geq t]} B(s_i) u^i_{a_i}\big(\theta, \alpha_{-i}(s_{-i}), a_i^{BR}(s_i;\rho')\big) d\boldsymbol{G}(\theta_{-j}, s_{-j}|\theta_j;\rho_{-j})}_{\triangleq \psi(\theta_j, s_j; t)} \Big[dG(\theta_j, s_j;\rho'_j) - dG(\theta_j, s_j;\rho''_j)\Big]$$

where the last equality follows from Assumption 4, and

$$B(s_i) = \left(- \int_{\Theta \times S_{-i}} u^i_{a_i a_i}\big(\theta, \alpha_{-i}(s_{-i}), a_i^{BR}(s_i;\rho')\big) d\boldsymbol{G}(\theta, s_{-i}|s_i,\rho'')\right)^{-1}.$$

Note that $B(s_i) > 0$ by the concavity of u^i in a_i. Additionally, it is an increasing function because $-u^i_{a_i}$ has decreasing differences in $(\theta, a_{-i}; a_i)$, is concave in a_i, and $\boldsymbol{G}(\tilde{\theta}, \tilde{s}_{-i}|s_i;\rho'')$ is increasing in FOSD as s_i increases by Assumption 1, Assumption 3, and Assumption 4.

First, we consider the case when $i = j$. For any $s''_i > s'_i$, we have

$$\psi(\theta_i, s''_i; t) - \psi(\theta_i, s'_i; t)$$

$$= \Big(\mathbf{1}_{[s''_i \geq t]} B(s''_i) - \mathbf{1}_{[s'_i \geq t]} B(s'_i)\Big) \int\limits_{\Theta_{-i}\times S_{-i}} u^i_{a_i}\big(\theta, \alpha_{-i}(s_{-i}), a_i^{BR}(s''_i;\rho')\big) d\boldsymbol{G}(\theta_{-i}, s_{-i}|\theta_i;\rho_{-i})$$

$$+ \mathbf{1}_{[s'_i \geq t]} B(s'_i) \int\limits_{\Theta_{-i}\times S_{-i}} \Big[u^i_{a_i}\big(\theta, \alpha_{-i}(s_{-i}), a_i^{BR}(s''_i;\rho')\big) - u^i_{a_i}\big(\theta, \alpha_{-i}(s_{-i}), a_i^{BR}(s'_i;\rho')\big)\Big] d\boldsymbol{G}(\theta_{-i}, s_{-i}|\theta_i;\rho_{-i}),$$

which is an increasing function of θ_i. To see why, notice that *(i)* strategies are monotone in signal realizations, *(ii)* $\mathbf{1}_{[s_i \geq t]} B(s_i) \geq 0$ and increasing in s_i, *(iii)* the first integrand is increasing in (θ, s_{-i}) since u^i has increasing differences in $(\theta, a_{-i}; a_i)$, *(iv)* the second integrand is increasing in (θ, s_{-i}) since $u^i_{a_i}$ has increasing differences in $(\theta, a_{-i}; a_i)$, and *(v)* $G(\tilde{\theta}_{-i}, \tilde{s}_{-i}|\theta_i, \rho_{-i})$ is increasing in FOSD as θ_i increases.

Next, we consider the case when $i \neq j$. For any $s''_j > s'_j$, we have

$$\psi(\theta_j, s''_j; t) - \psi(\theta_j, s'_j; t)$$

$$= \int\limits_{\Theta_{-j}\times S_{-j}} \mathbf{1}_{[s_i \geq t]} B(s_i) \Big[u^i_{a_i}\big(\theta, \alpha_{-i}(s''_{-i}), a_i^{BR}(s_i;\rho')\big) - u^i_{a_i}\big(\theta, \alpha_{-i}(s'_{-i}), a_i^{BR}(s_i;\rho')\big)\Big] d\boldsymbol{G}(\theta_{-j}, s_{-j}|\theta_j;\rho_{-j}),$$

with $\alpha_{-i}(s_{-i}) = (\alpha_{-i,j}(s_{-i,j}), \alpha_j(s_j))$. The difference is increasing in θ_j because (i) strategies are monotone in signal realizations, (ii) $\mathbf{1}_{[s_i \geq t]} B(s_i) \geq 0$ and increasing in s_i, (iii) the term in the square brackets is non-negative since u^i has increasing differences in $(a_j; a_i)$, (iv) the term in the square brackets is increasing in (θ, s_{-j}) since $u^i_{a_i}$ has increasing differences in $(\theta, a_{-j}; a_j)$, and (v) $G(\tilde{\theta}_{-j}, \tilde{s}_{-j} | \theta_j, \rho_{-j})$ is increasing in FOSD as θ_j increases.

In both cases, we have just shown that $\psi(\theta_j, s_j; t)$ has increasing differences in $(s_j; \theta_j)$ for all $t \in [0,1]$. We can therefore conclude that for each $t \in [0,1]$,

$$\int_t^1 \Big(a_i^{BR}(s_i; \rho') - a_i^{BR}(s_i; \rho'') \Big) ds_i \leq \int\limits_{\Theta_j \times S_j} \psi(\theta_j, s_j; t) \Big[dG(\theta_j, s_j; \rho'_j) - dG(\theta_j, s_j; \rho''_j) \Big] \leq 0$$

where the last inequality follows from Lemma A.2. Thus, $a_i^{BR}(\cdot; \alpha_{-i}, \rho'') \succeq_H a_i^{BR}(\cdot; \alpha_{-i}, \rho')$. ■

Lemma B.2 *Fix $i, j \in N$ with $j \neq i$, a monotone strategy $\alpha_{-i,j} \in \mathcal{A}_{-i,j}$, and an information structures Σ_ρ. For $\alpha''_j, \alpha'_j \in \mathcal{A}_j$ such that $\alpha''_j \succeq_H \alpha'_j$, $a_i^{BR}(\cdot; \alpha''_{-i}, \rho) \succeq_H a_i^{BR}(\cdot; \alpha'_{-i}, \rho)$.*

Proof. We suppress the dependence on Σ_ρ as it is held fixed. For any $t \in [0,1]$, we use the first order conditions argument to get the expression

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

$$\leq \int\limits_{\Theta \times S} \mathbf{1}_{[s_i \geq t]} \tilde{B}_i(s_i) \left[u^i_{a_i}\Big(\theta, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i; \alpha'_{-i})\Big) - u^i_{a_i}\Big(\theta, \alpha''_{-i}(s_{-i}), a_i^{BR}(s_i; \alpha'_{-i})\Big) \right] d\boldsymbol{G}(\theta, s),$$

where

$$\tilde{B}(s_i) = \left(- \int\limits_{\Theta \times S_{-i}} u^i_{a_i a_i}\Big(\theta, \alpha''_{-i}(s_{-i}), a_i^{BR}(s_i; \alpha'_{-i})\Big) d\boldsymbol{G}(\theta, s_{-i} | s_i) \right)^{-1}.$$

Note that $\tilde{B}(s_i)$ has the same properties as $B(s_i)$ in Lemma B.1 for the same reasons, i.e., $\tilde{B}(s_i) > 0$ and increasing in s_i.

By convexity of $u^i_{a_i}$ in a_j,

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

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

Thus,

$$\int_t^1 \left(a_i^{BR}(s_i;\alpha'_{-i}) - a_i^{BR}(s_i;\alpha''_{-i})\right)ds_i$$
$$\leq \int\limits_{S_j} \left(\alpha'_j(s_j) - \alpha''_j(s_j)\right) \int\limits_{\Theta\times S_{-j}} \mathbf{1}_{[s_i\geq t]}\tilde{B}(s_i)u^i_{a_ia_j}\Big(\theta, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)d\boldsymbol{G}(\theta, s_{-j}|s_j)ds_j.$$

As $\alpha''_j, \alpha'_j \in \mathcal{A}_j$, $\alpha''_j \succeq_H \alpha'_j$ if, and only if,

$$\int_t^1 \left(\alpha'_j(s_j) - \alpha''_j(s_j)\right)ds_j \leq 0, \ \ \forall t \in [0,1].$$

Furthermore,

$$\int\limits_{\Theta\times S_{-j}} \mathbf{1}_{[s_i\geq t]}\tilde{B}(s_i)u^i_{a_ia_j}\Big(\theta, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)d\boldsymbol{G}(\theta, s_{-j}|s_j)$$

is an increasing function of s_j because (i) strategies are monotone in signal realizations, (ii) $\mathbf{1}_{[s_i\geq t]}\tilde{B}(s_i) \geq 0$ and increasing in s_i, (iii) $u^i_{a_ia_j}\Big(\theta, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big) \geq 0$ since u^i has increasing differences in $(a_j;a_i)$, (iv) $u^i_{a_ia_j}\Big(\theta, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)$ is increasing in (θ, s) since $u^i_{a_i}$ has increasing differences in $(\theta, a_{-j}; a_j)$ and is convex in a_j, and (v) $\boldsymbol{G}(\tilde{\theta}, \tilde{s}_{-j}|s_j)$ is increasing in FOSD as s_j increases by Assumption 1, Assumption 3, and Assumption 4.

Applying Lemma A.3, we have

$$\int_t^1 \left(a_i^{BR}(s_i;\alpha'_{-i}) - a_i^{BR}(s_i;\alpha''_{-i})\right)ds_i$$
$$\leq \int\limits_{S_j} \left(\alpha'_j(s_j) - \alpha''_j(s_j)\right) \int\limits_{\Theta\times S_{-j}} \mathbf{1}_{[s_i\geq t]}\tilde{B}(s_i)u^i_{a_ia_j}\Big(\theta, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)d\boldsymbol{G}(\theta, s_{-j}|s_j)ds_j$$
$$\leq \underbrace{\int\limits_{S_j} \left(\alpha'_j(s_j) - \alpha''_j(s_j)\right)ds_j}_{\leq 0} \int\limits_{\Theta\times S_{-j}} \underbrace{\mathbf{1}_{[s_i\geq t]}\tilde{B}(s_i)u^i_{a_ia_j}\Big(\theta, \alpha'_{-i}(s_{-i}), a_i^{BR}(s_i;\alpha'_{-i})\Big)}_{\geq 0}d\boldsymbol{G}(\theta, s_{-j}|s_j=0)$$
$$\leq 0$$

for each $t \in [0,1]$. Thus, $a_i^{BR}(\cdot;\alpha''_{-i},\rho) \succeq_H a_i^{BR}(\cdot;\alpha'_{-i},\rho)$. ■

We now tackle the last step in the "if" part of the proof: comparative statics of the BNEs. We apply the comparative statics of fixed points from Villas-Boas (1997). To do so, we will need the following definitions. Let $BV([0,1],\mathbb{R})$ be the space of functions of bounded variation from $[0,1]$ to $\mathbb{R}$. Given a function $g \in BV([0,1],\mathbb{R})$, the bounded variation norm is given by

$$||g||_{BV} = \int_0^1 |g(s)|ds + \sup_{p\in P} \sum_{i=0}^{n_p-1} |g(x_{i+1}) - g(x_i)|$$

where P is the set of all partitions $p = \{x_0, x_1, \ldots, x_{n_p}\}$ on $[0,1]$. The space $BV([0,1],\mathbb{R})$ equipped with the $||\cdot||_{BV}$ norm is a Banach space.

Definition B.1 (Contractible Space) *A topological space X is contractible if there exists some $x^* \in X$ and a map $\Phi : X \times [0,1] \to X$ such that*

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

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

Theorem B.1 *[Theorem 6 & 7; Villas-Boas (1997)] Let X be a compact subset of a Banach space. Consider a transitive and reflexive order $\succeq$ on X such that, for all $x \in X$, the upper-sets $\mathcal{U}(x) = \{x' \in X : x' \succeq x\}$ and lower-sets $\mathcal{L}(x) = \{x' \in X : x \succeq x'\}$ are compact and contractible. Let $T_1 : X \to X$ and $T_2 : X \to X$ be two continuous mappings.*

A. Suppose $x' \succeq x \Rightarrow T_1(x') \succeq T_1(x)$, and suppose $T_1(x) \succeq T_2(x)$ for all $x \in X$. Then for every fixed point $x_2^\star$ of T_2, there is a fixed point $x_1^\star$ of T_1 such that $x_1^\star \succeq x_2^\star$.

B. Suppose $x' \succeq x \Rightarrow T_2(x') \succeq T_2(x)$, and suppose $T_1(x) \succeq T_2(x)$ for all $x \in X$. Then for every fixed point $x_1^\star$ of T_1, there is a fixed point $x_2^\star$ of T_2 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.

Lemma B.3 *For each $i \in N$, $\mathcal{A}_i$ is a compact subset of $\big(BV([0,1],\mathbb{R}), ||\cdot||_{BV}\big)$.*

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

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

Let $\mathcal{U}(\alpha_i) = \{\alpha_i' \in \mathcal{A}_i : \alpha_i' \succeq_H \alpha_i\}$ and $\mathcal{L}(\alpha_i) = \{\alpha_i' \in \mathcal{A}_i : \alpha_i \succeq_H \alpha_i'\}$ be the upper and lower-sets of $\mathcal{A}_i$ respectively.

Lemma B.4 *For each $i \in N$ and for any $\alpha_i \in \mathcal{A}_i$, $\mathcal{U}(\alpha_i)$ and $\mathcal{L}(\alpha_i)$ are compact and contractible.*

Proof. For a given $\alpha_i \in \mathcal{A}_i$, $\mathcal{U}(\alpha_i)$ and $\mathcal{L}(\alpha_i)$ are closed subsets of $\mathcal{A}_i$ (follows from the dominated convergence Theorem). Hence, they are compact. Let $\alpha_i^u : [0,1] \to A_i$ be a constant function with $\alpha_i^u(s_i) = \bar{a}_i$ for all $s_i \in [0,1]$. Note that $\alpha_i^u \in \mathcal{A}_i$. Furthermore, $\alpha_i^u(s_i) \geq \alpha_i(s_i)$, $\forall s_i \in [0,1]$ which implies $\alpha_i^u \succeq_H \alpha_i \Rightarrow \alpha_i^u \in \mathcal{U}(\alpha_i)$.

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

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

As λ increases from 0 to 1, Φ^u continuously deforms any strategy in $\mathcal{U}(\alpha_i)$ to the constant strategy α_i^u, which is itself in $\mathcal{U}(\alpha_i)$. Therefore, $\mathcal{U}(\alpha_i)$ is contractible.

Similarly, let $\alpha_i^\ell : [0,1] \to A_i$ be a constant function with $\alpha_i^\ell(s_i) = \underline{a}_i$ for all $s_i \in [0,1]$. Again, $\alpha_i^\ell \in \mathcal{A}_i$. Furthermore, $\alpha_i^\ell(s_i) \leq \alpha_i(s_i)$, $\forall s_i \in [0,1]$ which implies $\alpha_i \succeq_H \alpha_i^\ell \Rightarrow \alpha_i^\ell \in \mathcal{L}(\alpha_i)$. Then for each $\alpha_i \in \mathcal{A}_i$, define the mapping $\Phi^\ell : \mathcal{L}(\alpha_i) \times [0,1] \to \mathcal{L}(\alpha_i)$ such that

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

As λ increases from 0 to 1, Φ^ℓ continuously deforms any strategy in $\mathcal{L}(\alpha_i)$ to the constant strategy α_i^ℓ, which is itself in $\mathcal{L}(\alpha_i)$. Therefore, $\mathcal{L}(\alpha_i)$ is contractible. ■

Thus far, we have an order $\succeq_H$ on $\mathcal{A}_i$ that generates compact and contractible upper and lower-sets. We extend these properties to $\mathcal{A} = \times_{i\in N}\mathcal{A}_i$ by the product order: given $\alpha'', \alpha' \in \mathcal{A}$, $\alpha'' \succeq_H \alpha'$ if, and only if, $\alpha_i'' \succeq_H \alpha_i'$ for each $i \in N$. Along with the product topology, $\succeq_H$ is a partial order on $\mathcal{A}$ that generates compact and contractible upper and lower-sets.[4]

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

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

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

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

Consider two different games, $\mathcal{G}_{\rho''} = (\Sigma_{\rho''}, \Gamma)$ and $\mathcal{G}_{\rho'} = (\Sigma_{\rho'}, \Gamma)$, with $\rho'' \succeq_{spm} \rho'$. For all $\alpha \in \mathcal{A}$,

$$\rho'' \succeq_{spm} \rho' \quad \underbrace{\Rightarrow}_{\text{by Lemma B.1}} \quad a_i^{BR}(\alpha_{-i}, \rho'') \succeq_H a_i^{BR}(\alpha_{-i}, \rho'), \forall i \Leftrightarrow T_{\rho''}(\alpha) \succeq_H T_{\rho'}(\alpha).$$

Furthermore,

$$\alpha'' \succeq_H \alpha' \quad \underbrace{\Rightarrow}_{\text{by Lemma B.2}} \quad a_i^{BR}(\alpha''_{-i}, \rho) \succeq_H a_i^{BR}(\alpha'_{-i}, \rho), \forall i \Leftrightarrow T_{\rho}(\alpha'') \succeq_H T_{\rho}(\alpha').$$

We can now directly apply Theorem B.1 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 $a^\star(\rho'') \succeq_H a^\star(\rho')$, and for every fixed point $a^\star(\rho'')$ of $T_{\rho''}$, there is a fixed point $a^\star(\rho')$ of $T_{\rho'}$ such that $a^\star(\rho'') \succeq_H a^\star(\rho')$. Hence, players are more responsive with a higher mean under $\Sigma_{\rho''}$ than under $\Sigma_{\rho'}$. A symmetric argument establishes the analogous result for $\Gamma \in \boldsymbol{\Gamma_L}$.

($\Longleftarrow$) Given two profiles of information structures $\Sigma_{\rho''}$ and $\Sigma_{\rho'}$, $\rho'' \not\succeq_{spm} \rho'$ if there exists a player $i^* \in N$ such that $\rho''_{i^*} \not\succeq_{spm} \rho'_{i^*}$. From Lemma A.2, $\rho''_{i^*} \not\succeq_{spm} \rho'_{i^*}$ implies there exist a $(\theta^*_{i^*}, s^*_{i^*}) \in \Theta_{i^*} \times S_{i^*}$ such that

$$G(\theta^*_{i^*}, s^*_{i^*}; \rho''_{i^*}) < G(\theta^*_{i^*}, s^*_{i^*}; \rho'_{i^*}).$$

Consider a basic game $\Gamma = \langle \{A_i, u^i\}_{i\in N}, F\rangle$ such that $u^i : \Theta \times A \to \mathbb{R}$ is given by

$$u^i(\theta, a) = -\frac{1}{2}\Big(\bar{a}_i - \mathbf{1}_{[\theta_i \le \theta_i^*]}(\bar{a}_i - \underline{a}_i) - a_i\Big)^2$$

for all $i \in N$. Each player i's payoff depends only on her own action and whether her state is above or below some cutoff θ_i^*. Thus, each player acts as a single decision maker.

For all $i \in N$, $u^i(\theta, a)$ satisfies Assumption 2: It is continuous, twice differentiable, and strictly concave in a_i. It has increasing differences in $(\theta, a_{-i}; a_i)$. For each $(\theta, a_{-i}) \in \Theta \times A_{-i}$, the optimal action under complete information is $\underline{a}_i$ if $\theta_i \le \theta_i^*$ and $\bar{a}_i$ otherwise. Furthermore, the marginal utility $u^i_{a_i}(\theta, a) = \bar{a}_i - \mathbf{1}_{[\theta_i \le \theta_i^*]}(\bar{a}_i - \underline{a}_i) - a_i$ is

(i) linear in a_j, and

(ii) has constant differences in $(\theta, a_{-j}; a_j)$.

Therefore, $\Gamma \in \boldsymbol{\Gamma_H} \cap \boldsymbol{\Gamma_L}$. As each player acts as a single-decision maker, there is a unique BNE which is just a profile of each player's optimal choice. For any given Σ_ρ,

$$a_i^\star(s_i; \rho) = \bar{a}_i - (\bar{a}_i - \underline{a}_i) G(\theta_i^* | s_i; \rho_i).$$

Now consider player i^*; Given $\Sigma_{\rho'}$ and $\Sigma_{\rho''}$,

$$\begin{aligned}&\int_0^{s^*_{i^*}} \Big(a^\star_{i^*}(s_{i^*};\rho'') - a^\star_{i^*}(s_{i^*};\rho')\Big) dG_{S_{i^*}}(s_{i^*})\\ =&(\bar{a}_{i^*} - \underline{a}_{i^*})\Big(G(\theta^*_{i^*}, s^*_{i^*};\rho'_{i^*}) - G(\theta^*_{i^*}, s^*_{i^*};\rho''_{i^*})\Big) > 0,\end{aligned}$$

which implies $a^\star_{i^*}(\rho'') \not\succeq_L a^\star_{i^*}(\rho')$ (by Lemma A.1). By definition, the players are therefore not more responsive with a lower mean under $\Sigma_{\rho''}$ than $\Sigma_{\rho'}$. Notice that for any Σ_ρ,

$$\mathbb{E}[a^\star_{i^*}(\rho)] = \bar{a}_{i^*} - (\bar{a}_{i^*} - \underline{a}_{i^*}) \int_{S_{i^*}} G(\theta^*_{i^*}|s_i;\rho_i) dG_{S_{i^*}}(s_{i^*}) = \bar{a}_{i^*} - (\bar{a}_{i^*} - \underline{a}_{i^*}) F_{\Theta_{i^*}}(\theta^*_{i^*}),$$

which is independent of ρ. Thus,

$$\begin{aligned}&\int_{s^*_{i^*}}^1 \Big(a^\star_{i^*}(s_{i^*};\rho'') - a^\star(s_{i^*};\rho')\Big) dG_{S_{i^*}}(s_{i^*})\\ =&\underbrace{\int_{S_{i^*}} \Big(a^\star_{i^*}(s_{i^*};\rho'') - a^\star_{i^*}(s_{i^*};\rho')\Big) dG_{S_{i^*}}(s_{i^*})}_{\substack{=\mathbb{E}[a^\star_{i^*}(\rho'')] - E[a^\star_{i^*}(\rho')]\\ =0}} - \left(\underbrace{\int_0^{s^*_{i^*}} \Big(a^\star_{i^*}(s_{i^*};\rho'') - a^\star_{i^*}(s;\rho')\Big) dG_{S_{i^*}}(s_{i^*})}_{>0}\right) < 0,\end{aligned}$$

which implies $a^\star_{i^*}(\rho'') \not\succeq_H a^\star_{i^*}(\rho')$ (by Lemma A.1). By definition, the players are therefore not more responsive with a higher mean under $\Sigma_{\rho''}$ than $\Sigma_{\rho'}$. ■

Proof of Theorem 2. We prove the case for $\Gamma \in \boldsymbol{\Gamma_H}$; the other case is analogous. As $n = 1$, we suppress the indices for players.

Let $\mu \in \Delta(\Theta)$ be a probability measure representing an arbitrary belief the single-agent may hold. Define

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

Since $u(\theta, a)$ has increasing differences in $(\theta; a)$, we know that $a^*(\cdot)$ is an increasing function in the sense that $a^*(\mu_2) \geq a^*(\mu_1)$ whenever $\mu_2 \succeq_{FOSD} \mu_1$ (Athey, 2002).

Lemma B.5 *Let $\{\lambda_k\}_{k=1}^K \in [0,1]^K$ be a sequence of weights with $\sum_{k=1}^K \lambda_k = 1$ and let $\{\mu_k\}_{k=1}^K \in \Delta(\Theta)$ be a finite sequence of beliefs with $\mu_K \succeq_{FOSD} \mu_{K-1} \succeq_{FOSD} \cdots \succeq_{FOSD} \mu_1$. If $\Gamma \in \boldsymbol{\Gamma_H}$,*

$$a^*\left(\sum_{k=1}^K \lambda_k \mu_k\right) \leq \sum_{k=1}^K \lambda_k a^*(\mu_k).$$

If $\Gamma \in \boldsymbol{\Gamma_L}$, the opposite inequality holds.

Proof. Consider a basic game $\Gamma \in \mathbf{\Gamma_H}$. Suppose first that $K = 2$. (The case for $K = 1$ is clearly trivial). Let $a_k^* = a^*(\mu_k)$ for $k = 1, 2$, $a_\lambda = \lambda a_1^* + (1-\lambda)a_2^*$, and $\mu_\lambda = \lambda\mu_1 + (1-\lambda)\mu_2$. By the first order condition, we have that $U_a(\mu_k, a_k^*) = 0$.

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

where the first inequality follows from the convexity of u_a in a. Increasing differences of the utility $u(\theta, a)$ in $(\theta; a)$ along with $\mu_2 \succeq_{FOSD} \mu_1$ implies $a_2 \geq a_1$. By increasing differences of the marginal utility u_a in $(\theta; a)$, we have $u_a(\theta, a_1) - u_a(\theta, a_2)$ is a decreasing function of θ. The last inequality then follows from the definition of first-order stochastic dominance. Since the marginal value of a_λ is non-positive at μ_λ, we must have $a^*(\mu_\lambda) \leq a_\lambda$.

Next, suppose that $K > 2$. Since $\sum_{k=2}^K \frac{\lambda_k}{1-\lambda_1}\mu_k \succeq_{FOSD} \mu_1$, by the argument above, we have

$$
a^*\left(\lambda_1\mu_1 + (1-\lambda_1)\sum_{k=2}^K \frac{\lambda_k}{1-\lambda_1}\mu_k\right) \leq \lambda_1 a^*(\mu_1) + (1-\lambda_1)\, a^*\left(\sum_{k=2}^K \frac{\lambda_k}{1-\lambda_1}\mu_k\right).
$$

By induction,

$$
a^*\left(\sum_{k=1}^K \lambda_k\mu_k\right) \leq \sum_{k=1}^K \lambda_k a^*(\mu_k).
$$

A symmetric argument establishes the analogous result for $\Gamma \in \mathbf{\Gamma_L}$. ■

Consider any information structure Σ_ρ that satisfies Assumption 3. In particular, suppose Σ_ρ induces posteriors $\{\mu_k\}_{k=1}^K \in \Delta(\Theta)$ with corresponding probabilities $\{\tau_k^\rho\}_{k=1}^K$ such that $\mu_k \succeq_{FOSD} \mu_{k'}$ whenever $k > k'$. Let $\Sigma_{\rho'}$ be another information structure that induces posteriors $\{\nu_m\}_{m=1}^M$ with corresponding probabilities $\{\tau_m^{\rho'}\}_{m=1}^M$. Note that $\Sigma_{\rho'}$ does not have to satisfy Assumption 3.b-d.[5] If $\Sigma_{\rho'}$ is a garbling of Σ_ρ, then there exist weights $\{\{\lambda_k^m\}_{k=1}^K\}_{m=1}^M \in [0,1]^{K\times M}$ such that

(*i*) $\sum_{k=1}^K \lambda_k^m = 1$ for each $m = 1, \ldots, M$,

(*ii*) $\nu_m = \sum_{k=1}^K \lambda_k^m \mu_k$ for each $m = 1, \ldots, M$, and

(*iii*) $\tau_k^\rho = \sum_{m=1}^M \lambda_k^m \tau_m^{\rho'}$ for each $k = 1, \ldots, K$.

[5] All information structures have to satisfy Assumption 3a by Bayes-plausibility.

To show that the agent is more responsive with a higher mean under Σ_ρ than $\Sigma_{\rho'}$, take any increasing and convex function $\psi : \mathbb{R} \to \mathbb{R}$. Then

$$\begin{aligned}
\int \psi(z) dH(z; \rho') &= \sum_{m=1}^{M} \psi\left(a^*\left(\nu_m\right)\right) \tau_m^{\rho'} \\
&= \sum_{m=1}^{M} \psi\left(a^*\left(\sum_{k=1}^{K} \lambda_k^m \mu_k\right)\right) \tau_m^{\rho'} \\
&\leq \sum_{k=1}^{K} \sum_{m=1}^{M} \lambda_k^m \psi\left(a^*\left(\mu_k\right)\right) \tau_m^{\rho'} \\
&= \sum_{k=1}^{K} \psi\left(a^*\left(\mu_k\right)\right) \tau_k^{\rho} \\
&= \int \psi(z) dH(z; \rho)
\end{aligned}$$

where the inequality follows from Lemma B.5 and the convexity of ψ. By definition of increasing convex order, the agent is more responsive with a higher mean under Σ_ρ than $\Sigma_{\rho'}$. The above argument can be extended to the case of infinite posteriors following the methods in Khan, Yu, and Zhang (2020). ■

Proof of Proposition 1. We prove the case for $\Gamma \in \boldsymbol{\Gamma_H}$; a symmetric argument proves the case for $\Gamma \in \boldsymbol{\Gamma_L}$. Define $a_\lambda = \lambda a_1 + (1-\lambda) a_2$ and $\mu_\lambda = \lambda \mu_1 + (1-\lambda)\mu_2$. By convexity of U_a in a,

$$U_a(\mu_\lambda, a_\lambda) \leq \lambda^2 U_a(\mu_1, a_1) + (1-\lambda)^2 U_a(\mu_2, a_2) + \lambda(1-\lambda)\Big(U_a(\mu_1, a_2) + U_a(\mu_2, a_1)\Big).$$

There are two cases to consider:

1. $\boldsymbol{a_1 \leq a_2}$. Since u_a has increasing differences in $(\theta; a)$, so does U_a. As $\mu_2 \succeq_{FOSD} \mu_1$, we have

$$\begin{aligned}
U_a(\mu_\lambda, a_\lambda) \leq& \lambda^2 U_a(\mu_1, a_1) + (1-\lambda)^2 U_a(\mu_2, a_2) + \lambda(1-\lambda)\Big(U_a(\mu_1, a_2) + U_a(\mu_2, a_1)\Big) \\
\leq& \lambda U_a(\mu_1, a_1) + (1-\lambda) U_a(\mu_2, a_2) \\
\leq& \max\{U_a(\mu_1, a_1), U_a(\mu_2, a_2)\},
\end{aligned}$$

where the first equality follows from $U_a(\mu_1, a_2) + U_a(\mu_2, a_1) \leq U_a(\mu_1, a_1) + U_a(\mu_2, a_2)$.

2. $\boldsymbol{a_1 > a_2}$. Since u is concave in a, so is U. Therefore, $U_a(\mu, a_1) \leq U_a(\mu, a_2)$ for any $\mu \in$

$\Delta(\Theta)$. Additionally, since u has increasing differenes in $(\theta; a)$, so does U. As $\mu_2 \succeq_{FOSD} \mu_1$, we have $U_a(\mu_1, a_2) \leq U_a(\mu_2, a_2)$. In other words, $\max\{U_a(\mu_1, a_1), U_a(\mu_2, a_2)\} = U_a(\mu_2, a_2)$. We can then conclude that

$$\begin{aligned} U_a(\mu_\lambda, a_\lambda) \leq & \lambda^2 U_a(\mu_1, a_1) + (1-\lambda)^2 U_a(\mu_2, a_2) + \lambda(1-\lambda)\Big(U_a(\mu_1, a_2) + U_a(\mu_2, a_1)\Big). \\ \leq & U_a(\mu_2, a_2). \end{aligned}$$

In both cases, we get the desired quasiconvexity condition. ■

Proof of Proposition 2. We prove the case for when $v(\theta, a)$ is component-wise convex in a and has increasing differences in $(\theta, a_{-i}; a_i)$ for all $i \in N$. The proof extends to the respective component-wise concave and decreasing differences case by setting the sender's payoff to $-v(\theta, a)$.

Take two information structures $\Sigma_\rho, \Sigma_{\rho'}$ such that, without loss of generality, $\rho_1 \succeq_{spm} \rho'_1$ and $\rho_i = \rho'_i$ for all $i \neq 1$. The sender's ex-ante payoff difference is given by

$$\begin{aligned} V(\rho) - V(\rho') = & \int\limits_{\Theta \times S} v(\theta, a_1^\star(s_1; \rho), a_{-1}^\star(s_{-1}; \rho))\big[d\boldsymbol{G}(\theta, s; \rho) - d\boldsymbol{G}(\theta, s; \rho')\big] \\ & + \int\limits_{\Theta \times S} \big[v(\theta, a_1^\star(s_1; \rho), a_{-1}^\star(s_{-1}; \rho)) - v(\theta, a_1^\star(s_1; \rho'), a_{-1}^\star(s_{-1}; \rho'))\big] d\boldsymbol{G}(\theta, s; \rho'). \end{aligned} \tag{1}$$

We consider each of these integral terms separately and show that they are both non-negative.

Using Assumption 4, we can rewrite the first integral term in (1) as

$$\int\limits_{\Theta_1 \times S_1} \underbrace{\int\limits_{\Theta_{-1} \times S_{-1}} v(\theta, a_1^\star(s_1; \rho), a_{-1}^\star(s_{-1}; \rho)) d\boldsymbol{G}(\theta_{-1}, s_{-1} | \theta_1; \rho_{-1})}_{\triangleq \psi(\theta_1, s_1)} \big[dG(\theta_1, s_1; \rho_1) - dG(\theta_1, s_1; \rho'_1)\big].$$

$\psi(\theta_1, s_1)$ has increasing differences in $(\theta_1; s_1)$ because (i) the receivers' strategies are monotone in signal realizations, (ii) $v(\theta, a)$ has increasing differences in $(\theta, a_{-1}; a_1)$, and (iii) $\boldsymbol{G}(\tilde{\theta}_{-1}, \tilde{s}_{-1} | \theta_1; \rho_{-1})$ is increasing in FOSD as θ_1 increases. Thus, by definition of the supermodular stochastic order, the first integral term of (1) is non-negative.

When $v(\theta, a)$ is differentiable[6] and component-wise convex in a, the second integral

[6]If $v(\theta, a)$ is not differentiable in a, we can uniformly approximate it by a convex analytic function.

term of (1) satisfies

$$\int\limits_{\Theta\times S} \left[v(\theta, a_1^\star(s_1;\rho), a_{-1}^\star(s_{-1};\rho)) - v(\theta, a_1^\star(s_1;\rho'), a_{-1}^\star(s_{-1};\rho'))\right] d\boldsymbol{G}(\theta, s;\rho')$$

$$\geq \sum_{i=1}^{n} \int_0^1 \left(a_i^\star(s_i;\rho) - a_i^\star(s_i;\rho')\right) \underbrace{\int\limits_{\Theta\times S_{-i}} v_{a_i}(\theta, a_i^\star(s_i;\rho'), a_{-i}^\star(s_{-i};\rho')) d\boldsymbol{G}(\theta, s_{-i}|s_i;\rho')}_{\triangleq \varphi_i(s_i)} ds_i.$$

$\varphi_i(s_i)$ is an increasing function of s_i for all $i \in N$ because, for all $i \in N$, (i) strategies are monotone in signal realizations, (ii) $v(\theta, a)$ has increasing differences in $(\theta, a_{-i}; a_i)$, (iii) $v(\theta, a)$ is convex in a_i, and (iv) $\boldsymbol{G}(\tilde{\theta}, \tilde{s}_{-i}|s_i;\rho')$ is increasing in FOSD as s_i increases by by Assumption 1, Assumption 3, and Assumption 4. We have three cases to consider:

Case I: $v(\theta, a)$ is increasing in a and $\Gamma \in \boldsymbol{\Gamma_H}$.

From Theorem 1 and Lemma A.1, $\Gamma \in \boldsymbol{\Gamma_H}$ implies that for each $i \in N$,

$$\int_t^1 [a_i^\star(s_i;\rho) - a_i^\star(s_i;\rho')] ds_i \geq 0, \forall t \in [0,1].$$

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

$$\int_0^1 [a^\star(s_i;\rho) - a_i^\star(s;\rho')]\varphi_i(s_i) ds_i$$

$$\geq \underbrace{\varphi_i(0)}_{\geq 0} \underbrace{\int_0^1 [a_i^\star(s_i;\rho) - a_i^\star(s_i;\rho')] ds_i}_{\geq 0} \geq 0.$$

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

Case II: $v(\theta, a)$ is decreasing in a and $\Gamma \in \boldsymbol{\Gamma_L}$.

From Theorem 1 and Lemma A.1, $\Gamma \in \boldsymbol{\Gamma_L}$ implies that for each $i \in N$,

$$\int_0^t [a_i^\star(s_i;\rho) - a_i^\star(s_i;\rho')] ds_i \leq 0, \forall t \in [0,1].$$

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

$$\int_0^1 [a^\star(s_i;\rho) - a_i^\star(s;\rho')]\varphi_i(s_i) ds_i$$

$$\geq \underbrace{\varphi_i(1)}_{\leq 0} \underbrace{\int_0^1 [a_i^\star(s_i;\rho) - a_i^\star(s_i;\rho')] ds_i}_{\leq 0} \geq 0.$$

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

Case III: $\Gamma \in \boldsymbol{\Gamma_H} \cap \boldsymbol{\Gamma_L}$.
From Theorem 1 and Lemma A.1, $\Gamma \in \boldsymbol{\Gamma_H} \cap \boldsymbol{\Gamma_L}$ implies that for each $i \in N$,

$$\int_t^1 [a_i^\star(s_i;\rho) - a_i^\star(s_i;\rho')]ds_i \geq 0, \forall t \in [0,1].$$

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

$$\int_0^1 [a^\star(s_i;\rho) - a_i^\star(s;\rho')]\varphi_i(s_i)ds_i$$

$$\geq \underbrace{\varphi_i(0)}_{\geq 0} \underbrace{\int_0^1 [a_i^\star(s_i;\rho) - a_i^\star(s_i;\rho')]ds_i}_{=0} = 0.$$

Hence, $V(\rho) \geq V(\rho')$. ∎

Proof of Theorem 3. Each information structure Σ_ρ induces a distribution $\tau^\rho \in \Delta\left(\Delta(\Theta)\right)$ such that the sender's ex-ante payoff is

$$V(\rho) = \int\limits_{\Delta(\Theta)} \boldsymbol{v}(\mu)\tau^\rho(d\mu),$$

where $\boldsymbol{v} : \Delta(\Theta) \to \mathbb{R}$ is the sender's interim value function given by

$$\boldsymbol{v}(\mu) = \int\limits_{\Theta} v(\theta, a^*(\mu))\mu(d\theta).$$

Take any two beliefs $\mu_1, \mu_2 \in \Delta(\Theta)$ such that $\mu_2 \succeq_{FOSD} \mu_1$. Let $a_i^* = a^*(\mu_i)$ for $i = 1, 2$. Since $u(\theta, a)$ has increasing differences in $(\theta; a)$, $a_2^* \geq a_1^*$. For any weight $\lambda \in [0,1]$, let $\mu_\lambda = \lambda\mu_1 + (1-\lambda)\mu_2$. Consider the first case: $\Gamma \in \boldsymbol{\Gamma_H}$, and $v(\theta, a)$ is increasing and convex in a and satisfies increasing differences in $(\theta; a)$. We then have

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

where the first inequality follows because $a^*(\mu_\lambda) \leq \lambda a_1^* + (1-\lambda) a_2^*$ by Lemma B.5 and $v(\theta, a)$ is increasing and convex in a. The second inequality follows from

$$\int_\Theta \Big(v(\theta, a_1^*) - v(\theta, a_2^*) \Big) \big(\mu_2(d\theta) - \mu_1(d\theta) \big) \leq 0$$

since $v(\theta, a)$ satisfies increasing differences in $(\theta; a)$, $a_2^* \geq a_1^*$, and $\mu_2 \succeq_{FOSD} \mu_1$. By a similar induction argument as Lemma B.5, this convexity property of $\boldsymbol{v}$ can be extended to any finite sequence of beliefs $\{\mu_k\}_{k=1}^K \in \Delta(\Theta)$ with $\mu_K \succeq_{FOSD} \mu_{K-1} \succeq_{FOSD} \dots \succeq_{FOSD} \mu_1$ and any weights $\{\lambda_k\}_{k=1}^K$ with $\sum_{k=1}^K \lambda_k = 1$, so that

$$\boldsymbol{v}\left(\sum_{k=1}^K \lambda_k \mu_k \right) \leq \sum_{k=1}^K \lambda_k \boldsymbol{v}(\mu_k).$$

Consider any information structure Σ_ρ that satisfies Assumption 3. In particular, suppose Σ_ρ induces posteriors $\{\mu_k\}_{k=1}^K \in \Delta(\Theta)$ with corresponding probabilities $\{\tau_k^\rho\}_{k=1}^K$ such that $\mu_{k''} \succeq_{FOSD} \mu_{k'}$ whenever $k'' > k'$. Let $\Sigma_{\rho'}$ be a garbling of Σ_ρ that induces posteriors $\{\nu_m\}_{m=1}^M$ with corresponding probabilities $\{\tau_m^{\rho'}\}_{m=1}^M$. Again, $\Sigma_{\rho'}$ does not have to satisfy Assumption 3.b-d. Since $\Sigma_{\rho'}$ is a garbling of Σ_ρ, there exist weights $\{\{\lambda_k^m\}_{k=1}^K\}_{m=1}^M \in [0,1]^{K \times M}$ such that

(*i*) $\sum_{k=1}^K \lambda_k^m = 1$ for each $m = 1, \dots, M$,

(*ii*) $\nu_m = \sum_{k=1}^K \lambda_k^m \mu_k$ for each $m = 1, \dots, M$, and

(*iii*) $\tau_k^\rho = \sum_{m=1}^M \lambda_k^m \tau_m^{\rho'}$ for each $k = 1, \dots, K$.

Thus,

$$\begin{aligned} V(\rho') &= \sum_{m=1}^M \boldsymbol{v}(\nu_m) \tau_m^{\rho'} \\ &= \sum_{m=1}^M \boldsymbol{v}\left(\sum_{k=1}^K \lambda_k^m \mu_k \right) \tau_m^{\rho'} \\ &\leq \sum_{k=1}^K \sum_{m=1}^M \lambda_k^m \boldsymbol{v}\left(\mu_k\right) \tau_m^{\rho'} \\ &= \sum_{k=1}^K \boldsymbol{v}\left(\mu_k\right) \tau_k^\rho \\ &= V(\rho). \end{aligned}$$

The above argument can further be extended to the case of infinite posteriors following the methods in Khan, Yu, and Zhang (2020).

Therefore, if Σ_ρ dominates $\Sigma_{\rho'}$ in the Blackwell order and Σ_ρ satisfies Assumption 3, $V(\rho) \geq V(\rho')$. As $\Theta \subset \mathbb{R}$ is a totally ordered set when $n = 1$, the full information structure always satisfies Assumption 3 and always Blackwell dominates any other information structure, which concludes the proof. The remaining cases are analogous. ■

Appendix C. When Responsiveness Fails

In this section, we explore why a higher quality of information may not lead to more dispersed optimal actions in a single-agent setting when an agent's optimal action $a^*(\mu)$ is neither convex nor concave as in B.5, which necessarily implies $\Gamma \notin \boldsymbol{\Gamma_H} \cup \boldsymbol{\Gamma_L}$. We even assume that $a^*(\mu)$ is monotone over beliefs ordered by FOSD (so the agent's utility function may satisfy Assumption 2).

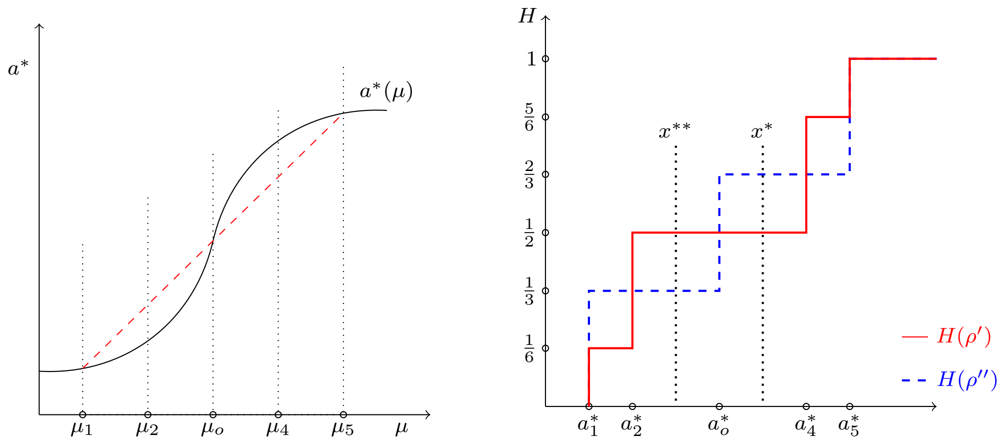
Consider a simple binary-states setting in which the agent's prior places mass on only two points $\{\underline{\theta}, \bar{\theta}\} \subset \Theta$ with $\bar{\theta} > \underline{\theta}$. Let $\mu = \mathbb{P}(\tilde{\theta} = \bar{\theta}) \in [0, 1]$ represent some posterior belief the agent holds. For some $\delta \in (0, 1/5)$, let the prior be $\mu_o = 3\delta$ and consider beliefs $\{\mu_n\}_{n=1,2,4,5}$ such that $\mu_n = n\delta$. Beliefs are ordered by first-order stochastic dominance with $\mu_5 \succeq_{FOSD} \mu_4 \succeq_{FOSD} \mu_o \succeq_{FOSD} \mu_2 \succeq_{FOSD} \mu_1$. Finally, let $a^*(\mu)$ be monotone but neither convex nor concave as in Figure 1a.

Let $\Sigma_{\rho'}$ induce posteriors $\{\mu_1, \mu_2, \mu_4, \mu_5\}$ with probability $\{1/6, 1/3, 1/3, 1/6\}$. Let $\Sigma_{\rho''}$ be an information structure that induces three posteriors $\{\mu_1, \mu_o, \mu_5\}$ with probabilities $\{1/3, 1/3, 1/3\}$ Notice that $\Sigma_{\rho'}$ is a equivalent to getting information from $\Sigma_{\rho''}$ with probability 0.5 and no information with probability 0.5. Thus, $\Sigma_{\rho''}$ is Blackwell more informative than $\Sigma_{\rho'}$, which implies $\rho'' \succeq_{spm} \rho'$.

Assume that the average action under $\Sigma_{\rho''}$ and $\Sigma_{\rho'}$ equal $a^*(\mu_o)$. In Figure 1a, this corresponds to the point of intersection of the dashed line and the solid curved line at μ_o. Figure 1b maps the distribution over optimal actions: $\Sigma_{\rho''}$ induces the dashed line while $\Sigma_{\rho'}$ induces the solid line.

If we start integrating from the right, then $\int_x^\infty H(z;\rho'') - H(z;\rho')dz \leq 0$ for all $x > a_4^*$ but the sign changes at some point $x^* \in (a_o^*, a_4^*)$. Thus, the agent is not more responsive with a higher mean under $\Sigma_{\rho''}$. If we instead integrate from the left, then $\int_{-\infty}^x H(z;\rho'') - H(z;\rho')dz \geq 0$ for all $x < a_2^*$ but the sign changes at some point $x^{**} \in (a_2^*, a_o)$. Thus, the agent is not more responsive with a lower mean under $\Sigma_{\rho''}$.

In fact, as the average action under $\Sigma_{\rho''}$ equals the average action under $\Sigma_{\rho'}$, we can conclude that the distributions of actions $H(\rho'')$ and $H(\rho')$ cannot be ordered by most



(a) Optimal action (b) Induced distribution

Figure 1: Non-convexity/concavity and non-responsiveness

univariate stochastic variability orders such as second-order stochastic dominance, mean-preserving spreads, Lorenz order, dilation order, and dispersive order.[7]

References

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

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

Susan Athey. 2002. Monotone Comparative Statics under Uncertainty. *The Quarterly Journal of Economics* 117, 1 (2002), 187–223.

Susan Athey and Jonathan Levin. 2017. The Value of Information in Monotone Decision Problems. *Research in Economics* (2017).

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

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

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

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

David Blackwell. 1953. Equivalent Comparisons of Experiments. *Ann. Math. Statist.* 24, 2 (06 1953), 265–272.

Yeon-Koo Che, Jinwoo Kim, and Fuhito Kojima. 2019. Weak monotone comparative statics. *Available at SSRN 3486365* (2019).

Piotr Dworczak and Giorgio Martini. 2019. The simple economics of optimal persuasion. *Journal of Political Economy* 127, 5 (2019), 1993–2048.

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

Olivier Gossner. 2000. Comparison of Information Structures. *Games and Economic Behavior* 30 (2000), 44–63.

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

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

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

M Ali Khan, Haomiao Yu, and Zhixiang Zhang. 2020. A Missing Equivalence Theorem on Information Structures on a General State Space. (2020).

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

E. L. Lehmann. 1988. Comparing Location Experiments. *Ann. Statist.* 16, 2 (06 1988), 521–533.

Laurent Mathevet, Jacopo Perego, and Ina Taneva. 2020. On information design in games. *Journal of Political Economy* 128, 4 (2020), 1370–1404.

Jeffrey Mensch. 2019. Monotone persuasion. *Available at SSRN 3265980* (2019).

Margaret Meyer and Bruno Strulovici. 2012. Increasing interdependence of multivariate distributions. *Journal of Economic Theory* 147, 4 (2012), 1460–1489.

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

Paul Milgrom and Chris Shannon. 1994. Monotone Comparative Statics. *Econometrica* 62, 1 (1994), 157–180.

Stephen Morris and Hyun Song Shin. 2002. Social Value of Public Information. *The American Economic Review* 92, 5 (2002), 1521–1534.

Alfred Müller and Marco Scarsini. 2006. Stochastic order relations and lattices of probability measures. *SIAM Journal on Optimization* 16, 4 (2006), 1024–1043.

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

John K.-H. Quah and Bruno Strulovici. 2009. Comparative Statics, Informativeness, and the Interval Dominance Order. *Econometrica* 77, 6 (2009), 1949–1992.

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

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

Nicolas Roux and Joel Sobel. 2015. Group polarization in a model of information aggregation. *American Economic Journal: Microeconomics* 7, 4 (2015), 202–32.

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

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

Ina Taneva. 2017. Information Design. *working paper* (2017).

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

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

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

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